Global Campus
Europe

Contemporary Threats to Human Rights in the Online Public Sphere

The case of Facebook

María José Segura

Supervisor: Rikke Frank Jørgensen

Acknowledgments

This thesis would not have been possible without the support of different people. I therefore take this opportunity to thank everyone who has been there for me in one way or another.

Foremost I want to thank my supervisor Dr. Rikke Frank Jørgensen, whose work and guidance inspired me to write this thesis. I am thankful for her vital feedback, asking insightful questions and offering invaluable advice.

I am also grateful to Dr. Eva Maria Lassen, E.MA Director at the Danish Institute for Human Rights, for her kindness and encouragement throughout this process.

Furthermore, I want to express my sincere gratitude to everyone else in the Research Department of the Danish Institute for Human Rights, for welcoming me in the team and teaching me a great deal about human rights research.

Lastly, I must thank my loving family and friends, particularly my parents and boyfriend, for their unending patience and support.

Table of Contents

Abstract

This thesis examines how the current use of social media platforms as forums for public discussions impacts human rights. The discussion makes use of Facebook as an example of a major platform that is globally used for public debate.

The thesis builds on the concept of the public sphere, developed by Habermas, to look at the nature of today's public forums in the online realm. It finds that the normative public sphere is subject to a dramatic change and it claims that this shift is rooted in the overlapping of the private and the public. Two main elements are challenged in social media platforms. In the first place, the public sphere is based on civic engagement and participation but content moderation policies by platforms owners are disrupting this idea. Secondly, Habermas envisioned a public sphere that is free from commercial interests but this ideal is being corrupted by the platforms' profit-driven goals. This thesis indicates that such a significant transformation of the public sphere is having, among others, a negative impact on the human rights associated to public life. Freedom of expression and the right to privacy, respectively.

An analysis of how social media like Facebook are taking on the duty of gatekeepers of public speech, effectively shaping user participation and the public discourse, illustrates the profound effects of platforms' policies on freedom of expression. Likewise, a study of the online business model of these corporations demonstrates the threat that this business model poses to users' right to privacy. The thesis concludes with the concern that users exercise their human rights and freedoms within privately owned platforms that operate outside the human rights protection framework. It looks at existing efforts, including regulation, and examines ways of overcoming the governance gaps that prevent human rights from being protected in the online realm of social media platforms.

Keywords: *public sphere, Facebook, right to privacy, freedom of expression, content moderation, online business model*

Abbreviations

CSR	Corporate Social Responsibility
GDPR	General Data Protection Regulation
GNI	Global Network Initiative
ICT	Information and Communication Technologies
OECD	Organization for Economic Co-operation and Development
RDR	Ranking Digital Rights
SNS	Social Network Sites / Social Network Services
SPS	Sponsored Stories
UGC	User Generated Content
UN	United Nations
UNGC	United Nations Global Compact
UNGPs	United Nations Guiding Principles on Business and Human Rights

List of Figures

List of Tables

Introduction

The UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression opened his last report with a quote from 1996 by John Perry Barlow, in which he stated that "the Internet would usher in a world where anyone, anywhere, may express his or her beliefs, no matter how singular, without fear of being coerced into silence and conformity[1]" (UN Human Rights Council , 2018). Kaye then declares that such online evangelism is hard to find in 2018 (Ibid). Indeed, social media platforms like Facebook serve as the indisputable global forums for public discussions but they seem to differ from the romanticized idea of the public sphere that Habermas envisioned[2]. The online public sphere is found within a commercial domain, where the host's policies and practices are mainly business driven. Because individuals rely on social media to exercise core civil and political rights, platforms hold immense powers over the boundaries and conditions of these rights. Yet, as private actors, the mere question of platform owners being liable for human rights abuses appears as a quixotic pursuit.

Due to recent scandals and increased public scrutiny, this discussion is gaining grounds beyond the technological and sociological field. However, there is scarce literature on the implications to human rights of social media serving as the public sphere. This way, this thesis expands the existing analysis of social media platforms, specifically Facebook[3], by drawing from communications theory and making use of a human rights approach. The goal is to flag some of the impacts on human rights that social media have when performing as hosts of public discussions. As platforms grow in size and scale, this issue becomes more relevant and the conversation on governance of human rights protection online becomes more critical.

All in all, this thesis aims at analyzing the nature of today's public sphere in social media platforms and its impacts on human rights, as well as the regulatory frameworks they abide by. More precisely, it looks into the right to privacy, enshrined in article 12 of the Universal Declaration of Human Rights

---

[1] This is John Perry Barlow, A Declaration of The Independence of Cyberspace, 8 February 1996.

[2] Jurgen Habermas conceptualized the public sphere in his *work The Structural Transformation of the Public Sphere: an Inquiry into a Category of Bourgeois Society* in 1989.

[3] This thesis focuses narrowly on Facebook but is intended to have wider significance as the analysis can be applied to similar social media platforms.

(UDHR)[4] and article 17 of the International Covenant on Civil and Political Rights (ICCPR); and the right to freedom of expression recognized in articles 19[5] of both of the mentioned instruments.

As shown in the following table, this thesis is divided in three chapters: Chapter 1 reviews the concept of the public sphere, as defined by Habermas, and introduces its key elements to be examined in the online public sphere. Chapter 2 analyzes the features of Facebook that may be in conflict with the normative idea of the public sphere and give raise to human rights concerns. Finally, Chapter 3 investigates these concerns and evaluates the relevant regulatory human rights framework. Throughout the thesis, special attention is given to the distinction of *private* and *public*, which is presented as root of the matters at hand.

Table 1: Outline of the thesis

Chapter 1 Public sphere - some key elements	Chapter 2 Key features of Facebook's platform	Chapter 3 Affected human rights
Free from commercial interests →	Online business model →	Right to privacy
Civic engagement and public participation →	Content moderation →	Freedom of expression

---

[4] Article 12 of the Universal Declaration of Human Rights: *No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.* Retrieved rom: https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf

[5] Article 19 of the Universal Declaration of Human Rights: *Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.* Retrieved from: https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf

Chapter 1 - The Public Sphere

This chapter is focused on the concept of the public sphere, hosted now by social media platforms. A holistic approach of the term is presented and questions are raised on the paradigm shift from a traditional public sphere to a new one. First of all, a definition is given, along with a historical background. Secondly, two main elements of the public sphere are considered: 1) civic engagement in the public sphere and 2) the freedom from commercial interests. The rise of a new public sphere that is digital and global by nature is then explored. The last section looks at the distinction between *private* and *public* and presents a comprehensive theoretical model to be used in the analysis of social media platforms in the following chapter. All in all, the aim of this chapter is to look at the public sphere, in a normative sense, and its contemporary digital form. Throughout the chapter, the shortcomings of platforms providing the space for public discussions are exposed.

1.1 Habermas' public sphere

The public sphere was first conceptualized by Jurgen Habermas in his *work The Structural Transformation of the Public Sphere: an Inquiry into a Category of Bourgeois Society (1989)*[6]. It represents a domain of social life in which public opinion is expressed by means of rational public discourse and debate (Papacharissi, A Private Sphere: Democracy in a Digital Age, 2010).

Descriptively, the aim of the public sphere is public accord and decision making but, from a normative standpoint, beyond a place for opinion formation, it is a space for civil society deliberations on how these opinions can influence the actions of the state (Habermas, The structural transformation of the public sphere : an inquiry into a category of bourgeois society, 2006). As noted by Jørgensen, public discourse is the articulation of the citizens' opinions and priorities but also, of the criticism with respect to the democratic rule of a country (Jørgensen, The Internet as public sphere, 2013). In the words of Calhoun, it helps understand the conditions for effective leverage by citizens on the state and other powerful institutions and, as such, the public sphere is closely linked to fundamental rights and freedoms (Calhoun, Civil Society and the Public Sphere, 1993). For this reason, and because the concept of the public sphere presupposes the participation of the public in decision making, it is foundationally important for democratic theory. It is also at the core of communication studies,

---

[6]The first edition of this book was published in 1962 in Germany under the tittle *Strukturwandel der Öffentlichkeit; Untersuchungen zu einer Kategorie der Bürgerlichen Gesellschaft*. It became widely popular in 1998, when it was translated in English.

specially political and cultural communication, and it sets the agenda for the debate on public discourse.

Although intertwined, the concept of the *public sphere* must not be mixed up with the *public space* and *public opinion*. The public sphere, as described above, provides a metaphor of a sphere which mediates between society and state, in which the public organizes itself as the bearer for public opinion (Habermas, Lennox, & Lennox, 1974). Habermas argues that public opinion can only come into existence when a reasoning of the public is presupposed. Thus, rational deliberation distinguishes public opinion from mere opinions made public or opinions related to public affairs. Finally, the public space serves as the expanse for the public sphere to convene.

While traditionally the public space was found in physical spaces such as cafés; contemporarily, information and communication technologies (ICTs) are removing all the temporal and spatial barriers to distanced communication (Khan, Gilani, & Nawaz, 2012). As a *public space*, the Internet provides another forum for political deliberation. However, as a *public sphere*, the Internet could facilitate discussion that promotes a democratic exchange of ideas and opinions (Papacharissi, The virtual sphere: the internet as a public sphere, 2002). It is then argued that a new public sphere is emerging in the virtual realm[7].

Habermas traces the liberal public sphere to the bourgeois political life of the 18th and 19th centuries. This period was characterized by Enlightenment values of equality, freedom and justice that came to replace feudal values of hierarchy, tradition and respect for authority (Mackee2005). The struggle against the feudal state in Europe, led to gatherings of private individuals in public spaces, where they would discuss public affairs and organize against the power of the state. It was the bourgeoisie that gradually replaced a public sphere in which the ruler's power was merely represented *before* the people with a sphere in which state authority was publicly monitored through informed and critical discourse *by* the people (McCarthy, 1991). Although advocating for the "basic rights of man", Habermas notes that the liberal public sphere was actually restricted to a certain class of men. A market economy was developing and Europe was seeing the maturation of capitalism. The free market, he argues, led to the inclusion in the public sphere of disadvantage workers who asked for state regulations. He additionally notes that it was the consequent intertwining of state and society

---

[7] Papacharissi claims that a digital medium like the Internet, with an infrastructure that promises unlimited and unregulated discourse that operates beyond geographic boundaries, would suggest a virtual reincarnation of the public sphere. Retrieved from: Papacharissi, *A Private Sphere: Democracy in a Digital Age*, 2010.

what put an end to the liberal public sphere. <u>The public sphere of social-welfare-state democracies</u> is rather a field of competition among conflicting interests, in which organizations representing diverse constituencies negotiate and compromise among themselves and with government officials, while excluding the public from their proceedings (McCarthy, 1991). All in all, Habermas develops a normative ideal of rational public discussion based on the bourgeois public sphere that, he argues, is now corrupted by capitalist forces and, thus, no longer viable.

Why this leap in history if the issue at hand is social media? As it is stated in the Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression: the Internet has become the central global public forum with profound value for human rights (UN Human Rights Council, 2015). Social media platforms are hosting today's public sphere, where individuals engage in public discussions. Yet, these are (legally) private actors taking on functions that were traditionally exercise in public spaces. Certainly, this shift has direct implications for the rights and freedoms associated with public life. But, before analyzing human rights impacts, we must look into the elements that make up the public sphere, as defined by Habermas, to signal the issues that are raising concerns. Namely, civic engagement and freedom from commercial interests.

How is the new public sphere different from the one that Habermas envisioned?

A. *Civic engagement and public participation*

As noted by Castells, the public sphere is an essential component of sociopolitical organization because it is the space where people come together as citizens and articulate their autonomous views to influence the political institutions of society. He then clarifies that the organized expression of these views conforms civil society (Castells, 2008)[8]. Despite being central to democratic practice, as Calhoun claims, a public sphere adequate to a democratic polity depends upon both quality of the discourse and quantity of participation (Calhoun, Habermas and the Public Sphere, 1992). Not only should discussion be constituted around rational critical argument, but the more people participate as citizens in politics, the closer one comes to the ideal of a public sphere (Grbesa, 2003). Thus, two

---

[8] Habermas too linked the notion of the *public sphere* with that of *civil society* in the sense that "civil society embraces a multiplicity of ostensible "private" yet potentially autonomous public arenas distinct from the state." Civil society ideally distils and transmits reactions to societal problems from the private sphere to the public sphere. Retrieved from: Jørgensen, *The Internet as public sphere*, 2013.

preconditions to an effective public sphere can be distinguished: 1) access to the public space and 2) a vibrant participation by civil society.

Dahlberg indicates that the public sphere model aims for all-inclusive, consensus seeking rational deliberation (Dahlberg, 2007). Inclusivity can only be achieved by granting access to all individuals and guarantying that asymmetries of power do not interfere in the deliberation process. But access does not ensure participation. Only ideally, all individuals engage in public discussion, which in turn gives place to an effective civil society that channels citizens' views and influences political institutions.

While an inclusive, consensus-seeking public sphere is a desirable goal; the more individuals engaging in public discussion, the harder the consensus seeking process gets. In Habermas' conception of the bourgeois public sphere, the merits of arguments were much more important than the identities of arguers. It was the best rational argument and not the identity of the speaker that mattered (Calhoun, Habermas and the Public Sphere, 1992). In fact, transformations on the public sphere are ascribed to its continual expansion to include more participants (Ibid). Habermas defends that inclusivity brought degeneration of the quality of the discourse. That is to say that an inclusive democratic public sphere would be imperatively detrimental to its quality. It is on these points that Habermas' theory has met strong criticism. The exclusionary character of the early public sphere, including only small numbers of elite men, was based on the assumption that transformation by "massification" will result in a negative outcome. The space was restricted not only to the upper class, but also just to men. Scholars like Benhabib, Fraser, Eley, and Ryan, argue that exclusion on gender grounds is not reasonable because literate women would not have transformed the bourgeois sphere into a mass (Ibid). In the same line, Grbesa notes that, although discussions among citizens of different social groups may not reach consensus, the inclusion of diversity of opinions and views is of greater importance (Grbesa, 2003). Hence, Habermas has been criticized for his idealistic vision of the 18th and 19th century public life, whose features are unacceptable in modern democracies. However, it is important to evaluate concepts in their historical context. In the early public sphere, for the first time, practical reason was institutionalized through norms of reasoned discourse in which arguments, not statuses or traditions, were to be decisive (Calhoun, Habermas and the Public Sphere, 1992).

Regarding social media, issues of inclusivity replicate offline restrictions. Shifting public discussions to social media excludes those with no access to these platforms or to the Internet, typically

developing countries and countries with Internet restrictions. Moreover, connectivity alone does not guarantee a more representative and vigorous public sphere (Papacharissi, The virtual sphere: the internet as a public sphere, 2002). The same way, it does not ensure rational deliberation and enlightened political discourse, nor that people from different backgrounds will be more understanding of each other (Khan, Gilani, & Nawaz, 2012). Similarly, online political conversations can be as easily dominated by elites as offline ones (Papacharissi, A Private Sphere: Democracy in a Digital Age, 2010).

It is argued that social media platforms facilitate civic engagement in political life. Yet, as host of public discussions, they also act as curators of speech. Doing so they effectively shape public discourse, for which their role as hosts of public debate must be questioned. Social media corporations could actually be undermining civic engagement and, arguably, constitute a deficient public sphere. What is more, by acting as such, platforms owners may have bigger responsibilities towards society than providing a profitable service.

B. Free from commercial interests

Given that an informed electorate is the basis for rational deliberation and, as such, a premise of any meaningful democracy, the media is central to the public sphere. According to Shirky, political freedom has to be accompanied by a civil society literate enough and densely connected enough to discuss the issues presented to the public (Shirky, The Political Power of Social Media: Technology, the Public Sphere, and Political Change, 2011). Interestingly, Sociologists Elihu Katz and Paul Lazarsfeld sustain that mass media alone do not change people's minds. Instead, there is a two-step process: 1) opinions are transmitted by the media, and 2) they are echoed by friends and family member[9] (Ibid). It is through the socialization of informed individuals that public opinion is formed. In fact, access to information is far less important, politically, than access to conversation (Ibid). For this reason, social media become especially relevant in today's public sphere: they provide the space for both, media and conversation and they do so, at a global level[10].

---

[9] Castells pointed in the same direction by claiming that "the will of the people emerges from people's minds. And people make up their minds on the issues that affect their lives from the messages and debates that take place in the public sphere". Retrieved from: Castells, *The New Public Sphere: Global Society, Communication Networks, and Global Governance*, 2008.

[10] Shirky claims that social media has the long-term potential of strengthening civil society and the public sphere. In contrast to the instrumental view of Internet freedom, Shirky is a supporter of the environmental view or the belief that positive changes in the life of a country, follow, rather than precede the development of a strong public sphere. Retrieved from: Shirky. *The Political Power of Social Media: Technology, the Public Sphere, and Political Change*, 2011.

One may assume that by promoting media consumption and production, social media strengthens civil society and public discussions. However, it is Habermas' argument that the commercialized mass media have turned the public sphere into a space where the rhetoric and objectives of public relations and advertising are prioritized (Papacharissi, A Private Sphere: Democracy in a Digital Age, 2010)[11]. Likewise, Carey believes that capitalist forces have damaged the public sphere by transforming it into a mass commercial culture[12]. In a liberal reaction to autocracy, and regarded as separate from the state, the public sphere might have fallen into the opposite sphere: that of "the private" and the market. Thus, along with the freedom of publicly discuss political issues, and fostered by capitalism, the public sphere started to include commercial freedoms.

The commercialization of the media is not new to the cyberspace. The radio and television also built up hope for the strengthening of democracy but ultimately proved to be corrupted by market forces. For this reason, Papacharissi reminds us that despite all the hype surrounding the innovative uses of the Internet as a public medium, it is still a medium constructed in a capitalist era and, as such, it is susceptible to the same forces (Papacharissi, The virtual sphere: the internet as a public sphere, 2002). In a culture based on commercial values, advertising revenue has more impact on programming than democratic ideals (Ibid). Habermas himself also doubts the democratizing potential of the Internet, as he saw the Internet developing in a commercial direction, with a political orientation that was largely circumstantial (Papacharissi, A Private Sphere: Democracy in a Digital Age, 2010). Likewise, Dalhberg insists that the growth of economic interests into areas of online life is resulting in the displacement of rational deliberation by instrumental rationality (Dahlberg, 2007).

Social media platforms consist of private actors and naturally, their operations are business driven. However, their free of charge services raise concerns on the methods by which platforms like Facebook make profit. As forums for the deliberation of the public, their commercial interests may have serious implications for the quality and neutrality of public discussions. Beyond developing a public forum largely influenced by advertisement, these private players are profiting from the

---

[11] He traces the shift in the media business all the way to the end of the 18th century literary journalism where newspapers acquired a commercial basis. The press was an intensifier of public discussion, no longer a vendor of recent news but a dealer in public opinion. In the transition from the literary journalism of private individuals to the public services of the mass media the public sphere was transformed by the influx of private interests, which received special prominence in the mass media. See: (Habermas, Lennox, & Lennox, The Public Sphere: An Encyclopedia Article (1964), 1974).

