CELE

Human Rights Impact Assessments: Trends, Challenges, And Opportunities for ICT Sector Adoption

October 2020

Facultad de Derecho
Centro de Estudios en Libertad
de Expresión y Acceso a la Información

UP
Universidad
de Palermo

Human Rights Impact Assessments: Trends, Challenges, And Opportunities for ICT Sector Adoption

Ramiro Álvarez Ugarte
CELE

In our previous work, we sought to place human rights impacts assessments (HRIA) within their history. We showed how the institute comes from a set of corporate practices that rest uncomfortably within different frameworks. On the one hand, HRIAs stem from the environmental impact assessments developed in the 1960s and 1970s as the legal and procedural mechanism to ensure that private companies respect minimum environmental standards defined by law and enforced by administrative agencies. On the other hand, HRIAs also operate within the obscure and ambiguous world of corporate social responsibility (CSR). Finally, we also showed how HRIAs are a preferred tool of a strange animal: the protect, respect and remedy (PRR) framework adopted by the United Nations in 2011, advanced by the UN Special Representative John Ruggie, who was actively supported by former Secretary-General, Kofi Annan.[1] Thus, our conceptual analysis placed HRIAs in a specific historical context: the one produced by the UN decision to step back from mandatory international treaties addressing the relationship between corporations and human rights and to adopt Ruggie's more relaxed framework. That initial analysis ended with the introduction of different paths through which HRIAs could evolve. From our perspective, they will either be swallowed by the self-regulation framework of CSR or will become an increasingly *juridified* practice, pushed by either international or national laws. The latter path, however, could look very different depending on how various regulatory possibilities are combined.[2]

---

[1] J. Ruggie, «Guiding Principles on Business and Human Rights. Implementing the United Nations "Protect, Respect and Remedy" Framework». Human Rights Council. Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, New York. HR/PUB/11/04. 2011.

[2] On this, *see* K. Abbott; D. SnidalWalter Mattli, Ngaire Woods (eds.) («The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State», in *The politics of global regulation*, Princeton University Press, Princeton, 2009.) and C. Marsden; T. Meyer; I. Brown («Platform Values And Democratic Elections: How Can The Law Regulate Digital Disinformation?», Computer Law & Security Review, vol. 36, 2020, Retrieved from http://www.sciencedirect.com/science/article/pii/S026736491930384X)

This paper proceeds in the following way. First, we briefly review Ruggie's PRR framework in which HRIAs seem—at this time—to be standing. We move then to frame the main goal of this paper: to outline the potential challenges and perils of deploying this tool in the ICT sector. We do this with some degree of hindsight, for several ICT companies have already moved in this direction and important multistakeholder initiatives such as the Global Network Initiative (GNI) have embraced it. But some caution is required, for it is still too early to tell where HRIAs in the ICT sector will move in the next few years. To an extent, their future path depends on broader regulatory challenges. The third section overviews how HRIAs were used by Facebook, one of the Internet companies that more heavily relied on the practice to study its impacts in different contexts. By closely looking at these HRIAs or their executive summaries we obtain some insights in terms of how the *assessment* is produced on the field. The fourth section outlines a few preliminary conclusions regarding the limitations, weaknesses, challenges, and potential of the use of HRIAs in the ICT sector. In particular, we focus on a certain lack of transparency that seems to affect the HRIAs we studied, and a lack of consensus on the scope of liability ICT companies should be subjected to, a problem that derives—in part—from a difficulty in establishing causal mechanisms between certain technological developments and specific human rights wrongs. The fifth and final section discusses these findings and advances a few hypotheses of the challenges lying ahead.