[12] This is Carey (1995) in Papacharissi, *The virtual sphere: the internet as a public sphere*, 2002.

commercialization of personal data and user generated content (UGC). This way, they hold special powers towards public discussions that may also infringe individuals' rights and freedoms.

Habermas' normative idea of the public sphere does not appear to be realized in its contemporary digital form. However, before making any blunt statements, the nature and scope of today's public sphere should be looked into. Can we speak of a unique public sphere and, most importantly, is it static or rather subject to change?

1.2 The new public sphere: global and digital

ICTs are removing spatial barriers to distanced communication and opening a discussion on the boundaries and scope of the public sphere. To tackle this question, an initial examination of the rationale of the public sphere must be made: if its aim is to influence the state, then the expansion of the public sphere will be limited to state borders. Yet, public interests today may no longer be confined to national institutions. The process of globalization has shifted the debate from the national domain to the global debate, prompting the emergence of a global civil society and of ad hoc forms of global governance (Castells, 2008). Social global interests are found in multinational businesses, world religions, cultural creators, or public intellectuals, among others[13]. Contemporary challenges such as climate change, terrorism or the current migration crisis are also international by nature and hence, they affect individuals globally and transcend national political institutions. The management of interests affecting residents of more than one state is referred to as global governance (Borock, 2017).

Castells is one of the most prominent scholars behind global governance and a global civil society. He considers that not everything or everyone is globalized, but the global networks that structure the planet affect everything and everyone (Castells, 2008). If current economic, communicative, and cultural activities are globalized, public interests will then be globalized too. As follows, it is not absurd to speak about a "global public sphere". In fact, it is widely recognized by *cosmopolitanism* or the idea that all human beings are citizens of a single community. One of the most common objections to it is that, without a "world state", cosmopolitanism is nothing more than a metaphor for a way of life or normative ideal to strive for. It can be argued that social media platforms attempt to fill that void. In fact, Facebook claims that its aim is to "build a global community that works for all

---

[13] This is Beck, "*Power in the Global Age*", retrieved from: Castells, *The New Public Sphere: Global Society, Communication Networks, and Global Governance*, 2008.

of us", a platform "where humanity comes together not just as citizens or nations but as a global community"[14].

Indeed, given that globalization is dominating the agenda, public debate has to acquire a new vision and a wider spectrum where, beyond national interests, there is room for international concerns. Without a flourishing international public sphere, the global sociopolitical order becomes defined by the realpolitik of nation-states that cling to the illusion of sovereignty despite the realities wrought by globalization[15]. Thus, while globalization provides the rationale for a global public sphere, social media platforms provide the means for it. By enacting a global, horizontal network of communication, social network sites (SNS) serve as both an organizing tool and a means for debate, dialogue, and collective decision making (Castells, 2008).

All in all, we are now witnessing a shift from a national public sphere to a communication-based public sphere that transcends territorial limitations. Castells explains this transformation by pointing to the Internet and horizontal networks of communication. He claims that the current media system is local and global at the same time: "people make up their minds on the issues that affect their lives, as well as the future of humankind, from the messages and debates that take place in the public sphere" (Ibid).

Discussions taking place in social media like Facebook develop around the demands of the global civil society and hence, are quite different to those previous to the existence of global networks. Therefore, rather than replacing the traditional public sphere, it can be said that a new (global) public sphere has emerged in the online realm. Linking it to the notion of governance, Castells names it "public diplomacy" or the projection in the international arena of the values and ideas of the public. In this sense, public diplomacy intervenes in the global space equivalent to what has been traditionally conceived as the public sphere in the national system. (Ibid)

Speaking normatively, Papacharissi makes an important point: she clarifies that a new public *space* is not synonymous of a new public *sphere*. A virtual space, she argues, enhances discussion; a virtual sphere enhances democracy (Papacharissi, The virtual sphere: the internet as a public sphere, 2002). She then questions the democratization potential of the Internet and analyzes whether it amounts to a

---

[14] CEO of Facebook, Mark Zuckerberg's speech on *Building a Global Community*, Feb 1th, 2017. Retrieved from: https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634/
[15] This is Held, 2004, referenced in Castells, *The New Public Sphere: Global Society, Communication Networks, and Global Governance*, 2008, page 80.

virtual reincarnation of the public sphere or, on the other hand, it is far from the romanticized ideal envisioned by Habermas. As analyzed previously, and agreeing with Papacharissi's conclusion, the Internet and related technologies do provide a public space for political discussion but, even with greater participation, a healthier democracy cannot be inferred. The content, diversity, and impact of public discussion need to be considered carefully. (Ibid)

1.3 The *public* and the *private*

It is my claim that the differences between Habermas' public sphere and its main contemporary form, found in social media, are rooted in the overlap of the *public* and the *private*, which are social constructs that conceptualize different domains of life.

According to Habermas, in its ideal form, the public sphere is made up of private people gathered together as a public and articulating the needs of society with the state (Habermas, The Structural Transformation of the Public Sphere: An Inquiry into a category of Burgeois Society, 1991). Indeed, the notion of "the public" is closely tied to democratic ideals that call for citizen participation in public affairs (Papacharissi, The virtual sphere: the internet as a public sphere, 2002). Because of their private nature, social media platforms challenge Habermas' normative idea. The distinction between *public* and *private* stands out as one of the grand dichotomies of Western thought[16]. Yet, it is somewhat troublesome for the contemporary form of the public sphere in the online domain.

Weintraub stresses that any notion of *public* and *private* makes sense only as on element in a paired opposition. He notes that these terms are generally reduced to two fundamental, and analytically quite distinct, kinds of imagery. Thus, he presents two criteria in which public/private are often discussed. The first one refers to *visibility*: what is hidden versus what is what is open, revealed or accessible. Secondly, he introduces the *collectivity* sense: what is individual or pertains only to the individual versus what is collective or affects the interests of a collectively of individuals (Weintraub, 1997). Other scholars have discussed and enlarged this distinction to apply it to a given approach: an economical, feminist, political, or sociological one. In addition to the ones suggested by Weintraub, a third sense of private/public that eases the examination of social media and helps explain contemporary issues in this realm is put forward in this thesis. This thesis proposes that may be framed

---

[16] This is a quote of Norberto Bobbio, "*The Grand Dichotomy: Public/Private*". The binary opposition of private/public is used to subsume a wide range of other important distinctions and that attempts to dichotomize the social universe in a comprehensive and sharply demarcated way (Weintraub, 1997).

as *governance* representing motivated by commercial interests versus what is driven by individual rights and freedoms.

This way, expanding on Weintraub's work, the following criteria that will serve as the basis for the analysis of today's public sphere as provided by social media platforms is advanced.

Table 2 Senses in which the public/private distinction is invoked.

Criteria / senses	**Public**	**Private**
Visibility	Accessible	Inaccessible
Collectivity	Society (collectivity of individuals)	Individuals
Governance	Freedoms and rights	Commercial interests

Visibility

This first criterion is the most evident one. It stands for "public" as that which is visible, known or observed, in contrast to what is unknown and unobserved. Sociologically, it is employed when discussing "privacy" as things that we are able and/or entitled to keep hidden, sheltered, or withdrawn from others (Weintraub, 1997). For instance, Erving Goffman, renowned sociologist, sees all activity carried out in the presence of others as "behaviour in public spaces" while the domain of the "private" is restricted to the "backstage", where individuals prepare to enact their roles in social interaction[17]. The private realm corresponds to the personal life[18]. It is the intimate domain of family, friendship and emotionality. In comparison, "public" connotes ideas of citizenship, commonality, and things not private, but accessible and observable by all (Papacharissi, The virtual sphere: the internet as a public sphere, 2002).

Collectivity

Under the criterion of collectivity, the private equates to what concerns the individual and, in contrast, the public stands for what affects the collective or group of individuals. This distinction is partly drawn from notions of intimacy and anonymity. Dewey specifies that the private is the domain of self-defined interests and values, while the public is the domain of the shared interests and values

---

[17] Retrieved from Weintraub (1997). See Erving Goffman, *Behaviour in Public Places: Notes on the Social Organization of Gatherings (1963).*

[18] Gavison also makes this distinction: "the private is that which is unknown and unobserved and the public is that which is known or observed" (Gavison, 1992).

(Dewey, 1954). Indeed, when talking about a "private interest" it is clear that the individual is pursuing a "special interest" that concerns him, rather than a public interest, which pertains the collective of individuals. On this point, Weintraub clarifies that seeking a "private interest" does not mean the action is undertaken in secret (visibility criterion). Instead, the private is the particular (collectivity criterion) (Weintraub, 1997). Another application of the *collectivity* criterion is the term "public good" as used by economists to refer to an indivisible collective benefit. Likewise, when "public" describes the actions and agents of the state (public/private = state/nonstate), it does so based on the state's responsibility to look after the general interests and affairs of a politically organized collectivity, as opposed to "private" interests (Ibid).

It is important to note that this distinction is not equivalent to individual versus social. Dewey highlights that many private acts are social, in the sense that their consequences contribute to the welfare of society. (Dewey & Rogers, 2012). For instance, an individual may serve others even the community at large, in carrying a private business[19]. Thus, the line between private and public cannot be drawn based on consequentialism. This is where the third added criterion becomes relevant. Even though, the individual might be contributing to society through a private business, the sole driving factor is private interest, which can be equated to commercial interests.

Governance

It would seem there is a gap to be filled in the private/public distinction. The visibility criterion points out the obvious aspects of the public/private spheres, what is accessible and what is inaccessible, and the collective criterion refers to the interests and concerns of individuals or the collective. But what is the motor of the private/public realms? A third criterion can be added to expose what is it that guides the private domain, in contrast to the public. The former one is framed by commercial interests and, conversely, the latter rests upon legally guaranteed rights and freedoms. Each domain is governed by fundamentally distinct causes. Accordingly, when individuals operate in the sphere of the "public" they have certain expectations and behaviors, shaped by the rights and freedoms guaranteed in this realm. On the other hand, when individuals act in the "private", they are subjected to capitalist forces and commercial interests, governed by the free market, "the law of the jungle".

---

[19] Dewey and Rogers make an interesting link with Adam Smith's ideas on the benefits of the free market: "to some extent it is true, as Adam Smith asserted, that our breakfast table is better supplied by the convergent outcome of activities of farmers, grocers, and butchers carrying on private affairs with a view to private profit than it would be if we were served on a basis of philanthropy or public spirit." (Dewey & Rogers, 2012, s. 47).

As Weintraub accurately reveals, the previous concepts of "private" and "public" overlap giving place to mixed constructs. For instance, if market exchange is considered a "private" act – on the grounds of being, in principle, self-interested, nongovernmental and unconcerned with collective outcomes – then it does not cease to be private when it is carried out "in public" (Weintraub, 1997). And likewise, voting in an election does not necessarily cease to be a "public" act if it is carried out "in private" by secret ballot (Ibid). This overlapping is central to new media like Facebook's platform, which is developing as a public space for civic engagement, in despite of its private ownership.
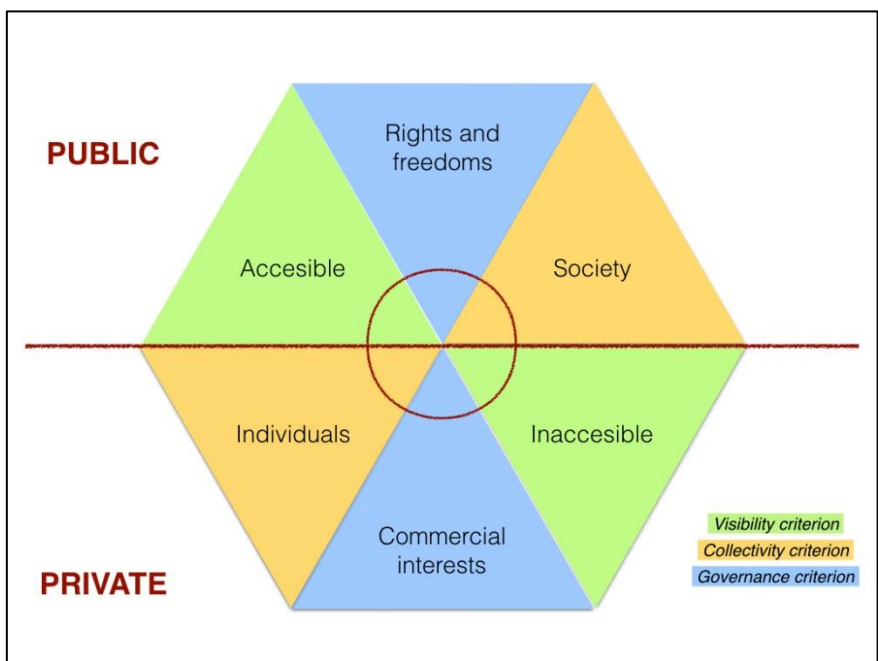

Figure 1: Diagram of the overlapping of the public and the private (self-made).

The boundaries between what is *public* and what is *private* are being redefined (if not erased) in the information era. In the words of Papacharissi, "cyberspace offers a new terrain for the playing out of the centuries of friction between individual (private) and collective (public) identity, the individual and the community" (Papacharissi, The virtual sphere: the internet as a public sphere, 2002). Interestingly, Gavison argues that the labelling of things as "private" or "public" is based on the existence of some normative structures which, if altered, may cause change in the adequacy of labelling (Gavison, 1992). Are we witnessing a redefinition of our normative structures led by social media platforms?

Papacharissi answers affirmatively: online media acquire agency as they enable the renegotiation of what is considered private and what is considered public in public life (Papacharissi, A Private

Sphere: Democracy in a Digital Age, 2010). She further sustains that, in the truest form of democracy, negotiation of that which is considered public and that which is considered private takes place within the public sphere (Ibid). Yet, we should not be fooled into equating platform owners with platform users. Platform owners, as hosts of the public sphere, should allow users to negotiate the meaning of private and public. However, as private actors, they should not be leading this discussion and the follow up transformation of our normative structures.

In addition to the criteria developed above, Weintraub identifies the major ways in which the previous distinctions are drawn in social and political analysis (Weintraub, 1997). Some of these theoretical roots help frame the human rights concerns of social media operating as the public sphere: the liberalist economic model and the republican model.

1. The liberalist economic model

This model draws the line between state administration and the market economy. It conforms the terms of "public sector" and "private sector", being the former one all governmental affairs and the later one the non-governmental matters. It is based on jurisdiction criteria. Thus, that which is the authority of the state corresponds to the domain of the public and that which pertains matters between private individuals corresponds to the private domain. The private is then equated to the market. The liberalist economic model arises in popular discussions of public policy and the "privatization" debate. The discussion focuses on whether the provision of some services and goods should be left to the market (the invisible hand) or, on the contrary, regulated by the state (the visible hand of administrative regulation)[20].

Although often raised in neoclassical economics, the liberalist economic model becomes relevant when discussing the role that social media plays in the public sphere. The space for rational deliberation of the public has traditionally been "public" and hence, free from commercial interests. Conversely, public opinion today is developed in the private domain of platforms that are profit driven. In addition, as Facebook grows in ambition and scale, it takes on duties traditionally placed under governmental affairs such as policing speech to safeguard democracy. All in all, from the liberalist economic perspective, the overlap of public and private is found in the bigger role of

---

[20] The "visible hand" is a term coined by Alfred D. Chandler. See*: The Visible Hand: The Managerial Revolution in American Business* (1977).

platforms in public (governmental) matters. What are the impacts on human rights of having market-led corporations regulating the public sphere?

2. *The republican model*

The republican or classical model sees the public realm as "the political", which stands for the active engagement of the community in society. The public is understood as a realm in which public debate and deliberations take place, directly or indirectly influencing the distribution of wealth and public services within a society (Jørgensen, Revisiting public and private, 2013). Public life makes use of notions of citizenship and active participation in collective decision making. This model differs from the previous one in the sense that it conceives the political community (public) as distinct from both the market and the state (private). The public is not the government because it is not formalized in the institutions of the state. By the public, we usually mean what is common to a given social organization that transcends the private (Castells, 2008).

Hannah Arendt builds on the concept of citizenship to equate public life to political life. Political opinions, she claims, can never be formed in private; rather, they are formed, tested and enlarged only within a public context of argumentation and debate (d'Entreves, 2016). She believes that politics are a public activity and, as such, one cannot participate in them without being in a public space. Furthermore, Arendt distinguishes one's life as an individual, pursuing private interests, and one's life as a citizen, pursuing public interests. According to Arendt, public interests are those of a public world which we share as citizens and which we can pursue and enjoy only by going beyond our private self-interest (Ibid). The spatial quality of public life is also an essential feature to Arendt. She argues that political activities are located in physical public spaces where individuals must be able to see and talk to one another in public, to meet in a public-political space (Ibid). Do the anonymity features of social media platforms then compromise such distinction of private/public? What is more, if social media platforms are to provide the space to develop one's citizenship, access to these tools should be as available as access to public plazas or parks. Instead, access to social media networks is provided by private companies and subjected to commercial terms.

Habermas does as well focus on the public. In his conceptualization of the early public sphere, the rational-critical discourse grew in the salons and coffee houses and was dependent on the rise of power states and the capitalist economy (Calhoun, Habermas and the Public Sphere, 1992). This process led to an idea of society separate from the ruler (or the state) and of a private realm separate

from the public (Ibid). In this way, Habermas divides society into a private realm of lifeworld relations (the family), a private realm of system relations (market economy), a public realm of system relations (the state), and a public realm of lifeworld relations (the public sphere) (Jørgensen, Revisiting public and private, 2013). It is arguable that, by blurring the distinction of public/private, social media platforms are integrating different realms, specifically the market economy and the public sphere.

The conception of the public as equal to the political that the republican model presents is the basis for most of the issues touched upon in this thesis. Tom Highfield suggests that, in social media platforms, the personal and the political are highly interlinked[21]. "The political and the personal are in concert in social media, as platforms are used to document experiences which are every day for some and to present users' identities, lifestyles and other interests. The articulation of gender, race, sexual orientation, religion, and other aspects of individuals' lives and their communities can make for a politicized life, by design or necessity" (Highfield, 2016). He adds that platform policies about acceptable content and practices further politicize everyday experiences[22]. In the same line, Papacharissi argues that the discourse surrounding the political potential of online new media could be located in the tension between the "private" and the "public", as articulated in contemporary democracies" (Papacharissi, A Private Sphere: Democracy in a Digital Age, 2010).

The republican model sees in social media a clear public role because it provides a forum for citizens' political discussions. Yet, it is not distinct from the market but hosted by it. This aspect evidences the intertwining of *private* and *public*. It also raises concerns over the freedom that the online political community enjoys to deliberate on public matters without interferences. How do platforms (the private), as hosts of public discussions, influence the political community (the public) and what impacts does this have on human rights?

1.4 Conclusion
Normatively, as described by Habermas, the public sphere is the space where public opinion is developed to influence political powers and is based on the critical deliberation of an informed public.

---

[21] To illustrate how the personal becomes politicised, he takes the example of voters taking selfies while in polling booths and uploading them to social media. Besides being a threat to the validity of the vote, the act is both personal and political. It is personal in the sense that it is about the individual and it is political as it is a clear engagement with a definite political activity. See: Highfield, T. *Social Media and Everyday Politics* (2016).

[22] Highfield presents the following example: A picture of a woman that features, among other physical attributes, the subject's nipples is highly personal and mundane, rather than political per se. However, the way in which platforms deal with this content, politicises the issue. The question of inappropriate nipples reflects wider debates that play out around feminism, sexism misogyny and gender. Furthermore, it may lead to social movements such as #freethenipple, which generated online but took the leap to the offline world. See. Highfield, T. *Social Media and Everyday Politics* (2016).

As public forums, social media platforms are functioning as today's online public sphere, one that is digital and global by nature. In this way, users rely on these sites for exercising their rights and freedoms associated with public life. Yet, as private actors, platforms make a profit from their services and effectively moderate content to regulate activity in their business. Thus, the public sphere of social media differs greatly from Habermas' ideal, which was based on open civic engagement and its independence from commercial interests.

This paper claims that this shift is rooted in the overlap of *the public* and *the private*. These concepts are being redefined online, causing major changes in our social and cultural norms. What is more, their reconceptualization is the source of the alterations in the public sphere that call into question the effective protection of human rights.

Essentially, private actors such as Facebook are taking on duties that were traditionally exercised in public spaces like public debate hosting. Through the mixture of the means in which *private* and *public* are generally invoked (visibility, collectivity, governance), concerns over the guarantees of freedoms and rights in the social media platforms are being raised.

Public sphere - some key elements	Facebook's features
Free from commercial interests	Online business model
Civic engagement and public participation	Content moderation

Table 3: Outline of the first and second chapters

The paper will now examine specific features of Facebook that are in conflict with the normative idea of the public sphere and risk, ultimately, having a negative impact on human rights.

Chapter 2 – The Public Sphere of Facebook

This chapter addresses the case of Facebook. The rationale behind this choice is first presented and followed by an examination of some of Facebook features in relation to the public sphere. This chapter draws mainly on the work of José van Dijck and Tarleton Gillespie, prominent scholars in new media, while drawing critical views from relevant voices within the tech world, social sciences, and human rights law. The analysis distinguishes two key features of platforms that may be in conflict with human rights: (1) the online business model, directly connected to right to privacy, and (2) content regulation policies, which raise questions on freedom of expression and the limitations of public discussions. Finally, the public/private criteria, developed in the previous chapter, is applied to the analysis to better understand the impact of the platform's features on human rights.

2.1 Why Facebook?

There are four main reasons for which Facebook was chosen to exemplify the impacts on human rights that social media has when acting as a public sphere.

Firstly, its size and dominance. Facebook is easily the largest of its kind in the world and the most widely used. In fact, if Facebook was a country[23], it would be the most populated country of the world with over two billion citizens (users)[24]. Its power and dimension are attributed to the platform's ability to grow as well as the speed in which it scales up. As Schmidt and Cohen depict: almost nothing short of a biological virus can spread as quickly, efficiently or aggressively as these technology platforms, and this makes the people who build, control and use them powerful too (Schmidt & Cohen, 2013) . Because of its leading position in social media services, Facebook is substantially influencing social and cultural norms and shaping practices in relation to privacy and control over data. The platform dominates the Western world and it's the preferred engine for social networking worldwide. Dijck points out that, by implementing various coding technologies and interfaces strategies, Facebook inscribes how online social interaction should be conducted (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). The platform's worldwide influence translates into power to set the standard for other similar platforms and the ecosystem of social media as a whole. Although

---

[23] In 2012, Rebecca MacKinnon coined the company as *Facebookistan* arguing that Facebook is not a physical country but its population comes third after China and India (it has now surpass them). "It may not be able to tax or jail its inhabitants, but its executives, programmers, and engineer do exercise a form of governance over people's online activities and identities" (MacKinnon, Ruling Facebookistan, 2012).

[24] Facebook counts with 2.20 billion monthly active users as of March 31, 2018. Retrieved from Facebook Newsroom: https://newsroom.fb.com/company-info/

not the pioneer, Facebook constitutes a true paradigm shift and it has managed to impose its ideology on an unprecedented scale, while naturalizing a business model that offers a free service, collects data from its users, and sells advertising on the basis of this data.

Secondly, Facebook serves as a great object of study when discussing human rights threats, precisely because the company makes use of human rights narratives. Of course, the platform has had many positive impacts on human rights but, based on a complex mix of company policies and legal standards, Facebook also enforces boundaries for freedom of expression and the right to privacy (Jørgensen, What Platforms Mean When They Talk About Human Rights, 2017). The company repeatedly claims that its mission is pursued in the interest of society: "We have really big aspirations around making the world a more open and transparent place. We define our aspirations more in terms of that mission than in terms of the company aspirations". (Sheryl Sandberg, CCO at Facebook)[25]. However, as it is only natural, the platform's main aim is not to protect the freedoms associated with public life, but instead, to make a profit from this service.

The third reason of choosing Facebook as the object of this study is its combination of multiple services that arguably resemble a virtual public sphere. The design and architecture of this online platform is often seen as filling the void of the already (offline) existing public sphere and revitalizing it. All features of the platform promote communication and civic engagement in one form or another: photo and video sharing, group chatting, private messenger, events development, etc. While other SNS are specialized in a specific service like microblogging (Twitter), video sharing (Youtube) or user generated content (Wikipedia), Facebook's ambitious strategy has been the constant expansion of its services. Note that the social network company has acquired over 50 companies including *Instagram* (in 2012) and *Whatsapp* (in 2014). This way, it has managed to stay innovative and develop to be engrained to our society and everyday lives. Hence, it is this platform, among all social network sites, that deserves the most scrutiny when examining "the virtual public sphere".

Lastly, Facebook is at the center of public discussion. The company has been subject of controversy for a litany of issues, including Russians interferences with the intent to sway the 2016 US elections and user data sharing with third parties (Cambridge Analytica scandal). Although leading the SNS

---

[25] *Evangelical networker who wants Facebook to open up the world.* The Guardian, 20th August 2009. Retrieved from: https://www.theguardian.com/business/2009/aug/20/facebook-ceo-sheryl-sandberg-interview

market and breaking records on users worldwide, Facebook practices and business model are (finally) now being called into question.

2.2 Facebook and human rights

As defined by Dana Boyd, social network sites are web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system (Boyd, 2018). Boyd stresses the *visibility* aspects of social media by highlighting the publicness of user profiles and positioning social connections at the center of the concept of social network sites. Kaplan and Haenlein further highlight two common features that all social media have in common: 1) they are Internet-based applications and 2) their services are based on user-generated content (UGC) (Kaplan & Haenlein, 2010). These features can be linked to the *collectivity* criterion in the sense that social media is a domain of the shared interests. From a different angle, Gillespie points out to the intermediary role that social media play: "what unites all platforms is their central offer of hosting and organizing user content for public circulation, without having produced or commissioned it" (Gillespie, Governance of and by platforms, 2017). It is in this aspect that the *governance* criterion becomes important. As social media platforms are privately owned, they are governed by the market. Commercial interests drive decision making of platforms and the way in which social media hosts user generated content.

It is obvious that social media, along with all the other new information and communication technologies, brings both benefits and costs to any given domain (Latonero, 2018) . Yet, when it comes to the human rights context, the potential risks arising from using social media platforms as forums for public discussions deserve greater consideration. Particularly relevant are the right to freedom of expression and freedom of information, the rights of peaceful assembly and association, the right to take part in the conduct of public affairs, and the right to privacy.

As a private corporation, Facebook operates outside the human rights protection framework and accordingly, the legal responses to the interference with fundamental human rights in this platform have yet to be found. It is for this reason that the (negative) effects of platforms on human rights must be specially considered.

This paper looks at two specific features of social media that are believed to have an impact on fundamental rights and freedoms. Firstly, the online business model which, based on the

commercialization of personal information, may be in conflict with the right to privacy. And secondly, content regulation policies of Facebook that shape public discourse and may limit freedom of expression.[26]

Table 4: Outline of the second and third chapters

Facebook's features	Human rights
Online business model →	Right to privacy
Content moderation →	Freedom of expression

"Facebook tries to disappear as a technology: users are steered like Truman in the movie "The Truman Show", but the fact that they are in this show is no part of their lifeworld" (Heyman & Pierson, 2015). As Heyman and Pierson accurately depict, we have embraced Facebook's platform and it is has gradually become part of our everyday. From buying second-hand things, to getting together to protest unfair laws, or simply sharing pictures of the holidays with friends, Facebook is the place where it all happens. Of all these functions, this thesis focuses on the use of social media as a forum for public discussions: the public sphere. Impacts on human rights are particularly interesting when Facebook takes on "public duties". In this case, it provides the space for discussion, which was traditionally located in public spaces.

How is social media affecting the freedoms and rights associated with public life?

2.3 (A) Online business model

Facebook's aspirations are greater than being a forum for public discussion; they company looks to profit from it. The way in which the platform makes a business out of information exchange is heavily downplayed to regular users behind an "information is free" attitude. Advertising funds this service to a great extent. But, certainly, this is not a neat ideological fit with the ethos of the participatory web – not to mention that, for the most part, the users generating the content do not enjoy any revenue in return (Gillespie, The Politics of Platforms, 2010).

As the cultural industry suffered major changes by shifting its emphasis from products (CDs, books, DVDs) to services (downloadable music, books and movies through media platforms), media

---

[26] Note that this thesis examines some features, among many others, that impact human rights and deserve further scrutiny. However, these are merely two case studies that exemplify some of the impacts of social media on human rights but there are many other aspects that, for the sake of conciseness, have been left out of this paper.

companies had to develop new ways of monetizing online creativity and sociality (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). The arrival of social media platforms particularly altered the industry by maximizing user generated content to offer free content and services. Attempting to earn money by treating social networks as just another entertainment medium was doomed to failure (Clemons, 2009). That is not to say that owners do not have monetary intentions, but that their business models have developed in a much complex way. Some UGC platforms have tried to establish a business model based on subscription fees, where the consumer could choose from a "free" model sustained by advertising or an advertising-free subscription option for "premium" users. However, this strategy is fundamentally opposed to Facebook's mission of empowering people and connecting the world. Facebook claims it is bringing freedom of expression to a whole new level, for which their services cannot be restricted to those that can financially afford them. Hence, the platform offers "free" services and bases its business model almost entirely on advertising[27].

Adapting advertisement to fit the model of social media

Advertising is not new to the industry and it goes way back in traditional broadcast media. But the selling of audience attention juxtaposed with cultural content had to be reinvented to fit the web 2.0. Conventional advertising strategies appealing to mass audiences no longer applied to a world dominated by UGC and social networking; in online social environments populated by "friends", users neither expected nor tolerated commercial activities (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). Clemons accurately indicates that the business model of social media is a delicate harmonizing act between users' trust and owners monetizing intentions (Clemons, 2009). In traditional entertainment users felt comfortable being exposed to paid messages in exchange for being entertained. Differently, Facebook is seen as a space where people congregate to exchange information and entertain themselves. As users take on a more active role and platforms act as the mere milieu, mass advertisement is not so well received. If users feel that they are being manipulated or exploited, they simply quit the site, causing the platform to lose its most important asset (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). But

---

[27] Although very limited, there are non-profit social networks that run in a non-commercial basis. Diaspora is one of them. Bases on a free software, it is not owned by one person or entity. Consequently, there is no need to surrender your privacy to a big business as a prerequisite to use these services. See: *Four Nerds and a Cry to Arms Against Facebook.* New York Times (May 11th, 2010). *https://www.nytimes.com/2010/05/12/nyregion/12about.html*

obviously, the norms for how much commercial activity users tolerate are gradually stretched and expanded to include new tactics and business models (Ibid).

The revival of advertisement came with automated customization and personalization of ads. Instead of suffocating users with pop-ads and mass advertisement, social media platforms would personalize ads to the liking of the user. The strategy is brilliantly presented as a win-win: companies have their product shown to targeted potential customers, users are only exposed to those advertisement that might actually catch their eye, and the multi-billion dollar company gets to keep its image of freedom of expression enabler by de-emphasizing its commercial interest. The corporation claims that personalized ads enrich user experience and constitute part of the offered service: "Data helps us show you better and more relevant ads. Data lets a local coffee shop survive and grow amid larger competitors by showing ads to customers in its area. And it lets a non-profit promote a diabetes fundraiser to those interested in the cause" (Vice President of Facebook Ads, Rob Goldman)[28]. While data may, on the one hand, serve a good purpose, on the other hand, it is massively processed and shared with third parties for Facebook's profitability.

Gillespie accurately catches the essence of Facebook's balance: "the business of being a cultural intermediary is a complex and fragile one, oriented as it is to at least three constituencies: end users, advertisers, and professional content producers. This is where the discursive work is most vital. Intermediaries must present themselves strategically to each of these audiences, carve out a role and a set of expectations that is acceptable to each and also serves their own financial interests, while resolving or at least eliding the contradictions between them." (Gillespie, The Politics of Platforms, 2010)

Targeted advertisement: the exploitation of data

The lifeblood of Facebook's online business model and advertisement strategy is data. Essentially, actions of individuals in the platform (from conversations to search of information) are captured, retained, and used for advertisement purposes, and as such, they constitute products that feed into the online business model (Jørgensen, Human Rights and Private Actors in the Online Domain, 2018). Facebook monetizes the information it holds about its users by allowing third parties to target potential customers based on profiles derived from this data. This business model is also known as

---

[28] Retrieved from Newsroom, Facebook: https://newsroom.fb.com/news/2018/04/data-and-advertising/

personal information economy (PIE) and it makes use of every transfer of data that takes place at every instance of online participation. Facebook knows much more than whether you are spending time in Farmville or watching cat videos. The company probably knows that you are getting married soon, that you take certain medication or that you despise your job. By noting behavioral patterns, social habits, personal preferences, and online performances; social media companies collect all kind of personal information and assemble databases that are later used to market to specific target populations. Zuboff has coined this as surveillance capitalism (Zuboff, 2015) based on the idea that personal data is now a tradable commodity in capitalist societies. The main concern here is that surveillance capitalism places the free market economy inherently at odds with privacy (Papacharissi & Fernback, 2005). Through targeted advertising, users are tracked extensively and large amounts of their personal information are being retained (Ghosh & Scott, 2018). The scale of information storage should set off the alarms and raise concerns over privacy. However, social media corporations spare no efforts at branding its advertisement strategy as a social network experience that brings convenience for users and advertising potential for companies: "You get better ads, advertisers get a tailored audience and everyone on Facebook gets our firm commitment to privacy" (Vice President of Facebook Ads, Rob Goldman)[29].

There are two main aspects that Facebook relies on to captivate customers: attention and popularity. *Attention* applies to the traditional approach of pop-ups, banners and explicit ads. The classic "selling eyeballs to advertisements". However, this time they are customized to meet one's interest and, as such, are presented in a much subtler way. Users are constantly exposed to this kind of advertisements but, often, they are unconscious about it. If you are planning a holiday for the summer, you might be shown articles on "the best destinations of 2018" followed by some convenient deal by an airline company that seems too tempting to let go. Because these ads are tailored to users' needs (and based on their personal data), they are not perceived as aggressive or bothersome as massive advertisement. The second aspect, *popularity,* brings the disguise to a whole new level by using social bonds for commercial ends. Contrary to pop-ads, popularity is not simply out there. It is rather engineered through algorithms that prompt users to rank things, ideas, or people (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). Popularity is based on a new metric: engagement; the glue of social media. This way, instead of airline banners, you might be encouraged to click on a travel agency page through a message that says: "your friend likes this and recommends

---

[29] Retrieved from Newsroom, Facebook: https://newsroom.fb.com/news/2018/04/data-and-advertising/

it to you". Note that popularity is distinct from values such as trust, objectivity or quality. Yet, it has been proved that individuals trust more a friend's suggestion than a side banner. Facebooks makes use of this fact as a sales pitch: users are shown recommendations that are ostensibly sent by "friends" but are in fact propelled by third parties using an unwitting friend's Like as an advertising ploy (Ibid). In digital marketing terms, this strategy is referred as "sponsored stories" (SPS) and it has transformed the industry by camouflaging advertisement as UGC. "The initial goal of sharing experiences between friends online in the News Feed is extended towards advertising, without being visible to users" (Heyman & Pierson, 2015). This feature distinguishes online media from traditional media such as television or newspapers where promotional content needs to be explicitly separated from editorial content.

One of the most known and successful tools of SPS is the "Like button"[30], the feature used in the platform through which users can express that they like, enjoy or support certain content. It was launched by Facebook in 2009 and has expanded over the years to other social media and services alike. Because it is a social plugin[31], the Like button is available outside of Facebook´s platform enabling the company to gather enormous quantity of data from other apps and websites that is later used for personalized spamming[32] and other marketing strategies in Facebook. Its purpose is to connect people, things, and ideas and, this way, create an enormous information database. "Facebook's Like buttons enable multiple data flows between various actors, contributing to a simultaneous de- and re-centralization of the web" (Gerlitz & Helmond, 2013). User engagement is then easily translated into numbers or, in van Dijck's words, into an automated quantifiable commodity. This practice has become so massive and essential to social media platforms and their business model that it has been coined the "Like economy". More importantly, it has turned personal data sharing by third parties into an accepted practice in the online universe (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). In fact, only three months after its

---

[30] Facebook explains this feature serves to "give feedback and connect with things you care about." It further claims that is an easy way of showing care: "Clicking Like below a post on Facebook is an easy way to let people know that you enjoy it without leaving a comment." Retrieved from: Facebook Help Center,
https://www.facebook.com/help/452446998120360/
31 Social plugins are, as defined by Facebook, "tools that let you share your experiences on other websites with your friends on Facebook". Examples include the Like button and the Share button. Retrieved from: Facebook Help Center, https://www.facebook.com/help/203587239679209?helpref=related
[32] Studies show the accuracy of predictions based on "Likes": "A computer knows more about a person's personality than their friends or flatmates from an analysis of 70 likes, and more than their family from 150 likes. From 300 likes it can outperform one's spouse". Retrieved from: *Why Facebook's Business Model Is Incompatible With Human Rights* Independent, April 3rd, 2018. https://www.independent.co.uk/life-style/gadgets-and-tech/facebook-business-human-rights-data-breach-social-media-personal-information-a8286021.html

introduction, more than 350,000 external websites had already installed the Like button, much to Facebook's benefit (Ibid).

Rightfully, Van Dijck remarks that the use of Likes to promote brands takes place without user's consent, "users automatically approve of this tactic by signing Facebook's terms of service (ToS)" (van Dijck, Connective Memory: How Facebook takes Charge of your Past, 2016). In other words, to be able to use the platform, users have no choice but to agree to the (privacy) terms that make targeted advertisement possible. Further, privacy policies of social network sites are often complex and written in legalese that is incomprehensible to the general user. Can we then speak of user consent?