A. The Corporate Responsibility to Respect Human Rights

Unanimously endorsed by the Human Rights Council in 2011, the United Nations Guiding Principles on Business and Human Rights (UNGPs) provide the first "globally recognized and authoritative framework for the respective duties" of states and businesses in preventing and addressing adverse human rights impacts resulting from business activities across sectors and around the world.[3] Grounded in the foundational principles that states have a "duty" to "protect" human rights, while businesses have a "responsibility" to "respect" human rights, this framework neither creates new international law obligations nor limits the existing "legal obligations a state may have undertaken or be subject to under

[3] J. Ruggie, "Guiding Principles on Business and Human Rights. Implementing the United Nations «Protect, Respect and Remedy» Framework", cit.

international law."[4] Rather, the UNGPs "should be read, individually and collectively, in terms of their objective of enhancing standards and practices with regard to business and human rights so as to achieve tangible results for affected individuals and communities."[5]

The state's duty to protect human rights is derived from states' international human rights law obligations.[6] According to the first principle of the UNGPs, this duty requires that states protect against human rights abuses by third parties, including businesses, within their territory and that they take "appropriate steps to prevent, investigate, punish and redress such abuse through effective policies, legislation, regulations and adjudication."[7]

In contrast, while the corporate responsibility to respect human rights "refers to internationally recognized human rights," it "is distinct from issues of legal liability and enforcement."[8] The corporate responsibility to respect human rights requires businesses to "[a]void *causing* or *contributing* to adverse human rights impacts through their own activities" and "address such impacts when they occur" (emphasis added).[9] When businesses have not contributed to adverse human rights impacts but are nonetheless directly *linked* to those impacts through their operations, the responsibility to respect human rights requires that they "seek to prevent or mitigate," those impacts.[10] According to the UNGPs, the corporate responsibility to respect human rights "exist[s] independently of states' abilities and/or willingness to fulfill their own human rights obligations."[11] This responsibility applies to all businesses, "regardless of their size, sector, operational context, ownership and structure."[12] However, Principle 14 stipulates that the "scale and complexit[y] of the means" used to meet this responsibility "may vary according to these factors" and

---

4 *Ibid.*, paras. 1-13.

5 *Ibid.*, para. 1.

6 *Ibid.*, p. 1.

7 J. Ruggie, "Guiding Principles on Business and Human Rights. Implementing the United Nations «Protect, Respect and Remedy» Framework," cit., principle 1.

8 *Ibid.*, p. 13.

9 J. Ruggie ("Guiding Principles on Business and Human Rights. Implementing the United Nations «Protect, Respect and Remedy» Framework," cit.), 14. Principle 19 elaborates on how businesses should address impacts and instructs that "appropriate action" will vary according to two factors: (1) the determination of whether the businesses "causes or contributes to an adverse impact,"or rather is "directly linked," to the impact by "its operations, products or services by a business relationship;" and (2) the business's "leverage in addressing the adverse impact."

10 *Ibid.*, p. 14.

11 *Ibid.*

12 *Ibid.*, p. 15.

the "severity of the enterprise's adverse human rights impacts."[13]

The UNGPs counsel that businesses should have policies and processes in place to "know and show that they respect human rights," including a "human rights due diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights."[14] While the UNGPs grant businesses discretion over how to implement human rights due diligence in the context of their business operations,[15] Principles 18 through 21 elaborate on its four key components.[16] According to these Principles, human rights due diligence involves identifying and assessing "any actual or potential human rights impacts with which they may be involved,"[17] integrating findings from impact assessments and taking appropriate action in response to the impacts identified,[18] verifying whether adverse human rights impacts identified are being addressed by tracking the effectiveness of the actions taken in response,[19] and communicating this work with appropriate stakeholders.[20] Within this framework, human rights impacts assessments (HRIAs) have emerged as the UN's most highly-recommended tool for businesses to begin the ongoing human rights due diligence process, and in turn "know and show that they respect human rights."[21]

B. Human Rights Impact Assessments in the ICT Sector

In the first paper in this series, we examined the origins of HRIAs and argued that this history can help us better understand the strengths and weaknesses of HRIAs

[13] *Ibid.*, p. 14.