Redefining privacy: from connectedness to connectivity

Because of its leading position within social media, Facebook practices are setting the standard for other platforms and influencing social and cultural norms related to privacy and control over data. José van Dijck develops an exceptional framework that helps understanding social media as a techno-cultural and socio-economic phenomenon (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). In her work, *The Culture of Connectivity*, she examines Facebook (among other platforms) and sheds some light on the meaning of *sharing*[33], strategically pitched by the platform against the concept of *privacy*. Van Dijck highlights the ambiguity of the former term: "it relates to users distributing personal information to each other, but also implies the spreading of that personal information to third parties" (Ibid).

Platform owners emphasize the social meaning of *sharing* and they do so in opposition to opacity, non-transparency, and secrecy. Privacy is perceived and presented as negative, highlighting ideas of individualism and reluctance to share and socialize. By making people accustomed to openness and hesitant towards privacy, Facebook has orchestrated a major cultural shift. The fact that we feel comfortable sharing so much of our lives on social media is causing the worlds of online to interpenetrate. This way, today's interactions on an individual, community and societal levels, are often influenced by the multi-billion dollar platform of Facebook (Ibid).

---

[33] Van Dijck distinguishes other terms such as "friending" or "liking" that, along with "sharing", have acquired an ideological meaning orchestrated by Facebook in their own interest.

To better explain the meanings given to the term *sharing*, Van Dijck talks about a gradual shift from "connectedness" to "connectivity". <u>Connectedness</u> stands for the conscious sharing of information by users with other users (user-centered sharing), while <u>connectivity</u> refers to the sharing of user data by the platform with third parties (owner-centered sharing). The most important point she makes is that Facebook has a stake in promoting the first type of mechanism while diverting attention from the second type (Ibid). The argument behind it is that the more users know about the management of their personal data by social media corporations, the more likely they are to raise objections against these practices.

From a technological point of view, coding structures of both concepts differ greatly. Facebook's platform is designed in a way that incites users to share as much information as possible. Features like the *Wall* or *People You May Know*, for instance, motivate users to disclose personal information and increase connections with other users. Individuals are also encouraged to share their whereabouts or make announcements on personal life events such as relationship or professional status. Facebook is in a perpetual beta, and by constantly reinventing the service with new tools, it encourages and boosts social sharing: "Facebook is always working on new ways to connect and share" (Facebook privacy basics)[34]. More recently, the platform added *Did You Know*, *Fun Fact*, and *Lists* features. Through an innocent message that says "answer a question to help people get to know you" users are encouraged to disclose childhood stories, personal fears, or travel wish lists. And while updates are presented as positive changes because they enrich social experiences, developers are working day and night to aggregate and process all this new data and, through targeted personalization strategies, convert it into profit. "Platform owners have a vested interest in complete openness on the side of users; the more they know about users, the more information they can share with third parties" (Ibid). However, this second meaning of *sharing (*from the platform with third parties) is less visible to users. "The ability of users to exercise control is limited to adjustments on how they share information, while they have no means of limiting the data collection that takes place as a premise for using the services" (Jørgensen, What Platforms Mean When They Talk About Human Rights, 2017).

Facebook does not disclose the way in which its algorithms work nor the manner in which personal information is processed. Yet, it reassures users that they are in full control of their personal information: "You are in charge. You have control over who sees what you share on Facebook"

---

[34] Retrieved from Facebook: https://www.facebook.com/about/basics/get-to-know-facebook/new-ways

(Facebook privacy basics).[35] This is a risky statement as it leaves privacy up to the user. Rather than being the default set by the network, the burden is on the user to protect himself. Furthermore, privacy settings have an illusory effect as they do not cover *sharing* in the second meaning of the term (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). In fact, user control over information might even be another "coding strategy" based on the "paradox of control". Behavioral studies[36] have shown that "providing users of modern information-sharing technologies with more granular privacy controls may lead them to share more sensitive information with larger, and possibly riskier, audiences" (Porter, 2018). Hence, paradoxically, these tools might end up leaving users more vulnerable than before, and in turn, they might enable Facebook to gather more data to monetize. As van Dijck concludes, the power of owners over coding technologies gives them a distinct advantage over users in the battle for informational control (Ibid).

This power transcends the technological sphere and it entails, as well, a cultural transformation. The double logic of user centered-connectivity versus owner-centered connectedness penetrates in society through our digital lives to the point that it is profoundly affecting social and cultural norms. The change, driven by Facebook, is of course shaped conveniently to fit the company's profit-driven interests. Facebook's philosophy of making the world social is highlighted in all its operations as the company claims its mission is foremost social. Doing so, values of connectedness and community are equaled to connective commercial values, smoothly aligning business models with user interests (van Dijck, Connective Memory: How Facebook takes Charge of your Past, 2016). "Capital has subsumed social communication (connectedness) and changed it into a more profitable form (connectivity)" (van Dijck, 2013)[37]. Thus, a new category emerges where customers are matched to friends: "frustomers". Referring to customers as friends, power relations are rendered invisible. The company presents itself as being on the side of customers, instead of profiting from them. It is shown as a social experience that aims at connecting the world, not a business strategy. This philosophy has enabled Facebook to conduct its frictionless sharing stratagem. Progressively, the company has been negotiating and shaping the meaning of *sharing* while users have gradually adapted to a model of

---

[35] Retrieved from Facebook: https://www.facebook.com/about/basics#youreincharge
[36] This study can be found in: Brandimarte, Acquisti, & Loewenstein, *Misplaced Confidences: Privacy in the Control Paradox*, 2012.
[37] Christian Fuchs talks about prosumer commodity: "In social network sites, information consumption becomes productive, creative, and an active process of surplus value production. Users, their personal data, and their usage behavior become object of permanent economic surveillance and commodification so that profit can be accumulated by selling the users, their data, and their usage behavior as commodity to advertising clients."

online interaction where "the default is social"[38]. The fact that we feel comfortable with such an irrational level of information disclosure reflects a tremendous cultural shift. "In less than eight years, the meaning of sharing, once understood as user-to-user information exchange, has subtly been replaced by a meaning that naturalizes the sharing of personal data with anyone in the world" (van Dijck, The Culture of Connectivity: A Critical History of Social Media, 2013). Facebook's ideology of *sharing*, *friending* and *liking* has spread deeply into the pores of our (digital) lives, affecting the very fabric of sociability (Ibid)[39].

The hidden dangers behind data

The primary concern with the collection of data is its misuse. Van Dijck explains that correlations between personal data and web-browsing habits could be used to reveal consumers' political affiliations, religious beliefs, sexual orientations, or health issues (Ibid). Huberman goes even further by exposing that this power can also be exploited in a perverse way to monitor individuals, expose them, blackmail them or, to a bigger extent, change political regimes (Huberman, 2001). And while such cases have taken place and they have caused great outrage in the public[40], the mere fact that these corporations have the power to act in such a way should already be worrisome. The exploitation of data in social media is very invasive, for which concerns over privacy and data control cannot only be the response to massive breaches. Users should question the "default is social" deal that private bodies like Facebook are imposing and doubt of its intentions, instead of just coping with the new tailored privacy standards set by Facebook itself and adapting to new norms.

Beyond the misuse of data, which is an obvious concern, it is important to reflect on all the implications that the online business model may have and the "ordinary" problems that are intrinsic to it. Lawrence Lessig points to two worries: manipulation and discrimination. They are rooted in targeted advertisement, also known as "profiling". Each individual on Facebook carries a user profile created by the company thorough aggregation of data. The profile includes a huge amount of information about the person that narrows the content he/she is exposed to. Under the pretext of advertisement, vast efforts go into behavioral data collection to segment audiences with precision and

---

[38] This is a quote that Facebook's CEO, Mark Zuckerberg, used to explain the company's mission: to build a Web where the default is social in order to make the world more open and connected. Retrieved from: Dan Fletcher. *How Facebook is Redefining Privacy*. Time, May 20, 2010. http://content.time.com/time/magazine/article/0,9171,1990798,00.html
[39] An example of this puzzling change in society is doctors keeping patient's private information confidential and then discovering that patients themselves publish the most intimate details of their illnesses on their Facebook pages.
[40] The most recent scandal that Facebook has been involved in is the data breach by the company Cambridge Analytica, which might have been used to influence elections in the United States.

target them with different messages most likely to produce a response (Ghosh & Scott, 2018). This is however a slippery slope because, as Lessig advances, there is a possibility that "profiles" begin to normalize the population from which the norm is drawn: "The system watches what you do; it fits you into a pattern; the pattern is then fed back to you in the form of options set by the pattern; the options reinforce the pattern; the cycle begins again" (Lessig, 1999). In other words, "the observing might affect the observed". Whereas in television advertisement the motives are clear, this is not the case in social network sites. Targeted advertisement techniques are undertaken by complex algorithms in an opaque process that is made invisible to the users. This way, platforms are detached from the commercialization sphere that they actually constitute and, mixed with UGC, advertisements are not perceived as such.

Discrimination concerns are also rooted in the backstage algorithmic processes. To illustrate the issue of discrimination, Lessig goes back in time and looks at the bigger picture: "the market has traditionally carried and idea of equality, where everyone could purchase and sell equally. Yet, equality has been displaced by economic zonings that aim at segregation" (Ibid). By categorizing individuals on the pretext of tailored ads, the system may limit one's capabilities promoting distinctions based on social or economic criteria. In fact, machine learning algorithms adjust to people's behavior and may reinforce their prejudices and bias[41]. For instance, a study has found that Google ads on high-income jobs were shown to men much more often than they were shown to women (Datta, 2015). Labelling individuals, with the alleged reason of improving user experience, emphasizes distinctions among people and nourishes discriminatory practices.

It is imperative that we question the business model of technological companies such as Facebook and revisit the implications they may have to us individually but also as a society. In addition, we must reject deterministic claims about technology, and look further. The collection of personal data sustains the business model of corporations like Facebook but it also puts power in hands of private actors. Furthermore, beyond the misuse of data, the management of vast collections of personal information by private companies may already be in conflict with the right to privacy. Hence, questions are being raised on the obligations of platforms to respect but also to protect and remedy

---

[41] See: *When Algorithms Discriminate*, New York Times, July 9th, 2015. Retrieved from:
https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html

our right to privacy. Is this a feasible demand or is Facebook's business model simply incompatible with human rights?

2.4 (B) Content regulation

Harrowing testimonies of executions and rapes by Myanmar security forces were provided by Rohingya Muslims when insurgent attacks sparked a security crackdown in Myanmar a year ago. The events were so appalling that the United Nations suspected acts of genocide had taken place. It turns out that, by facilitating hate speech and misinformation, Facebook played a "determining role" in these events. Fabricated stories were spread out in the platform. They soon went viral, causing ethnic divide and fueling violence against the Rohingya minority (The Guardian, 2018). In Myanmar, Facebook is so dominant that it is considered as the Internet itself[42] and it is the only source of information for many. In fact, as UN Myanmar investigator Yanghee Lee stated: "Facebook is a huge part of public, civil and private life, and the government uses it to disseminate information to the public" (Reuters, 2018). The case raised big concerns on the role and responsibility of Facebook in passively enabling human rights abuses[43]. Why did this hate propaganda avoid the company's community standards while others do not make the threshold?

The UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, David Kaye[44,] declared that "Internet companies provide unprecedented space for communication and access to information, but they are also the global speech regulators of our time" (UN Human Rights Council , 2018)[45]. Given Facebook's dimension and role in public life[46], it is time to question their responsibilities as publishers of information. Why and how do platforms curate

---

[42] Surveys show this is a widespread misconception: more people say that they use Facebook than say they use the Internet. Retrieved from: Quartz (2015).
https://qz.com/333313/milliions-of-facebook-users-have-no-idea-theyre-using-the-internet/

[43] There are several cases where platform owners have abstained from preventing the spread of human rights abuses. For instance, last year, Facebook's life video capability was used to broadcast torture of a man in Chicago and the video went viral before it got removed. Likewise, the hashtag #GolpearMujeresEsFelicidad (Beating Women Is Happiness) led a harassment campaign on Twitter promoting violence against women. See: *Do Social Media Platforms Really Care About Online Abuse?* Forbes (2017). https://www.forbes.com/sites/kalevleetaru/2017/01/12/do-social-media-platforms-really-care-about-online-abuse/#653fb62b45f1

[44] Find here David's Kaye biography and a summary of his work:
https://www.ohchr.org/en/issues/freedomopinion/pages/davidkaye.aspx

[45] Find his declarations here:
https://www.ohchr.org/RU/NewsEvents/Pages/DisplayNews.aspx?NewsID=23218&LangID=E

[46] Although online sociability has never replaced the offline one, the use of social networking sites increases at an unprecedented scale. Statistics show that user base has grown more representative over the years. Furthermore, social media is part of most users 'daily routine. Retrieved from: Pew Research Center, (2018). http://www.pewinternet.org/fact-sheet/social-media/

content? Where do they draw the line and how do they strike a balance between freedom of expression and a healthy community?

Tarleton Gillespie provides an interesting angle to the content regulation analysis. His last book includes a quote from a worker of a SNS content policy team: "In the ideal world, I think that our job in terms of moderating function would be really to be able to just turn the lights on and off and sweep the floors… but there are always the edge cases, the grey" (Gillespie, Custodians, 2018). Gillespie illustrates how content moderation was intended to be simple maintenance; just like a custodian in the janitorial sense, who sweeps the floors. However, unexpectedly, "knowing what needs scrubbing" became a difficult task, rather than the mundane and ordinary routine that was first conceived as. Continuing the parallelism, as janitorial staff are instructed to work at night or with as little intrusion as possible, content moderation was supposed to be mostly invisibly (Ibid).

The promises of the participatory web quickly began to sour as fake news, ISIS beheadings, or GamerGate[47] emerged in social media. As the platforms grew, participation in social network sites saw the noble and the reprehensible. There are many benefits to the use of social media, but its perils are becoming more and more obvious. The impact of these platforms on public life has become evident and corporations have had to respond by taking on responsibilities that were not foreseeable from the start. "As a society, we have handed over to private companies the power to set and enforce the boundaries of appropriate public speech for us. That is an enormous cultural power, held by a few deeply invested stakeholders, and it being done behind doors, making it difficult for anyone else to inspect or challenge" (Ibid). As digital intermediaries, these corporations make the call on what content to distribute and to whom, how to connect users, and what to disapprove from appearing in the platforms. Gillespie suggest that beyond acknowledging the occurrence of content moderation, we must rethink it and revisit difficult questions on how social media platforms tune the public discourse they host.

Firstly, it is important to demystify the claim that platforms are mere intermediaries in user communication, that they are impartial, that they do not take sides. Social media platforms promised the unmediated public sphere, the ultimate form of unhindered participation. Yet, they soon began to look at ways to profit from it and started intervening in the participation they hosted. Because value

---

[47] GamerGate is the name given to a campaign of online abuse in the global video game industry that took place in 2014. The harassment campaign targeted several women in social media. Retrieved from: *Feminist Critics of Video Games Facing Threats in 'GamerGate' Campaign.* New York Times (2014).

comes from the data users leave behind, platforms seek to maximize communication and get more and more users to join the "community". But doing so at a global level entails maintaining a healthy environment where everyone feels comfortable to take part in the discussion, a user friendly space. This is why Gillespie argues that moderation is the commodity: "content moderation is essential, definitional and constitutional to platforms" (Ibid). In the name of impartiality, platforms heighten some forms of discourse and heavily moderate others. In fact, they are purposely designed to shape participation towards particular ends, for which platforms cannot be considered neutral actors.

Yet, content moderation is not an easy task. It requires making difficult and often unsound distinctions. In addition owners face many vocal calls to both permit contentious speech and limit it. For instance, from policymakers, from users, from civil society, from the press and from the international community. SNS are in constant conflict by having to find the balance between being merely facilitators of public expression and being protectors of a community embodying democratic practices. The two principles, speech protectors/ community protectors, are at odds when deciding how to moderate content. "Policing public expression and social behavior at this scale requires weighing competing, irreconcilable values: freedom of speech vs protection of harm, avoiding offense vs raising awareness, hiding the obscene vs displaying the newsworthy" (Gillespie, Facebook can't moderate in secret anymore, 2017).

This dichotomy is the central tension that dominates social media practices. Moreover, it explains the incoherence of their policies and their often ad hoc application. In Gillespie's words: "Platforms face a double-edged sword: too little curation, and users may leave to avoid the toxic environment that has taken hold; too much moderation, and users may still go, because what was promised to be an open platform feels either too intrusive or too antiseptic" (Gillespie, Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media., 2018).

The rationale behind content moderation is twofold: on one hand, national legislation may require social media companies to police the content and, on the other, social media corporations regulate curate content for business reasons. This chapter explores hereafter both aspects: the governance of platforms (regulations imposed by governments on social media) and the governance by platforms (content curation regulations on their own accord). Lastly, it examines and questions a new form of governance led by users through the flagging mechanism.

Governance of platforms

"Since platforms gather people and collect traces of their activity, they present a compelling opportunity to policymakers to govern users through them" (Gillespie, Governance of and by platforms, 2017). How does Facebook, an American corporation operating transnationally, cope with conflicting national regulations? Content moves easily across regional jurisdictions and different states have different policies towards harassment, terrorism, hate speech, or even political speech. What position does Facebook take when mediating between citizens, law enforcement and policy makers? This is obviously defined by the extent to which the platform is hold responsible for the content it hosts, the speech of its users and their actions. "Intermediary liability"[48] differs from country to country, as so do their cultural norms.

The United States, country where most of SNS companies reside, has a safe harbor agreement as part of its telecommunications law[49]. It consists of two parts: intermediaries cannot be held liable for the speech of their users and, if intermediaries decide to police what the users say or do, they don't lose their safe harbor protection by doing so (Ibid). This is to say that platforms are not required to adhere to some standard of effective policing or content moderation. It is quite interesting how, in spite of setting the threshold of the speech that needs to be moderated or removed, platforms are still not considered "publishers" or liable for this content. In the United States, social media owners orchestrate public discussions but wriggle out of the responsibilities that come with it[50].

Mackinnon et.al distinguish three degrees of liability: broad immunity, conditional liability, and strict liability (MacKinnon, Hickok, Bar, & Lim, 2014). Whereas jurisdiction in the United States grants intermediaries broad immunity; European Union states, along with Russia and most part of South America, apply conditional liability terms. In these regions, SNS companies are exempt from liability if they remove the user content when asked to do so by the state or the courts, or if they had no

---

[48] "Intermediary liability" is the name given to the regulatory approach that examines the extent to which intermediaries can or should be held legally responsible – or "liable" – for the activities of their users. Retrieved from: (MacKinnon, Hickok, Bar, & Lim, 2014).

[49] Section 230 of the Communications Decency Act.

[50] Gillespie develops the challenges that reveal the limitations of the safe harbor provision, when applied to social media companies. Firstly, most of these laws were not designated with social media platforms in mind. Contemporary ICT like Facebook were not anticipated when the law was passed but, in the absence of any other law, platforms claim they enjoy safe harbor. Secondly, while intermediary liability is typically nation-specific, platforms are largely not. And, finally, some categories of content particularly abhorrent to users and governments (i.e.: terrorism) are challenging presumptions on how much liability platforms should face. Retrieved from: Gillespie, *Governance of and by platforms*, 2011.

knowledge of such content and were unaware of its illicit nature. Examples of states that count with a strict liability jurisdiction are China and most of the countries in the Middle East.

Although the broadest, the U.S. safe harbor approach embodies conflicting views of social media platforms. Mueller accurately exposes this conflict by claiming that the safe harbor "intended both to immunize platforms who did nothing to restrict or censor their users' communication and to immunize those platforms who took some effort to restrict online pornography and other forms of undesirable content" (Mueller, 2015). The former ones were protected in order to promote freedom of expression, while the later were immunized to enhance their ability to remove inappropriate behavior. This tension continues to shape the way we think about the role and responsibility of intermediaries such as Facebook and how we regulate them (Gillespie, Governance of and by platforms, 2017).

Governance by platforms

As we have just seen, content moderation policies are somewhat motivated by government demands on restrictions of undesired content. However, on top of this, platforms have wide self-regulation policies designed to take care of the communities they host. And, interestingly, social media corporations curate public speech to a much bigger extent that what they are legally required to do so. Interventions are ceaseless and systematic, and they cut much deeper than the law requires (Ibid). Why? There are two major reasons: 1) they want to protect the public brand, 2) they want to maximize profit. Both of these arguments explain why platforms moderate content to the extent that they do but, indeed, one follows the other: they want to protect the public brand because they want to profit from it. It is reasonable to combine them as one and discuss the economic motive of their practices. Yet, Facebook's role in our everyday lives is so important now that content moderation policies and practices cannot longer be just about profitability. Their indisputable role in public life has complicated what initially was a simple task. They prefer self-regulation that allows them to keep their promises of being "good" actors than government restrictions that make Facebook appear an extension of state power over its users.

As with the Myanmar case, Facebook has been involved in several incidents that reveal that profitability is no longer a feasible driver for their content regulation practices. Fake news and hate speech, among others, are dismantling the company's mantra of making the world better. The ethos of the participatory web cannot longer be sustained when users are being harassed or misinformation

in these platforms is causing political turbulence. Hence, economic considerations are intertwined with a commitment to nurture a healthy community, a sense of public obligation, and attention to criticism leveled by angry users, journalists, or activists (Ibid).

Platforms face all types of challenges, from interpersonal conflict to obscene speech but where do they draw the line? Gillespie notes that generally, all social media platforms regulate in some way or another, at least, content on the following: sexual content and pornography, representation of violence and obscenity, harassment of other users, hate speech, representation of or promotion of self-harm, and representations of or promotion of illegal activity (particularly drug use) (Ibid)[51].

However, there is only one accurate answer to the question above: there is no line. Platforms have had to develop their own rules and sophisticated means of policing their sites, which still are often applied in an ad hoc fashion. The practice of content curation takes place in a big grey area where no two cases are alike. Attempting to strike a balance between protecting freedom of expression and maintaining a healthy online community, Facebook applies its "Community Standards" in an, inconsistent, case to case basis: "In some cases, and when we are provided with additional context, we make a decision based on the spirit, rather than the letter, of the policy" (Facebook Community Standards, 2018). As Gillespie rightly points out, beyond simply guiding enforcement, platforms do important discursive work (Gillespie, Governance of and by platforms, 2017). They are somewhat public arbiters. The problem lies in the obscurity of this process. Facebook acts as a gatekeeper of public speech but the way in which content is curated or encouraged is invisible to users. Their Community Standards are rather vague when it comes to content moderation and their internal guidelines generally opaque.

Another aspect worth looking at is the progression of such guidelines. Community standards are created not just in anticipation to undesired content but in response to it. Content moderation policies are adjusted in reaction to outcries and controversies: "in these guidelines, we can see the scars of the past challenges" (Ibid). There is so much content in Facebook's platform that, reviewing it before it is posted, has become an impossible task[52]. It is the scale of the problem that hinders a before

---

[51] These guidelines are similar for most of social media platforms because they adjust them in relation to each other. In this matter, it is Facebook, who sets the standard for other platforms and the ecosystem as a whole.
[52] Clay Shirky has coined it "publish-then-filter" model: "The media landscape is transformed, because personal communication and publishing, previously separate functions, now shade into one another. One result is to break the older pattern of professional filtering of the good from the mediocre before publication; now such filtering is increasingly social, and happens after the fact." Retrieved from: (Shirky, Publish, then filter , 2008).

enforcement process. Therefore, the reality is that undesired content will always be there, even if it is for a short period of time. The challenge of policing is too big for these platforms to handle in advance. Rather than being proactive, it is a reactive process: "even heinous content may get published, at least briefly, and criminal behavior may occur (and have its intended impact) before anything is done a response" (Ibid).

Jonathan Zittrain exposes the complexity of the situation by comparing it to the automotive industry: "if a company could repair the car as it's crashing - that's what Facebook is asked to do. Though tweaking speech on the fly, of course, has implications beyond improving product safety" (The Atlantic , 2017). This issue raises various legal questions and ethical concerns. What to do when the presence of criminal behavior is intrinsic to social media platforms? What is the responsibility of the platforms in the allowance and/or dissemination of this content?

The question that follows is: how do platforms approach already published undesired content? Facebook counts with two methods: it either removes the content or marks it as such and guides users on how to avoid it. The tension is again present: an impossible balance between protecting speech and committing to the protection of its users from something offensive or inappropriate. Removal is a blunt tool that directly interferes with speech. Inconsistency in its application can also be seen as hypocrisy and bias on the side of Facebook. Moreover, users might feel personally judged if what they consider a positive expression of themselves is removed. Gillespie gives the example of women that, while exposing their bodies, they post things that represent physical triumph such as giving birth, breastfeeding, or surviving a mastectomy surgery; and this content is removed because it is depicted as inappropriate or pornographic (Gillespie, Governance of and by platforms, 2017). In a way, social media content moderation policies depict society and its concerns. Thus, disagreements over acceptable content have become opportunities for substantive public debate. The #Freethenipple[53] campaign serves as an example of how social media regulations have triggered important discussions that transcend the online world.

Marking content as sensitive but allowing it to remain in the site is a less intrusive measure. Still, platforms may be criticized from being to permissive or harboring pornographers and terrorists (Ibid). In order to provide users with the experience they each seek, social media companies have, again,

---

[53] The censorship policy on female nipples (and not male) in social media platforms sparked a greater discussion on gender equality and femininity. Women in social media photoshopped male nipples to get around the strict policy and the social movement surpassed digital platforms when female activists walked topless the streets of New York City.

made use of algorithmic filters. They patrol users and spare unsafe material from those who want a "clean experience". Yet, as every algorithmic process, this system has proven to be flawed. The social media platform Tumblr blocked the tag #gay arguing that it was generally linked to pornographic content. However, the measure resulted as well in the blocking of all other posts that were not pornographic in nature and caused retaliation from the LGBT community[54]. Likewise, in order to restrict minors in their access to pornography, the platform of Flickr made certain countries' access to the site only possible in "safe mode". Although intending to respond to regional demands, the law protecting minors became technical measures restricting adults (Ibid). In another case, upon receiving complaints about a Facebook page inviting people to submit drawings of the prophet Muhammad, the platform restricted access to these pages in certain countries. The content was not removed but, instead, made invisible to those who considered it offensive[55].

User governance: the practice of "flagging"

Corporations like Facebook have to deal with conflicting worldviews and constant conflict. Yet, being the mediator has proven to be a difficult task and one too large to be covered entirely by the platform itself. As a solution the flagging mechanism was developed, an alternative method of content curation that allows users to "flag" material that they deem offensive. By handing over some of these chores to users themselves, platforms are alleviated from the huge burden of curating vast collections of UGC. What is more, flagging serves as a justification for platform owners to remove content, easing the process. Finally, users participate (or appear to do so) in governance of social media platforms.

But user feedback is collected and processed just as any other expression: it is transformed into data points. And while flags represent a willingness to listen to users, deliberation of flagged content are highly obscure. Moreover, users are given little room to articulate their concerns. Flags can only signify vocabulary of complaint. They are a basic gesture of objection but cannot depict the degree of concern, or the reasons why people might be flagging the content. One cannot indicate that something is troubling but nonetheless worth preserving. "Flags do not allow a community to discuss the concern, nor is there any trace left for future debates" (Crawford & Gillespie, 2014).

---

[54] *Yahoo Censors Tumblr's Gay And Lesbian Tags* — #Why? Refinery29 (2013). Retrieved from:
 https://www.refinery29.com/2013/07/50340/tumblr-yahoo-blocking-lgbt-tags
[55] *Pakistan lifts Facebook ban but 'blasphemous' pages stay hidden.* The Guardian (2010). Retrieved from: https://www.theguardian.com/world/2010/may/31/pakistan-lifts-facebook-ban.

Another issue to be concerned about is that users can use this mechanism not just for content that violates the platform's rules, but for personal ideological reasons. In fact, it is a widely and systematic practice. For instance, content expressing homosexuality might be flagged, expressing homophobia or hatred instead. Flags are not only deployed as a policing mechanism but also as a tactic tool for online bullying or harassment. Inappropriate use of this reporting tool undermines the value of flags as the community's expression of offense and undercuts its legibility as a sign of the community's moral temperature (Ibid).

In their paper, Crawford & Gillespie emphasize the need for transparency in decision making processes by platforms owners. Algorithmic datafitcation of complaints and opaque processes undermine the intended purpose of flagging, the empowerment of users. In addition, while flags allow users to point out what needs to be removed, there is no way of discussing what needs to remain. "Without a public record or space for constetation, flags leave us little chance to defend the right for those things to be said, or to rally agaisnt something so egregious that it warrants more than a quiet detention"(Ibid).

All in all, platforms have obligations not just to the individual who posts and the individual who views, but the broader public life constituted by their interaction. Yet, the practice of flagging may fall short with regard to their obligations to public discourse.

2.5 Reconceptualizing the *public* and the *private*

As put forward in the previous chapter, it is my view that social media practices that pose a threat to human rights are rooted in the overlap of the public and the private. Among them, the online business model and platform content regulation policies are particularly interesting for this discussion.

Facebook is for many people the only source of online information. It is often equated to the Internet, as if there was nothing else in the online realm besides this site. Indeed, the platform provides a space for users, companies, media, and government officials alike. Yet, the platform is more than a mere intermediary for communications between all of these actors. Facebook accumulates vast collections of information and designs the space and conditions under which users engage in public discussions. Because SNS are operating as public spaces, questions are being raised on whether relevant protections should be afforded in these environments and whether platform owners should bear this duty.

When United States President Trump blocked some Twitter users from his account, the actions where declared unconstitutional by court on the grounds that this account was being used as a public forum and blocked users were unable to take part in public life[56]. It was a clear violation of their right to access information. The distinction between private and public spaces is disappearing in social media platforms as corporations hold more power and they take on public duties.

The private curated walled gardens of social media have increasingly replaced other free public spaces. This overlap of public and private results in users thinking they are operating in a somewhat public sphere where rights and freedoms are guaranteed[57]. What happens when platforms monetize our interactions transforming the public sphere into a private money machine? (Leetaru, 2017) What are the implications of having public speech online privately governed?
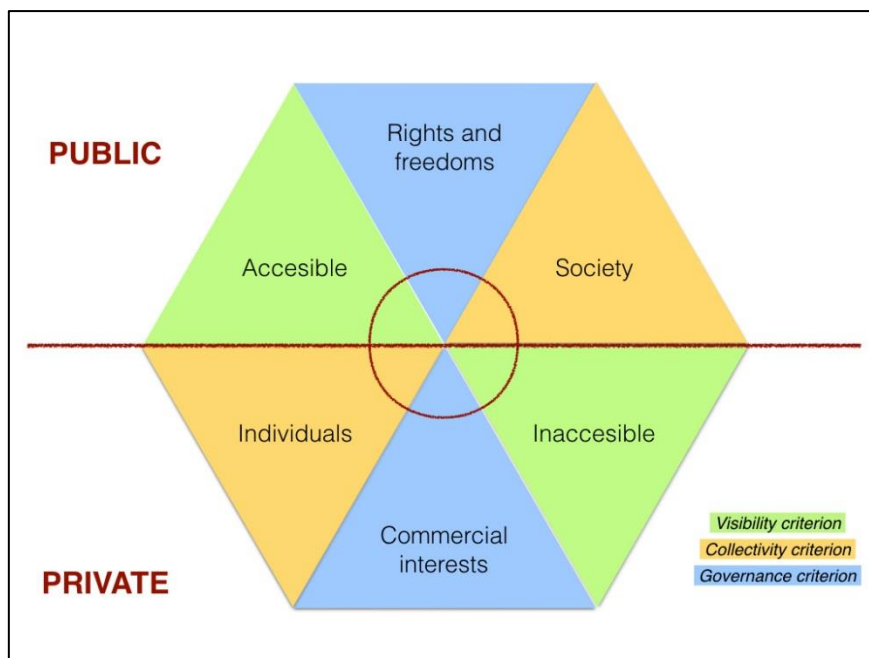

Figure 1: Diagram of the overlapping of the public and the private (self-made).

---

[56] See: *Trump's Blocking of Twitter Users Is Unconstitutional, Judge Says.* The New York Times (2018). https://www.nytimes.com/2018/05/23/business/media/trump-twitter-block.html?login=smartlock&auth=login-smartlock

[57] Contrary to the claim that social media are private spaces, studies show that users view and treat online social networks as public venues. See: (Burkell, Fortier, Yeung Cheryl Wong, & Lynn Simpson, 2014).

The public/private criteria developed in the first chapter clarifies the issues at stake and serves as basis for the analysis on impacts to human rights. Namely, the right to privacy and freedom of expression.

Right to privacy

The depiction of Facebook as a <u>public</u> forum upholds its business model. Firstly, by promoting openness and typifying privacy as negative, social network sites appear accessible and transparent *(visibility criterion)*. Secondly, the platform claims that it operates in good faith, in the interest of society, the users of the world, the collective *(collectivity criterion)*. The business presents itself as an engine that aims at making the world a better place underpinned by a mantra of a more open and connected world. This imaginary confuses users by giving them the feeling that the platform operates in the public realm. As a consequence, the expectations are that rights and freedoms are leading Facebook's agenda. In fact, the platform narrative is riddled with human rights language making it seem that way *(governance criterion)*.

Unfortunately, the reality is quite another. While connectivity (user sharing of data) is aligned with openness; connectedness (platform sharing with third parties) is hidden to the customer. Transparency is one sided: encouraged on the side of the user but minimal by the corporation itself *(visibility criterion)*. Furthermore, their claim of having a social mission is only true to the extent that it benefits the company. As a private corporation, Facebook is driven by profitability. And because often a connected world is in line with these interests, the company's strategy is to claim that their operations look out for the well-being of all its users *(collectivity criteria)*. Lastly, social media platforms do not operate in the human rights protection framework that states are bound by. For this reason, it would be naive to think that rights and freedoms are their main concern. As any other private business, information technology companies are driven by market forces *(governance criteria)*.

Platform's business model is based in the collection, process, and exchange of personal data with third parties. That is to say that all interactions between individuals in these sites constitute products that feed Facebook's business model. As Jørgensen indicates, contributing to the online economy via online expressions, habits, and preferences has become a premise for participation in the networked public sphere (Jørgensen, Human Rights and Private Actors in the Online Domain, 2018). This way, to take part in public discussions, users have no option but to hand over their personal information. However, the use of this data by platforms might be in violation of their human right to privacy.

Freedom of expression

"Whereas the ability of individuals to exercise privacy rights is closely connected to the online business model, the boundaries for freedom of expression are defined in a gray zone between legal frameworks and company norms" (Jørgensen, What Platforms Mean When They Talk About Human Rights, 2017). Objectively, Facebook today is the online space where individuals come together to discuss important issues that concern society (*collectivity criterion*). In this sense, the platform is the means for the exercise of our public life. However, this is not an all-free space. Instead, there are prerequisites to take part in it and limits to what can be said. "Social media platforms do much more than transmitting what we do, they constitute what we see" (Gillespie, Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media., 2018). It is through content moderation that, hidden behind a veil of neutrality, platforms interfere with freedom of expression. By curating content in their sites, social media platforms are not just providers of spaces for online activity, but keepers of the public discourse (Crawford & Gillespie, 2014).

In close relation to the redefinition of privacy, owners of SNS are drawing the line between what constitutes free speech and what should be censored. Doing so, platforms pick and choose the content that will be accessible, while the quietly remove the rest (*visibility criterion*). Similar to Orwell's Ministry of Truth were government approves versions of events, Facebook curates speech in a very obscure process, in which users have no say, nor eyes. In order for this model to survive, the platform needs to be portrayed as the freedom of expression "gatekeeper". Yet human rights standards are iterated as a benchmark for take down of content only in relation to government requests (Jørgensen, What Platforms Mean When They Talk About Human Rights, 2017). Maintaining user confidence is key for the survival of platforms and their business model. Thus, human rights are only invoked when it serves the interest of platform owners. Decision making on content moderation, and everything else in the platform, is driven by commercial interests (*governance criterion*).

In a nutshell, legally speaking, social network sites like Facebook do not constitute public forums, but rather private property made open to the public: "the online sphere holds no public streets or parks" (Jørgensen, 2018). The overlap of the public and private distinction in social media misleads the user into thinking he/she is operating in a public space where freedoms and rights are guaranteed. However, commercial drivers of corporations such as Facebook challenge the basic assumptions and freedoms associated with public life (Ibid). On one hand, social media companies have the power of information and, on the other, they set the conditions by which public discussions develop. These

private actors have human rights impacts but yet, they operate only in a regime of guidelines and soft law. The lack of legal human rights obligations on the side of corporations such as Facebook presents a governance gap that needs to be addressed.

Chapter 3 – The Protection of Human Rights in Facebook

This chapter examines the state of affairs of human rights regulations of social media corporations, using Facebook as the example. First of all, existing frameworks are considered including private initiatives, non-profit efforts, international soft law instruments and regional legal frameworks. Secondly, the main drawbacks of the current human rights regulatory framework are pinned down: (1) its soft law nature, (2) its general focus and (3) the absence of remedy channels. And, finally, the chapter concludes by advancing Emily Laidlaw's governance model as a potential way forward for a stronger protection of human rights in social media platforms.

3.1 Human rights regulatory framework of social media platforms

As we have seen, social media platforms have an impact on human rights. What is more, because these companies operate globally, their human rights impacts are global too. Traditionally, the protection of human rights was conceived as duty of states only, which created a regulatory gap on human rights abuses that took place in the private sector. It is for this reason that there are increasingly more international efforts attempting to include private actors in human rights regulatory frameworks. Present instruments that highlight human rights obligations of businesses take different forms: corporate social responsibility projects, non-profit initiatives, soft law tools and even hard law instruments in some regions. These are the most significant:

Corporate Social Responsibility: The Global Network Initiative (GNI)

The scale of human rights expectations of social media corporations is on the rise because technological developments are making human rights risks very significant in the industry[58]. There are a number of corporate social responsibility (CSR)[59] initiatives that aim at adhering business practices to human rights standards, being the Global Network Initiative a significant one[60].