[14] *Ibid.*, pp. 15-16. This principle instructs that the policies and processes businesses «should have in place» include: «(a) A policy commitment to meet their responsibility to respect human rights; (b) A human rights due diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights; and (c) Processes to enable the remediation of any adverse human rights impacts they cause or to which they contribute».

[15] *Id.* at 18, Principle 17(b). According to this principle, human rights due diligence "[w]ill vary in complexity with the size of the business enterprise, the risk of severe human rights impacts, and the nature and context of its operations." The Commentary further elaborates, "[h]uman rights due diligence can be included within broader enterprise risk-management systems, provided that it goes beyond simply identifying and managing material risks to the company itself, to include risks to rights-holders."

[16] *Id.* at 16, Commentary on Principle 17.

[17] *Id.* at 19, Principle 18.

[18] *Id.* at 19 – 20, Principle 19.

[19] *Id.* at 22 – 23, Principle 20.

[20] *Id.* at 23, Principle 21. This principle instructs that the communication expected of businesses should vary according to the severity of human rights impacts of said business. At least, all businesses "should be prepared to communicate…how they address their human rights impacts," externally, and businesses whose operations pose severe human rights risks should formalize reporting on how they are addressing those risks.

[21] R. Álvarez Ugarte; L. Krauer, «El sector TICs y derechos humanos: hacia un marco conceptual de las evaluaciones de impacto en DDHH». Centro de Estudios para la Libertad de Expresión, Buenos Aires, Argentina. October 2020.

in practice.[22] This paper further examines the effectiveness of HRIAs and specifically considers what trends in HRIA adoption and implementation in the ICT sector[23] can offer to our understanding of the tool's ability to yield "tangible results for affected individuals and communities," as envisioned by the UNGPs.[24] As Jørgensen, Veiber and ten Oever recall, several ICT companies "have worked with human rights topics for years, as part of corporate social responsibility (CSR)."[25]

This paper was informed by a desk review of how five different ICT companies—Facebook, Google, Telia, Telefonica, and Microsoft—have adopted and implemented HRIAs. More specifically, we analyzed the ways the various HRIAs evaluating these companies meet, and fail to meet, the human rights due diligence expectations elaborated in Principles 18 through 21 of the UNGPs. From the onset, we note that research in this area is necessarily limited to what companies themselves chose to make accessible to the general public.[26] Accordingly, to supplement our HRIA review, we also examined the language these companies use in relation to human rights more broadly as well as the information they disclose to "know and show that they respect human rights."[27]

We begin by presenting a brief overview of secondary literature on HRIAs in the ICT sector. We continue by describing the release of four HRIAs evaluating Facebook's presence in Myanmar, Cambodia, Indonesia, and Sri Lanka. This case study aptly captures the three trends that surfaced in our review that will guide the remainder of our discussion. Our analysis of these trends begins in section IV.1, where we discuss the observation that ICT companies are more likely to be transparent about human rights due diligence when the human rights risks related to their operations have a clear bad actor, such as a state. In section IV.2,

---

[22] *Ibid.*

[23] The ICT sector, or information and communications technology sector, is particularly relevant to CELE because of its dominance over information flow in modern society. The ICT sector includes telecommunications infrastructure and service providers, Internet infrastructure and service providers, and Internet platforms, like social media companies and other user-generated content platforms.

[24] J. Ruggie, "Guiding Principles on Business and Human Rights. Implementing the United Nations «Protect, Respect and Remedy» Framework", cit., p. 1.

[25] R. F. Jørgensen; C. B. Veiberg; N. ten Oever, «Exploring the role of HRIA in the information and communication technologies (ICT) sector», in *Handbook on human rights impact assessment*, Edward Elgar Publishing, 2019, p. 4.

[26] *See* J. Ruggie, «Human rights impact assessments --- resolving key methodological questions». Human Rights Council. Report of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises, Geneve, Switzerland. A/HRC/4/74. 2007 (noting HRIAs are notoriously difficult to come by. We address this issue at greater length in our discussion of transparency in Part IV.).