The GNI is a multi-stakeholder effort launched in 2008 (on the 60th anniversary of the Universal Declaration of Human Rights) to protect and advance freedom of expression and privacy in the ICT

---

[58] Human rights systems have traditionally treated human rights as a government responsibility but the blurring of the public-private divide is the fissure in which corporate social responsibility has been flourishing. Retrieved from: (Laidlaw, 2015).

[59] The European Commission has defined CSR as the responsibility of enterprises for their impact on society. See: https://ec.europa.eu/growth/industry/corporate-social-responsibility_en

[60] Although not included in this thesis for the sake of conciseness, there are other major international initiatives in the ICT industry concerning human rights such as the Global e-Sustainability Initiative (GeSI), and the Industry Dialogue. Retrieved from: (Laidlaw, 2015)

industry by setting a global standard for responsible company decision making[61]. It was a response to increased governmental pressures on ICT companies and a lack of guidance to resolve these tensions while, at the same time, protecting and advancing online privacy and expression[62]. With a focus on human rights, the GNI is composed by a multi-stakeholder group of ICT companies, civil society organizations, investors, and academics. After a two-year deliberation process[63], they developed a framework to unite ICT businesses around core values for freedom of expression and privacy online. Some of its current members include Facebook[64], Google, and Microsoft (companies), Human Rights Watch (civil society) and Berkman Klein Center for Internet & Society at Harvard University (academic organizations).

The initiative aims at easing the untenable position in which business must balance their obligation to respect local law with their responsibility to protect the rights of their users (Maclay, 2010). The situation is quite complicated because it transcends the issue of transnational companies following (or not) regional laws. It is also about understanding how and when to challenge the existing law to mitigate human rights risks. Through the Global Network Initiative, companies get together with human rights groups and academics to share and develop best practices and strategies for new media businesses to push back when governments make demands to remove content or requests for user information. The belief is that governmental pressure is in conflict with both, core elements of their business, and their users' fundamental rights to privately communicate and access information.

Yet, it is important to note that the GNI is not simply a form of altruism but rather it embodies the recognition of serious business' problems that have profound implications for all human rights (Ibid). In fact, the trade of near-term profits for human rights principles results in long-term gains for business in this industry. Responding responsibly to government requests and putting users' rights first is, not only in line with human rights standards, but also beneficial for the company's public brand and growth. In the case of ICT companies, challenges are often presented with respect to

---

[61] This is GNI's Mission statement, retrieved from the Global Network Initiative website: https://globalnetworkinitiative.org/team/our-mission/

[62] More specifically, two particular incidents pushed for the formation of the GNI: "First, Yahoo! handed over information on its email account holders to the Chinese authorities, thereby exposing the identity of a Chinese journalist and leading to its imprisonments for ten years. Second, Google launched a version of its search engine in China that censored search results." Retrieved from: Laidlaw, *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility,* 2015.

[63] The founder company members of the GNI were only Yahoo!, Google and Microsoft.

[64] Facebook joined as a full member in 2013, which entailed an enormous boost for the organization. Until then, it held observer status so it did not submit its operations to a human rights audit. The same year, mass online surveillance by governments was revealed by Edward Snowden.

freedom of expression and privacy. This is the reason why the GNI focuses exclusively on these two, though recognizing the interdependence of all human rights.

The Global Network Initiative is an ambitious project. In the end, it is a self-governing initiative that require its participants to <u>not</u> comply with the law when it conflicts with international human rights standards. Maclay attributes the tension between governments' abuse and corporate restrictions to the unprecedented technological progress that we are undergoing: "We are caught between the proverbial rock and a hard place, particularly when society undergoes the complicated process of integrating, adapting, and shaping the role of new media" (Ibid).

The GNI has four foundational documents: 1) the GNI Principles, 2) the Implementation Guidelines, 3) the Accountability, Policy and Learning framework, and 4) the Governance Charter[65]. In the *Principles,* companies articulate a common understanding for freedom of expression and privacy online, as drawn from the International Bill of Human Rights.[66] They recognize the complementary responsibility that ICT companies have, along with governments, in respecting, protecting, and promoting human rights. Principles are applied based on the UN Guiding Principles on Business and Human Rights, addressed later in this chapter. In addition, the following key elements are highlighted: responsible corporate decision making; multi-stakeholder collaboration; and governance, accountability, and transparency. The *Implementation Guidelines* are constituted by actionable steps for companies to comply with the Principles just mentioned. Essentially, participants evaluate human rights risks and seek opportunities to mitigate them when considering whether and how to enter a new market (Ibid). In other words, they undertake human rights impact assessment to identify potential threats to expression and privacy before they take place and work towards their minimization. This process has shared practices at its core and it is drawn upon collective experiences. Thirdly, the *Accountability, Policy & Learning Framework* provides the groundwork for basic institution-building responsibilities. These include independent monitoring, advocacy work, corporate accountability, and remediation where necessary. Finally, the *Governance Charter* defines the organizational structure of the Initiative and oversees its development.[67] (Ibid)

---

[65] These documents are available at https://globalnetworkinitiative.org/core-commitments/

[66] The International Bill of Human Rights is constituted by the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the International Covenant on Economic, Social, and Cultural Rights (ICESCR).

[67] This information has been retrieved from *MacLay,2010* and the website of the Global Network Initiative itself: https://globalnetworkinitiative.org/governance-charter/

The Global Network Initiative is certainly a promising breakthrough endeavor as it represents its members' commitment to free speech and privacy. Yet, it consists of a voluntary framework with no binding obligations on companies and hence, it falls short on the protection of human rights. It is for this reason that Amnesty International, for example, decided not to join the GNI: "Following careful consideration of these documents, Amnesty International has come to the conclusion that – while they represent a degree of progress in responding to human rights concerns – they are not yet strong enough to allow Amnesty International to endorse them[68]". Its lackluster participation is also object of criticism. The initiative lacks cultural and geographically diverse participation and it only counts with eleven ITC corporations with the glaring absence of companies like Twitter. In addition, the GNI is criticized for the absence of accountability or remedy mechanisms. Jørgensen accurately suggests that "these criticisms speak to the inherent challenge of having an industry define its own standards and procedures for respecting users' rights to privacy and freedom of expression" (Jørgensen, Human Rights and Private Actors in the Online Domain, 2018).

But perhaps the most important objection to the GNI is its state-centered scope. The Initiative is limited to the negative human rights impacts caused by external pressure from governments. It is focused on the state pressure that companies may be subjected to when they operate in contexts where the national legal framework does not comply with international human rights standards (Ibid). Yet, there is a lack of coverage of these rights when it comes to the companies' governance of content or tracking of user behavior: the focus is on the vertical conflict rather than in the relations between private actors, the horizontal conflict. This approach indicates that the human rights discourse by ICT companies tends to highlight push-back strategies against illegitimate government requests, with less attention being paid to the human rights impact of the company's own actions (Ibid).

Non-profit initiatives: Ranking Digital Rights (RDR)

Ranking Digital Rights is a non-profit research initiative[69]. Contrary to the GNI, it does not include ICT companies among its members. Instead, its partners are academic institutions and non-governmental organizations worldwide. It is for this reason that the human rights impact assessment

---

[68] See: *Amnesty criticizes Global Network Initiative for online freedom of speech.* The Guardian, October, 30th, 2008. Retrieved from: https://www.theguardian.com/technology/2008/oct/30/amnesty-global-network-initiative
[69] A similar major non-profit initiative is European Digital Rights, though not considered in this thesis for the sake of conciseness. It is the biggest European network and thought leader defending rights and freedoms online. Their mission is to promote, protect and uphold the right to privacy, data protection, freedom of expression and information, and the rule of law. For more information, see: https://edri.org/.

of ICT companies is broader, including human rights threats by companies' own practices as well. RDR evaluates the world most powerful Internet and telecommunications companies on freedom of expression and the right to privacy and it works out a global ranking system.

It does so by developing a "Corporate Accountability Index" based on thirty-one indicators focused on corporate disclosure of policies and practices that affect users' freedom of expression and privacy[70]. The Index aims at boosting corporate transparency and public scrutiny of business practices, which will encourage companies to strengthen their respect and protection of human rights. What is more, it provides measurable steps for platform owners, and other ICT companies, to improve the way in which they disclose information on policies and practices that affect human rights.

Ranking Digital Rights provides an overall score that signals companies' public commitments and disclosed policies affecting users' freedom of expression and privacy. This is important because the lack of disclosure on how platforms handle user information and their role in policing the flow of information online make difficult to assess the privacy, security and human rights risks of using these services. Facebook received a score of 55, being a 100 full disclosure and a score of 0 being non-disclosure at all. It has ranked fourth in the Index, disclosing less about policies affecting these rights than most of its U.S peers. (Ranking Digital Rights, 2018)
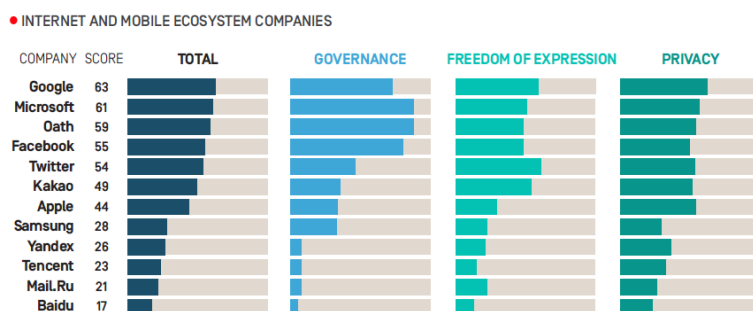

Figure 2: Ranking Digital Rights, 2018 Global Accountability Index, company score

In its 2018 Corporate Accountability Index, RDR found that Facebook is one of the least transparent social media companies when it comes to personal data and privacy issues: "Facebook provided users with limited options to control what information the company collects, retains, and uses, including for targeted advertising, which appears to be on by default" (Ranking Digital Rights, 2018). The

---

[70] Information retrieved from the Ranking Digital Rights website: https://www.newamerica.org/oti/ranking-digital-rights/about-ranking-digital-rights/

following graphs depicts the score on users' control over their own user information, being Facebook in the worst position with a score of 13%.
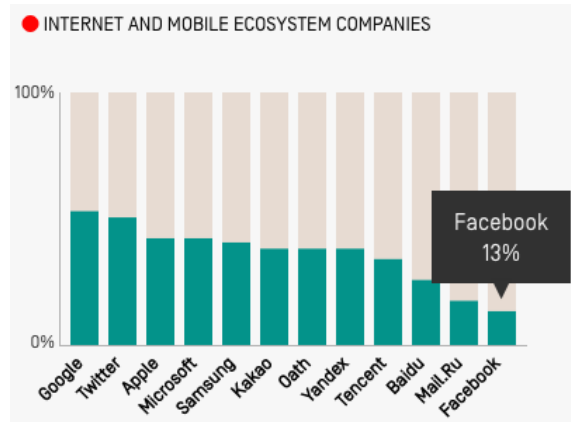

Figure 3: Ranking Digital Rights, users' control over their own user information

Furthermore, the 2018 Index results highlight a persistent and widespread lack of transparency by companies around the policing of content— especially about their terms of service enforcement (Ranking Digital Rights, 2018). The following graph shows the companies' score on disclosure about the volume and nature of content and accounts restricted for violating the company's rules. Although Facebook is in the third place, it only scored 4%.
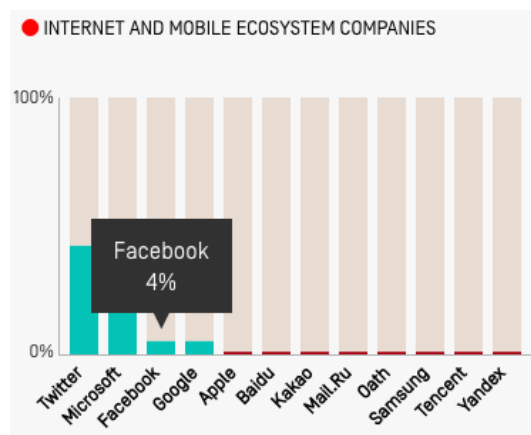

Figure 4: Ranking Digital Rights, Data about terms of service enforcement

Ranking Digital Rights is not a CRS initiative but its research is significant enough to have an impact and be considered a key effort for the protection of human rights in the ICT industry. Furthermore, its annual report includes recommendations to these companies on how to strengthen the protection of human rights in their business. Among other suggestions, it recommends companies to invest in

the development of new technologies and business models that strengthen human rights, and to publish comprehensive transparency reports[71].

Soft law – CSR: United Nations Global Compact (UNGC)

When discussing international efforts on business and human rights, one must mention the UN Global Compact[72]. It is the world's largest voluntary corporate sustainability initiative and it constituted a stepping stone in the business and human rights agenda. However, its focus is much broader than human rights, including as well environmental and anti-corruption principles. It was launched in 2000 under the leadership of UN Secretary General Kofi Annan to bring together companies and encourage sustainable practices in the private sector. The initiative counts with over 12,000 signatories representing nearly every country. Yet, as Laidlaw points out, the Global Compact is not popular among ICT companies and hence, not as relevant in the arena of internet governance (Laidlaw, 2015).

The UNGC is a ten-principle based framework for businesses, with only two being specifically directed to human rights: "Businesses should support and respect the protection of internationally proclaimed human rights[73]; and make sure that they are not complicit in human rights abuses[74]" (UN Global Compact, 2000). It offers no concrete guidance on companies' obligations[75]. The vagueness of the text facilitates *bluewashing,* or the practice of joining international organizations' initiatives for public relations purposes rather than actual commitment to the signed document, which constitutes its major criticism.

All in all, the Global Compact has done a poor job in setting up an effective governance system for businesses to comply with human rights obligations. However, it opened the conversation on business and human rights and paved the way for latter efforts such as the UN Guiding Principles on Business and Human Rights.

Soft law: OECD Guidelines for Multinational Enterprises

---

[71] More specifically, "companies should publish regular information and data on their official websites that helps users and other stakeholders understand the circumstances under which personal information may be accessed, speech may be censored or restricted, or access to service may be blocked or restricted" (Ranking Digital Rights, 2018).

[72] Find more information on the United Nations Global Compact here: https://www.unglobalcompact.org/what-is-gc

[73] Principle 1 of the UN Global Compact.

[74] Principle 2 of the UN Global Compact.

[75] The Global Compact makes use of the concept "sphere of influence", in the sense that businesses are responsible for human rights within their sphere of influence. Yet, this term is not defined in detail by human rights standards. R Retrieved from: Jørgensen, *Human Rights and Private Actors in the Online Domain*, 2018.

The Organization for Economic Co-operation and Development (OECD) developed a set of Guidelines for Multinational Enterprises[76]. These are recommendations by governments to multinational companies operating in or from OECD countries on responsible business conduct in a global context consistent with applicable laws and internationally recognized standards (OECD, 2018). The non-binding principles cover a range of issues beyond human rights such as environment, taxation, competition or consumer interests. Although they were first adopted in 1976, they have been revisited and upgraded a few times, the latter one being 2011 to align themselves to the UN Principles on Business and Human Rights.

The main reason why the Guidelines are worth mentioning is because, contrary to any other framework, they count with a body that, among other things, looks over the functioning and implementation of the guidelines: the OECD Investment Committee. The Guidelines are still considered as a soft law tool, but their observance can be followed up and even complained against through the National Contact Points[77] set up by governments of OECD members.

Soft law: United Nations Guiding Principles in Business and Human Rights (UNGPs)

In 2011 the Human Rights Council unanimously endorsed the Guiding Principles on Business and Human Rights[78], being now the most significant document on this issue in the UN framework. The aim and main focus of the text is to implement United Nations "Protect, Respect, and Remedy" Framework. Yet, the principles do not introduce a new approach or obligations for businesses. Instead, they serve as clarification of the respective duties and responsibilities that states and businesses already have and elaborate on them: "Nothing in these Guiding Principles should be read as creating new international law obligations, or as limiting or undermining any legal obligations a State may have undertaken or be subject to under international law with regard to human rights" (UN Guiding Principles on Business and Human Rights, 2011). Thus, the UNGPs provide an authoritative global standard for better preventing and addressing the risk of adverse impacts on human rights linked to business activity (Office of the High Commissioner for Human Rights, 2014).

---

[76] The OECD Guidelines for Multinational Enterprises can be found here:
http://www.oecd.org/daf/inv/mne/48004323.pdf
[77] National Contact Points are set up by governments to further the effectiveness of the OECD Guidelines by undertaking promotional activities, handing handling enquiries, and contributing to the resolution of issues that may arise from the alleged non-observance of the guidelines in specific instances. See: https://mneguidelines.oecd.org/ncps/
[78] The Human Rights Council endorsed the Guiding Principles in its resolution 17/4 of 16 June 2011.

The UNGPs were drafted by UN Special Representative on Business and Human Rights, John Ruggie, and are the product of six years of extensive research and consultations involving governments, companies, business associations, civil society, affected individuals and groups, investors and others around the world (Ruggie, UN Guiding Principles for Business & Human Rights, 2011). They were triggered by increasing stakeholder attention resulting from globalization and developed as a response to diverse industry initiatives. Aiming at unifying efforts and establishing a global standard, the Guiding Principles take a multi-stakeholder approach. They recognize that responsibilities are distinct but complimentary between states and companies. What is more, they enjoy wide support across states and the private sector.

There are 31 Guiding Principles[79], each of them followed by a commentary that clarifies its meaning and implications. The UNGPs rest on three pillars: protect, respect and remedy. The first one is <u>the state duty to protect human rights</u> and it refers to the obligations that states have to protect against human rights abuses committed by third parties through appropriate policies, regulation, and adjudication (Ruggie, UN Guiding Principles for Business & Human Rights, 2011). The second pillar is <u>the corporate responsibility to respect human rights</u>. John Ruggie introduced the business concept of "due diligence" into the human rights protection discourse. Business enterprises must carry out human rights due diligence to identify, prevent, mitigate and account for how they address their adverse human rights impacts[80]. The clarification of the concept of corporate human rights due diligence was a game-changer and a major contribution to the Guiding Principles[81.] It provided a common understanding of the essential elements that are required to operationalize respect for human rights in corporate practice (UN Working Group on Business and Human Rights, 2018). The use of this term appears to be a clever and deliberate tactic, as it is familiar to business people, human rights lawyers and states, among whom Ruggie sought to build a consensus on his approach (Bonnitcha & McCorquodale, 2017).

Thirdly, there is the <u>access to remedy</u> for victims of business related abuse to seek redress. The UNGPs include judicial and non-judicial remedies as well as state and non-state based grievance mechanisms. In his report, Ruggie explained the interrelatedness of the three pillars, each being an

---

[79] Principles 1 to 10 fall under the first pillar, Principles 11 to 24 belong to the second pillar and Principles 25 to 31 are under the third pillar. See: UN Guiding Principles on Business and Human Rights, https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf
[80] This is Principle 17 from the UN Guiding Principles on Business and Human Rights.
[81] Five of the 31 Guiding Principles appear under the tittle Human Rights Due Diligence and two others (Principles number 14 and 15) refer to due diligence.

essential component of a dynamic system: "the State duty to protect because it lies at the very core of the international human rights regime; the corporate responsibility to respect because it is the basic expectation society has of business in relation to human rights; and access to remedy because even the most concerted efforts cannot prevent all abuse" (Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business entreprises, John Ruggie, 2011).

The Guiding Principles operate in a grey zone between human rights law and soft law, mixing regulatory and voluntary approaches[82]. Its non-binding nature is a matter of discussion, seen as a strength by some and as its biggest weakness by others. It is argued that this framework does not set a sufficiently high standard for business because it is not legally enforceable and it lacks an accountability mechanism. Conversely, supporters claim that the Framework is remarkable because it achieved more consensus that any previous attempt at creating a global standard on business and human rights[83]. It is Ruggie's view that the Guiding Principles, as a soft-law instrument, is more appealing than a hard-law punitive approach and, while they do not bring business and human rights challenges to an end, they "mark the end of the beginning" by establishing a global platform for action (Ibid). On the contrary, a treaty would have to be developed in such a high-level of abstraction that it will diminish its applicability and effectiveness.

Ruggie recognizes that when it comes to their implementation, one size does not fit all: "the Guiding Principles are not intended as a tool kit, simply to be taken off the shelf and plugged in" (Ibid). While they are universally applicable, implementation must adapt to each industry and national context. Accordingly, the UN working group on human rights and transnational corporations and other business enterprises has produced a "Guidance" on the development of national action plans on business and human rights[84].

---

[82] Whereas pillars one and three combine existing state obligations under international human rights law with soft law recommendations, pillar two is soft law only. This reflects the lack of direct human rights obligations for companies under international human rights law. Retrieved from: Jørgensen, *Human Rights and Private Actors in the Online Domain*, 2018.

[83] Preceding the UNGPs, there was attempts to impose on companies, directly under international law, the same range of human rights duties that States have accepted for themselves under treaties they have ratified (Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business entreprises, John Ruggie, 2011). The initiative was called the *Norms on Transnational Corporations and Other Business Enterprises.* The final draft received significant opposition and it was determined by the Commission on Human Rights to have no legal standing.

[84] Find the Guidance on National Action Plans on Business and Human Rights here: https://www.ohchr.org/Documents/Issues/Business/UNWG_%20NAPGuidance.pdf

Many Internet platforms, as is the case of Facebook, advance the implementation of the UNGPs through collective industry initiatives like the Global Network Initiative. However, as discussed earlier in this chapter, the GNI has a limited focus on human rights matters arising from vertical conflicts. That is to say, it only looks at the relationship between ICT companies and governments but lacks assessment on negative impacts caused by the platforms' own actions. Companies involved praise the GNI because it is only about protecting users from government interferences. Yet, it appears as if they are not willing to engage in conversations about their own online business models violating human rights.

Despite its public commitment to human rights through multi-stakeholder initiatives, Facebook lacks a human rights policy per se, for which it falls short on the standard of conduct laid out in Principle 16 of the UNGPs. Hence, the existence of internal due diligence processes is unknown and doubted. It can be argued that Facebook, along with other platforms, have not lived up to its corporate responsibility to respect human rights as prescribed in the UNGPs.

Hard law: General Data Protection Regulation (GDPR)

The General Data Protection Regulation[85] is one of the strongest and most comprehensive endeavors globally to regulate the collection and use of personal data[86]. The law was adopted by the European Union in April 2016 and it came into force on May 25th this year. It was the most contested law in the E.U.'s history and a product of years of intense negotiations and thousands of proposed amendments (Powels, 2018). Yet, the result was outstanding as GDPR requirements apply equally to the public and private sector, including in this way businesses, governments and non-profit organizations.

The GDPR introduces three fundamental changes in relation to the previous data protection framework, the 1995 Data Protection Directive. First of all, it harmonizes data practices across Europe. The new regulation has an extra-territorial applicability as it applies to all companies dealing with personal data of individuals residing in the Union, regardless of the company's location and whether the processing takes place in the EU or not (Business & Human Rights Resource Centre, 2018). Secondly it is an enforceable regulation which means penalties are applied if not followed.

---

[85] Find the General Data Protection Regulation here:
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN
[86] The GDPR Article 4 defines personal data as "any information relating to an identified or identifiable natural person."

Regulators are able to fine any business in breach of the GDPR as much as four per cent of its total worldwide sales. Thirdly, the conditions for consent have been strengthen. According to the new regulation, consent now must be clear and provided in an intelligible and accessible form, using clear and plain language (EUGDPR, 2018).

Lots of things could be said and further examined regarding this innovative legislation but the aim of this thesis is to look at the bigger picture to better reveal why the GDPR sets a great example in the protection of human rights in the digital realm. Probably the biggest achievement of this piece of legislation is that it has managed to shift "the default is social" ideology behind platforms to "the default is privacy" philosophy. This way, it has given priority to human rights[87] over commercial interests, setting an important precedent for other regions in the world and other human rights lacking effective protection at the corporate level.

The GDPR also opens the discussion on whether the state has a positive obligation to legislate the obligations of companies. The European Union has done so in this case regarding issues of privacy. Yet, the right to privacy is not secured elsewhere. In fact, the entering into force of the GDPR has exacerbated how much weaker data protections regulations are in the United States. On a positive note, it is likely that this new regulation will gradually become a de facto global standard because it applies to any organization that collects or processes the data of EU citizens, regardless of where the organization is based or where the EU data is processed (Human Rights Watch, 2018)[88].

This regulation obviously has an impact on the right to access information and freedom of expression. For instance, it expands what has come to be known as "the right to be forgotten", or erasure of data that is no longer relevant to original purposes or has its consent withdrawn by the data subject. Under the GDPR, individuals have the right to ask companies to delete information on certain cases such as the ones just mentioned.

The European General Data Protection Regulation is a vital step towards privacy protections in the private sector and more precisely online platforms. In despite of being a core affected right, freedom

---

[87] More specifically it protects the right to private life as guaranteed by Article 7 of the Charter of Fundamental Rights. Find this Charter here: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:12012P/TXT&from=EN

[88] Appearing before Congress, Mark Zuckerberg, Facebook's CEO replied with a clear "yes" when asked if he commits to extends the same protection to Americans that Europeans receive under the GDPR. Yet, it was a misleading promise because the protection is only partially assured: some features designed to comply with the European regulation will be globally applied while other regulatory steps is expected to be only applied in the EU. Find Zuckerberg's testimony here: https://www.youtube.com/watch?v=mZaec_mlq9M

of expression has not received the same attention. No hard law regulation looks over content moderation policies and the human rights implications of social media platforms acting as curators of public speech. Jørgensen argues that, in fact, expression is less protected in the online domain: "Private companies in control of communicative platforms are free to decide the types of speech they support. This includes taking down or blocking and filtering expression that would otherwise be protected by the First Amendment" (Jørgensen, Human Rights and Private Actors in the Online Domain, 2018).

3.2 Shortcomings of the existing instruments

In the words of David Kaye, "human rights law gives companies the tools to articulate their positions in ways that respect democratic norms and counter authoritarian demands" (UN Human Rights Council , 2018). Still, the above-mentioned frameworks, which govern activities of Internet platforms, present a number of shortcomings in their implementation. These have been somewhat covered throughout this thesis but, for clarification purposes, they deserve a section of their own.

1. Soft law nature

In exception of the GDPR, all existing efforts to make businesses responsible of human rights are considered to be soft-law instruments. Hence, the first drawback is that they are voluntary by default[89]. In other words, it is up to businesses to take part in the conversation and make commitments to their human rights responsibilities. In addition, the absence of accountability – and sometimes follow-up- mechanisms, complicates the effective implementation of these commitments. Thus, soft law instruments are not filling up the human rights governance gap existing in the private domain. The lack of normative standards constitutes the second important shortcoming. When does a practice by a platform amount to a violation of human rights? Soft law tools do not establish a clear threshold that signals human rights abuses, which results in diverse interpretations and makes it more difficult to establish and sustain human rights standards.

Laidlaw accurately blames the human rights system for treating human rights as a government responsibility and having effectively privatized human rights in the digital environment (Laidlaw, 2015). As private parties, social media platforms run alongside the law while the exercise of human rights by users occurs extensively through them. It is Laidlaw's view that traditional conceptions of

---

[89] On the other hand, it is argued that voluntariness brings businesses to the negotiating table and has more potential to capture the spirit of commitment on the part of companies. As Laidlaw notes, this is particularly important in the field of human rights, where its moral force is often of greater weight than its legal dimensions. Retrieved from: Laidlaw, *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility,* 2015.

human rights as a relationship between citizens and state should be challenged (Ibid) to include private actors – i.e. corporations and citizens or corporations and states.

The main argument against binding instruments is that the imposition of duties would disrupt the market, hampering innovation. This concern is bigger when it comes to technological companies such as online platforms, where innovation appears to be at the core of the business. However, such criticism entirely misses the point. Indeed, it is very likely that the formalization of laws that impose human rights obligations on businesses will disrupt the market. Yet, corporations like Facebook hold great power over the exercise of human rights of individuals worldwide. It is only reasonable that they are governed in a human rights-compliant way, even if it places a burden on profit-making actors.

This discussion can be framed in a bigger one on whether technology should slow down and adapt to society and its norms or rather technological progress should be prioritized and users should catch up[90]. However, this thesis claims that when it comes to human rights there should be no question. The protection of human rights cannot be a reaction to technological progress stepping over these rights. Instead, human rights should be ensured by default and questions on potential threats must take place beforehand. In the words of Molly K. Land, the "it is better to ask for forgiveness than permission" mentality is not compatible with a human rights approach in social media technology[91].

3. General focus

Except ICT-focused initiatives, frameworks on business and human rights are very general in their content. Human rights obligations are vaguely laid out, lacking guidance on concrete steps that businesses are to follow. The imprecise nature of these documents make difficult the implementation of such obligations in the different sectors. This is especially relevant when addressing governance over online platforms because of the impact that they inflict on civil and political rights (right to

---

[90] Ingrid Burrington metaphorically discusses the powers of platforms over society: "power is not only something companies abuse but a force inadvertently unleashed by well-intentioned engineering teams, as pure and innocent as Pandora faced with an irresistible box of technical puzzles or as benevolent as Pandora's brother-in-law Prometheus, simply trying to bring lowly mortals treasures that they were wholly unprepared to wield. What laws can punish the innovation – nay, the generosity – of endowing the world with such accursed gifts? [...] But at the end of the day, the corporate campuses at Silicon Valley hold no gods. They hold merely men who, for now, primarily answer to a court of public opinion for the violence and horror that their platforms repeatedly enable". Read this article here: *Could Facebook be Tried for Human Rights Abuses?*, 2017, https://www.theatlantic.com/technology/archive/2017/12/could-facebook-be-tried-for-war-crimes/548639/

[91] This is part of Molly K. Land's speech during the presentation of her book "New technologies for human rights law and practice" available here: https://www.cambridge.org/core/books/new-technologies-for-human-rights-law-and-practice/A6473E8A4F6A9ED12675E54A03318802

privacy and freedom of expression). There is little elaboration provided on the duties regarding freedom of expression or privacy, and sometimes they are not mentioned at all (Laidlaw, 2015).

Jørgensen identifies the limited attention given to civil and political rights in CSR: "Corporate social responsibility primarily addresses social and economic rights, in particular as it relates to working conditions and environmental and community impact" (Jørgensen, Human Rights and Private Actors in the Online Domain, 2018). However, in the case of online companies like Facebook, responsibilities extend not only to its employees and community but also to its many users who are directly and indirectly affected (Ibid). The lack of attention given to civil and political rights was a main focus of Ruggie when drafting the Guiding Principles. Hence, the text reiterates that the responsibility of business enterprises to respect human rights refers to, at minimum, the human rights expressed in the International Bill of Human Rights[92].

Yet, interpretations of the UNGPs that focus on ICT companies are still flawed in the way they mainly look into human rights concerns resulting from state pressures on companies. This is the case, for instance, of the "ICT Sector Guide on Implementing the UN Guiding Principles on Business and Human Rights"[93] developed by the European Union Commission. In this guide, as in the GNI, the negative human rights impact that may flow from the company's governance of content or tracking user behavior is not addressed (Ibid).

To sum up, human rights responsibilities of social media corporations are often overlooked in international instruments on business and human rights; especially those arising from content regulation policies and use of personal data by platforms owners.

3. Absence of remedy channels

One of the key weaknesses in current corporate governance is the limited remedy channels that exist. In fact, this need was identified by Ruggie as a touchstone of his framework (Laidlaw, 2015) and placed as the focus of the Guiding Principles' third pillar. On the one hand, states should take steps to ensure the effectiveness of domestic judicial mechanisms when addressing business-related human

---

[92] UNGP Principle number 12.
[93] Find the Guide here:
 https://ec.europa.eu/anti-trafficking/sites/antitrafficking/files/information_and_communication_technology_0.pdf

rights abuses[94] and, on the other, enterprises should establish operational-level grievance mechanisms for individuals who may be adversely impacted[95].

Ruggie notes that access to <u>judicial mechanisms</u> for business-related human rights claims is often more difficult where the need is greatest as a result of both legal and practical obstacles (Ruggie, The UN "Protect, Respect and Remedy" Framework for Business and Human Rights, 2010). <u>State non-judicial mechanisms</u> include national human rights institutions or National Contact Points for signatory states of the OECD Guidelines on Multinational Enterprises. Finally, the Guiding Principles identify eight criteria that should be present in <u>non-judicial grievance mechanisms at the company level</u>: they should be legitimate, accessible, predictable, equitable, transparent, rights-compatible[96] and a source of continuous learning[97].

Mechanisms should of course be tailored to the industry, country and culture where they operate. Context adaptation becomes more relevant when dealing with Internet platforms operating globally. Unfortunately, Emily Laidlaw notes that, despite core democratic rights being engaged by the activities of Internet platforms, there is little available to users to address potential or actual infringements of such rights through corporate governance mechanisms (Laidlaw, 2015). Indeed, users face many challenges when making complaints and, generally, remain ignorant of the decision-making process of the claim. In addition, the procedures are not adjudicative because the complainant has no access to make representations to the decision makers or to hear the case being made against him (Ibid).

Much of this discussion has already been addressed in chapter two of this thesis so the practices whereby Facebook takes users' complaints will not be developed here. Instead the key problems in the existing grievance mechanisms will be pinned down. Namely that they are not accessible, predictable or transparent.

Although complaint mechanisms might seem <u>accessible</u> on online platforms, the issue is a bit tricky because it is worked algorithmically. Grievances are then classified in different types from which the user can chose one that resembles its issue of concern. There is no person you can talk to if your

---

[94] UNGP Principle number 26.
[95] UNGP Principle number 29.
[96] A rights-compatible mechanism integrates human rights norms and standards into its processes and is based on principles of non-discrimination, equity, accountability, empowerment and participation. Retrieved from: Ruggie & Nelson, *Rights-Compatible Grievance Mechanisms: A Guidance Tool for Companies and Their Stakeholders*, 2008.
[97] UNGP Principle number 31.

concern does not fit the "established box" and standard criteria. Therefore, there is very little room for users to correctly express their complaints. Consequently, users are often unable to complain for a violation of their rights because automated processes of grievances constitute a challenge impossible to overcome.

Secondly, complaint mechanisms of social media platforms clearly lack the required predictability to be efficient. Laidlaw attributes it to business being ill suited for the pseudo-judicial role they play (Ibid). Platforms owners have to weigh competing human rights and asses the lawfulness of diverse content. Due to the complexity of these tasks, their policing guidelines are often applied in an ad hoc fashion: "they are statements of policy and principle that are struggled over at times, are deployed when they are helpful and can be sidestepped when they are constraining" (Gillespie, Governance of and by platforms, 2017). Grievance mechanisms operate randomly and hence, there is no predictability whatsoever when users make complaints about their human rights being abused.

The third criterion, transparency, encompasses many of the mentioned weaknesses. The lack of access and predictability of grievance mechanisms is partly due to the opaque processes in place. Technology obscures agency and effectively obliterates duty bearers. In the context of social media, users might feel helpless and confused when claiming for their rights just because they do not know where and how to direct this claim or if it will be seriously considered or rather automatically processed by a standard algorithm. For this reason, transparency on grievance mechanisms is essential: "a grievance mechanism can only serve its purpose if the people it is intended to serve know about it, trust it and are able to use it" (UN Guiding Principles on Business and Human Rights, 2011)[98]. As shown by Ranking Digital Rights above, much work is needed in the disclosure and transparency of platform practices and processes in which grievances are dealt with.

In the presentation of his last report to the UN Human Rights Council, David Kaye emphasized these same concerns about social media standards and processes: "human rights concerns include a lack of clear and transparent rule-making and enforcement, weak and inconsistent consideration of the context, extensive reliance on automation, and a lack of any meaningful appeals process and remedies" ((UN Human Rights Council , 2018)[99].

---

[98] UNGPs Commentary of Principle 31.
[99] Find his declarations here:
https://www.ohchr.org/RU/NewsEvents/Pages/DisplayNews.aspx?NewsID=23218&LangID=E

3.3 Looking forward

The above challenges are rooted in the gray zone where human rights law ends and corporate social responsibility begins, and it is in this zone that online platforms operate (Jørgensen, Human Rights and Private Actors in the Online Domain, 2018). It is the same space where what was once considered private is ceasing to be so. Commercial interests are jumbled together with enormous responsibilities, which traditionally belonged to the public. The overlapping of private and public is the very nature of the business and human rights dilemma in this sector. It is now self-evident that privatized law enforcement and self-regulatory measures of these corporate platforms are not good enough. But is there a perfect model we should works towards?

It is argued that it is unrealistic that platforms adopt strict standards and apply them globally, just because standards on protected speech, for instance, differ depending on the location of the user. A strong answer to this claim is given by Jørgensen, who points out that platforms already adhere to different regimes for unlawful content depending on the national context in which they operate (Ibid). Therefore platforms could put in place mechanisms that ensure that no content is removed unless it abides with the criteria set out in human rights law. In fact, as stated by the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression: "National laws are inappropriate for companies that seek common norms for their geographically and culturally diverse user base. But human rights standards, if implemented transparently and consistently with meaningful user and civil society input, provide a framework for holding both States and companies accountable to users across national borders" (UN Human Rights Council , 2018).

Framing the solution

In her work, Emily Laidlaw puts forward a model that aims at building human rights safeguards into the governance structure[100]. She argues that there is no need for business to mirror the obligations that states have in the arena of human rights, nor to rely completely on voluntary company codes to ensure the respect of human rights (Laidlaw, 2015). Her point, more precisely, is that there is no need for such a split. Instead, she advances a model that is composed of an internal element, run by platforms themselves, and an external element, in the form of an independent regulatory body.

Externally, Laidlaw advocates for a strong partnership of business with governments to operationalize human rights commitments. She envisions a government with a "meta-regulatory capacity" that has

---

[100] Although her Governance Model can be extended to the protection of all human rights, it is important to note that Laidlaw has a specific focus on freedom of expression.

an oversight of the industry and even engage in self-regulation or in-house corporate governance of human rights (Ibid). Laidlaw goes further by suggesting the building of a public institution specially commissioned with overseeing issues of business and human rights in the information society[101].

The external efforts should be combined with a regulatory internal framework. However, according to Laidlaw, reform will be best achieved by enabling the environment through mainstreaming of human rights codes. The goal, she states, is to draw from the moral underpinning that links human rights and corporate social responsibility (Ibid). The external element completes the internal one by acting as an independent observer, avoiding risks of human rights being neglected by an internal regulator in favor of a cost-benefit analysis. Both elements would be linked in a complimentary and mutually reinforcing way.

Emily Laidlaw's Governance Model might seem too utopian to be feasible but in fact, the UN Guiding Principles on Business and Human Rights have already incorporated the needed complex narrative on which it can be based, a common conceptual language that frames the business and human rights debate (Ibid). In my view, by proposing a patchwork of state regulation and corporate frameworks, this model accurately tackles the roots of the present governance gap in the human rights regulatory framework: the blurring of the *private* and *public* in social media platforms. Only by adapting our mechanisms to these new social and cultural norms, we can correctly address human rights abuses associated with public life and taking place in the online private domain.

---

[101] Laidlaw herself recognizes the risks of having a body with power over business and little accountability or legitimacy. She explains that the powers, duties, and scope of this body would have to be clearly and narrowly defined by government. Retrieved from: Laidlaw, *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility,* 2015.

Conclusion

One of the most praised quotes by Facebook's founder and CEO, Mark Zuckerberg, is "move fast and break things. Unless you are breaking stuff, you are not moving fast enough"[102]. This philosophy has led his company to great success but, has Facebook broken too many barriers? Almost ten years after this statement, the platform makes headlines for interfering with elections or exposing users by sharing their personal information with third parties. Perhaps there are barriers that are not meant to be broken and perhaps there must be a speed limit on platforms' innovation. Indeed, this thesis claims that human rights provide the line in the sand that technological "progress" must not overstep.

More specifically, this thesis looks at Facebook as the new public sphere and its impacts on public life. Contrary to the normative ideal, today's public sphere is embedded in a commercial environment and public participation is shaped by the platform that hosts it. More importantly, the public sphere is found within the "business jungle", governed by private interests and self-regulatory measures. This is a matter of major concern because of the impact that platforms' services are having on fundamental rights.

The commercialization of users' personal information is at the core of platforms' business, posing a direct threat to the right to privacy. This thesis finds that Facebook's architecture and narrative steers users into believing that, when participating in public discussions, the platform acts as a "quasi-public" forum where privacy should not be a concern. In the same line, Jose van Dijck advances an interesting theory to explain how Facebook encourages the sharing of information between users while it purposely diverts attention from the sharing of this information by the platform with third parties; which actually feeds the business almost entirely.

As hosts of the public debate and as profit-driven actors, platforms have content moderation policies in place. While they aim at maintaining a "healthy community", they shape public discussions, interfering significantly with freedom of expression. This thesis agrees with Tarleton Gillespie's claim that human rights are only invoked when it serves the interest of platform owners. Generally, this corresponds to government request on content curation going against international human rights standards. Yet, platforms have self-regulatory mechanisms to curate speech to stricter extent than what is legally required. By analyzing content governance on platforms, this thesis sheds light on the

---

[102] This is a quote by Mark Zuckerberg during an interview in 2009. Retrieved from: *As Mark Zuckerberg Turns 30, His Best Quotes as CEO*. Entrepreneur, https://www.entrepreneur.com/article/233890

policies and practices by which Facebook curates public sphere and the way in which they threat freedom of expression.

The state-centered reality of human rights law is particularly challenged by technological platforms that, by taking on duties which traditionally belonged to public actors, are negatively impacting human rights. Attempting to fill the existing governance gap, Emily Laidlaw suggests a model that combines corporate self-regulation and monitoring by an independent public body. Because this proposal builds on existing frameworks and innovatively adds an outside body, it is a realistic approach with the potential of successfully adapting human rights protection mechanisms to the role that social media platforms play today.

All in all, the overwhelming role of social media such as Facebook in public life supports the idea of adoption and implementation of human rights in all its operations. Yet, the current available tools are ill equipped to persuade platform owners to adopt a true human rights based approach in all their policies and operations. Hence, there is an important governance gap that must be addressed. By framing the analysis in the overlap of the *private* and the *public,* this thesis helps understand the paradigm shift in public life taking place in social media platforms. Future research could build on this analysis and explore methods of mitigating human rights abuses online as well as look into effective human rights protection mechanisms that can adapt to the changes brought about by the information era.

Only by ensuring that human rights are protected and respected in this new public sphere can we get closer to achieving the vision of John Perry Barlow[103]. Only that way can we create a modern public sphere where internet users do not have to fear being coerced into silence and conformity.

---

[103] John Perry Barlow was referred in the introduction of this thesis for his statement in 1996: "the Internet would usher in a world where anyone, anywhere, may express his or her beliefs, no matter how singular, without fear of being coerced into silence and conformity".

Bibliography

Bonnitcha, J., & McCorquodale, R. (2017). The Concept of 'Due Diligence' in the UN Guiding Principles on Business and Human Rights. *European Journal of International Law, 28*(3), 899-919.

Borock, G. (2017). *Global Justice.* (E. N. Zalta, Ed.) Retrieved from The Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/archives/spr2017/entries/justice-global/

Boyd, D. (2018). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication* , 210-230.

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2012). Misplaced Confidences: Privacy in the Control Paradox. *Social Psychological and Personality Science*, 340-347.

Burkell, J., Fortier, A., Yeung Cheryl Wong, L., & Lynn Simpson, J. (2014). Facebook: public space, or private space? *Information, Communication & Society, 17*(8), 974-985.

Business & Human Rights Resource Centre. (2018). *General Data Protection Regulation: Issues of compliance and non-compliance.* Retrieved from Business & Human Rights Resource Centre: https://www.business-humanrights.org/en/the-general-data-protection-regulation-what-the-new-eu-privacy-law-means-for-businesses

Calhoun, C. (1992). *Habermas and the Public Sphere.* Cambridge: The MIT Press.

Calhoun, C. (1993). Civil Society and the Public Sphere. *Public Culture, V*(2), 267-80.

Castells, M. (2008). The New Public Sphere: Global Society, Communication Networks, and Global Governance. *ANNALS, AAPS*, 78-93.

Clemons, E. K. (2009). The complex problem of monetizing virtual electronic social networks. *Decision Support Systems, 48*(1), 45-56.

Crawford, K., & Gillespie, T. (2014). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society, XVIII*(3), 410-428.

Dahlberg, L. (2007). Rethinking the Fragmentation of the Cyberpublic: From Consensus to Contestation. *New Media & Society, SAGE, IX*(5), 827 - 847.

Datta, A. (2015). Automated Experiments on Ad Privacy Settings. *Proceedings on Privacy Enhancing Technologies*, 92-112.

d'Entreves, M. P. (2016). *Hannah Arendt.* (E. N. Zalta, Ed.) Retrieved from The Stanford Encyclopedia of Philosophy: \url{https://plato.stanford.edu/archives/win2016/entries/arendt/}

Dewey, J., & Rogers, M. L. (2012). *The Public and Its Problems: An Essay in Political Inquiry.* Pennsylvania: The Pennsylvania State University Press.

EUGDPR. (2018). *GDPR Key Changes*. Retrieved from EU GDPR Portal : https://www.eugdpr.org/key-changes.html

Facebook Community Standards. (2018). *Facebook*. Retrieved from
https://www.facebook.com/communitystandards/

Fuchs, C. (2014). *Social Media: A critical introduction.* London : SAGE Publications.

Garnham, N. (2000). *Emancipation, the Media, and Modernity.* Oxford: Oxford University Press.

Gavison, R. (1992, Nov). Feminism and the Public/Private Distinction. *Standford Law Review, 45*(1), 1-45.
Retrieved from http://www.jstor.org/stable/1228984

Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*, 1348-1365.

Ghosh, D., & Scott, B. (2018). *#DigitalDeceit: The Technologies Behind Precision Propaganda on the Internet.* Cambridge, MA: Harvard Kennedy School.

Gillespie, T. (2010). The Politics of Platforms. *New Media & Society, XII*, 347-364.

Gillespie, T. (2015). Platforms Intervene. *Social Media + Society*, 1-2.

Gillespie, T. (2017, May 23). *Facebook can't moderate in secret anymore*. Retrieved from Culture Digitally : http://culturedigitally.org/2017/05/facebook-cant-moderate-in-secret-any-more/

Gillespie, T. (2017). Governance of and by platforms. *SAGE Handbook of Social Media*.

Gillespie, T. (2018, June 6). *Custodians*. Retrieved from Social Media Collective: https://socialmediacollective.org/author/tarleton/?blogsub=confirming#blog_subscription-2

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media.* Yale University Press.

Gillespie, T. (2018, February 6). *Moderation is the Commodity*. Retrieved from Techdirt: https://www.techdirt.com/articles/20180206/09403839164/moderation-is-commodity.shtml

Goldman, R. (2018, April 23). *Hard Questions: What Information Do Facebook Advertisers Know About Me?* Retrieved from Newsroom Fb: https://newsroom.fb.com/news/2018/04/data-and-advertising/

Grbesa, M. (2003). Why if at all is the Public Sphere a Useful Concept? *Politicka misao, XL*(5), 110-121.

Habermas, J. (1991). *The Structural Transformation of the Public Sphere: An Inquiry into a category of Burgeois Society.* (T. B. Lawrence., Trans.) Cambridge, MA: The MIT Press.

Habermas, J. (2006). *The structural transformation of the public sphere : an inquiry into a category of bourgeois society.* (T. Burger, Trans.) London: Polity.

Habermas, J., Lennox, S., & Lennox, F. (1974). The Public Sphere: An Encyclopedia Article (1964). *New German Critique*(3), 49-55.

Heyman, R., & Pierson, J. (2015). Social Media, Delinguistification and Colonization of Lifeworld: Changing Faces of Facebook. *Social Media + Society*, 1-11.

Highfield, T. (2016). *Social Media and Everday Politics.* Cambridge: Polity Press.

Huberman, B. A. (2001). *The Laws of the Web: Patterns in the ecology of information.* . Cambridge, MA: MIT Press.

Human Rights Watch. (2018, June 06). *The EU General Data Protection Regulation*. Retrieved from Human Rights Watch: https://www.hrw.org/news/2018/06/06/eu-general-data-protection-regulation

Joseph, S. (2018, April 02). *Why the business model of social media giants like Facebook is incompatible with human rights*. Retrieved from The Conversation: https://theconversation.com/why-the-business-model-of-social-media-giants-like-facebook-is-incompatible-with-human-rights-94016

Jørgensen, R. F. (2013). Revisiting public and private. In *Framing the Net* (pp. 25-32). Northampton, MA: Edward Elgar Publishing.

Jørgensen, R. F. (2013). The internet as public sphere. In *Framing the Net* (pp. 81-106). Northampton, MA: Edward Elgar Publishing.

Jørgensen, R. F. (2017). What Platforms Mean When They Talk About Human Rights. *Policy & Internet*, 280-295.

Jørgensen, R. F. (2018). Human Rights and Private Actors in the Online Domain. In M. K. Molly K. Land, & J. D. Aronson (Eds.), *New Technologies for Human Rights Law and Practice* (pp. 243-269). Pennsylvania: Cambridge University Press.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Kelley School of Business*, 59-68.

Kellner, D. (2014). Habermas, the Public Sphere, and Democracy. In D. Boros, & J. Glass, *Re-Imagining Public Space* (pp. 19-43). New York: Palgrave Macmillan.

Khan, M. Z., Gilani, I. S., & Nawaz, A. (2012, March). From Habermas Model to New Public Sphere: A Paradigm Shift. *Global Journal of Human Social Science, XII*(5).

Kleingeld, P., & Brown, E. (2014). *Cosmopolitanism.* (E. N. Zalta, Ed.) Retrieved from The Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/archives/fall2014/entries/cosmopolitanism/

Laidlaw, E. B. (2015). *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility.* Cambridge: Cambridge University Press.

Latonero, M. (2018). Big Data Analytics and Human Rights. In *New Technologies for Human Rights Law and Practice* (pp. 149-161). Pennsylvania: Cambridge University Press.

Leetaru, K. (2017, August 1). *Is Social Media Really A Public Space?* . Retrieved from Forbes: https://www.forbes.com/sites/kalevleetaru/2017/08/01/is-social-media-really-a-public-space/#ae076762b805

Lessig, L. (1999). *Code And Other Laws of Cyberspace.* New York: Basic Books.

MacKinnon, R. (2012, June 14). *Ruling Facebookistan*. Retrieved from Foreign Policy: http://foreignpolicy.com/2012/06/14/ruling-facebookistan/

MacKinnon, R., Hickok, E., Bar, A., & Lim, H.-i. (2014). *Fostering Freedom Online: The Role of Internet Intermediaries.* Division for Freedom of Expression and Media Development. Paris: United Nations Educational, Scientific, and Cultural Organization (UNESCO).

Maclay, C. M. (2010). Protecting privacy and expression online: Can the Global Network Initiative embrace the character of the net. *Access controlled: The shaping of power, rights, and rule in cyberspace*, 87-108.

McCarthy, T. (1991). Introduction. In J. Habermas, *The structural transformation of the public sphere* (pp. xi-xiv). Cambridge: The MIT Press.

Mueller, M. (2015). Hyper-transparency and social control: Social media as magnets for regulation. *Telecommunications Policy, 39*(9), 804-810.

OECD. (2018). *OECD Guidelines for Multinational Enterprises*. Retrieved from Responsible Business Conduct: https://mneguidelines.oecd.org/guidelines/

Office of the High Commissioner for Human Rights. (2014). *The UN Guiding Principles on Business and Human Rights: Relationship to UN Global Compact Commitments.* United Nations Global Compact.

Papacharissi, Z. (2002). The virtual sphere: the internet as a public sphere. *New Media Society, IV*(1), 9-27.

Papacharissi, Z. (2010). *A Private Sphere: Democracy in a Digital Age.* Cambridge: Polity Press.

Papacharissi, Z. (2010). The Virtual Sphere 2.0: The Internet, the Public Sphere and beyond. In A. Chadwick, & P. Howard, *The Routledge Handbook of Internet Politics* (pp. 230-245). New York: Routledge.

Papacharissi, Z., & Fernback, J. (2005). Online Privacy and Consumer Protection: An analysis of Portal Privacy Statements. *Journal of Broadcasting & Electronic Media*, 259-281.

Porter, E. (2018, April 24). The Facebook Fallacy: Privacy is Up to You. . *The New York Times*. Retrieved from https://www.nytimes.com/2018/04/24/business/economy/facebook-privacy.html

Powels, J. (2018, May 25). *The New Yorker.* Retrieved from The G.D.P.R., Europe's New Privacy Law, and the Future of the Global Data Economy: https://www.newyorker.com/tech/elements/the-gdpr-europes-new-privacy-law-and-the-future-of-the-global-data-economy

Ranking Digital Rights. (2018). *2018 Corporate Accountability Index.*

Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business entrepreses, John Ruggie. (2011). *Guiding Principles on Business adn Human Rights: Implementing the United Nations "Protect, Respect, and Remedy" Framework.* Human Rights Council.

Reuters. (2018, March 12). *U.N. investigators cite Facebook role in Myanmar crisis*. Retrieved from Reuters: https://www.reuters.com/article/us-myanmar-rohingya-facebook/u-n-investigators-cite-facebook-role-in-myanmar-crisis-idUSKCN1GO2PN

Ruggie, J. (2010). *The UN "Protect, Respect and Remedy" Framework for Business and Human Rights.* United Nations.

Ruggie, J. (2011, April 9). *UN Guiding Principles for Business & Human Rights*. Retrieved from Harvard Law School Forum on Corporate Governance and Financial Regulation: https://corpgov.law.harvard.edu/2011/04/09/un-guiding-principles-for-business-human-rights/

Ruggie, J., & Nelson, J. (2008). *Rights-Compatible Grievance Mechanisms: A Guidance Tool for Companies and Their Stakeholders.* John F. Kennedy School of Government, Harvard University.

Schmidt, E., & Cohen, J. (2013). *The New Digital Age: Reshaping the future of people, nations and business.* New York: ALFRED A. KNOPF.

Shirky, C. (2008). Publish, then filter . In *Here comes everybody: the power of organizing without organizations* (pp. 81-108). London: Penguin Books.

Shirky, C. (2011). The Political Power of Social Media: Technology, the Public Sphere, and Political Change. *Foreign Affairs, 90*(1), 28-41.

The Atlantic . (2017, December 20). *Could Facebook Be Tried for Human-Rights Abuses?* Retrieved from The Atlantic : https://www.theatlantic.com/technology/archive/2017/12/could-facebook-be-tried-for-war-crimes/548639/

The Guardian. (2018, April 3). *Revealed: Facebook hate speech exploded in Myanmar during Rohingya crisis*. Retrieved from The Guardian: https://www.theguardian.com/world/2018/apr/03/revealed-facebook-hate-speech-exploded-in-myanmar-during-rohingya-crisis

UN General Assembly, *International Covenant on Civil and Political Rights*, 16 December 1966, United Nations, Treaty Series, vol. 999, p. 171, available at: http://www.refworld.org/docid/3ae6b3aa0.html

UN General Assembly, *Universal Declaration of Human Rights*, 10 December 1948, 217 A (III), available at: http://www.refworld.org/docid/3ae6b3712c.html

UN Human Rights Council . (2018). *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression.*

UN Human Rights Council. (2015). *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye.*

UN Working Group on Business and Human Rights. (2018). *Corporate human rights due diligence – identifying and leveraging emerging.* Note prepared by the UN Working Group on Business and Human Rights for the consultation process.

United Nations. *Guiding Principles on Business and Human Rights: Implementing the United Nations "protect, Respect and Remedy" Framework.* 2011.

van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media.* New York: Oxford University Press.

van Dijck, J. (2016). Connective Memory: How Facebook takes Charge of your Past. In L. Bond, S. Craps, & P. Vemeulen, *Memory Unbound: Tracing the Dynamics of Memory Studies* (pp. 151-172). New York: Berghahn Books.

van Dijck, J., & Powel, T. (2015). Social Media and the Transformation of the Public Space. *Social Media + Society*, 1-5.

Weintraub, J. (1997). The Theory and Politics of the Public/Private Distinction . In J. Weintraub, & K. Kumar, *Public and private in thought and practice* (pp. 1-42). London: University of Chicago Press.

Wessler, H., & Freudenthaler, R. (2011, Feb 23). *Oxford Bibliographies Datasets.* Retrieved from Public Sphere: http://dx.doi.org/10.1093/obo/9780199756841-0030

Wolfe, A. (1997). Private and public in theory and practice: Some Implications of an Uncertain Boundary. In J. Weintraub, & K. Kumar, *Public and Private in Thought and Practice* (pp. 182-203). London: University of Chicago Press.

Wright, D. C. (2012). Theorizing History: Separate Spheres, the Public/Private Binary and a New Analytic for Family Law. *University of Florida, Law Scholarship Repository* .

Zimmer, M. (2018). Addresing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity. *Social Media + Society*, 1-11.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75-89.