[27] J. Ruggie, "Guiding Principles on Business and Human Rights. Implementing the United Nations «Protect, Respect and Remedy» Framework," cit., p. 1.

we discuss the lack of understanding and consensus around the extent to which ICT companies should be held responsible for the promotion and protection of freedom of opinion and expression impacts in the international community. We argue that the absence of resolution on this issue presents a significant obstacle to the ability of HRIAs to yield "tangible results". We focus this discussion on the difficulty of determining whether a company caused, contributed, or is directly linked to adverse human rights impacts related to their operations. Finally, in section IV.3 we discuss how, despite a general lack of transparency and consistency around process, methodology, and stakeholder involvement, HRIAs may still meaningfully contribute to the discourse around the respective duties of states and businesses in preventing and addressing adverse human rights impacts resulting from business activities across sectors and around the world.

We conclude that these trends are entirely consistent with the voluntary nature of the corporate-responsibility-to-respect framework. With or without HRIAs, in the absence of legally enforceable duties, businesses lack the incentives necessary to thoroughly and consistently operate in a manner that does not adversely impact human rights, at least at this moment. As demands for corporate accountability—from civil society, investors, and governments alike—continue to increase, we argue that businesses may soon find increased incentives to demonstrate that the voluntary corporate-responsibility-to-respect framework can work.

C. The ICT Sector and Special Challenges

As we argued in the first paper of this series[28] and briefly repeated before, the place of transnational corporations *vis à vis* international human rights law is a contested one. Here we would like to highlight some special difficulties in that relationship that are particular to the ICT sector, as well as a sketch of relevant milestones in this history.

For instance, some companies of the ICT sector have a really weak physical footprint, which allows for different decisions that directly affect how much they *impact* a given country. As Maclay recalls, when Yahoo had to decide on an investment in Vietnam they produced an HRIA that led them into making choices that shield them from uncomfortable requests of an authoritarian state, such as storing in-

[28] R. Álvarez Ugarte; L. Krauer, "El sector TICs y derechos humanos: hacia un marco conceptual de las evaluaciones de impacto en DDHH," cit.

formation in a server in Singapore rather than locally or hiring a handful of sales employees with no operational responsibilities.[29] This was possible because some companies—especially those that can provide cloud-based services—can enter a country without deploying anything: cables, servers, employees, and so on. Most ICT companies that fall within this category do not need—and usually, do not have—neither local corporate personhood nor representation, unless markets become particularly important. This feature of some of the most influential companies in the ICT sector is related to another one: as they do not have a local presence, the *governance gap* that affects TNCs and the respect for human rights is bigger in these companies. Not only is it difficult for local governments to implement and create their own rules regarding these companies; but it is also often impossible.

The Yahoo story previously introduced was called—according to Maclay—by the Global Network Initiative (GNI), a multistakeholder initiative aimed at ensuring that companies comply with some standards. As Maclay puts it, participants "evaluate human rights risks and seek opportunities to mitigate them when considering whether and how to enter a new market."[30] In the story ahead, the GNI will have a prominent place, for the companies that generally conduct HRIAs usually do it as part of their engagement with GNI. Some of the limits we will discuss are necessarily connected to the structure and reach of this initiative.

D. Case Study: Facebook and HRIAs

On November 5, 2018, Facebook released an HRIA of the company's presence in Myanmar. The 62-page assessment, commissioned by Facebook and conducted by Business for Social Responsibility (BSR), identified "security; privacy; freedom of expression, assembly, and association; children's rights; nondiscrimination; standard of living; and access to culture" as priority human rights impacts related to Facebook's presence in Myanmar.[31] Of these impacts, BSR determined that Facebook contributes to three.[32] "By developing innovative products and services," the assessment states, "Facebook will have contributed to" positive im-

[29] C. M. Maclay, *Can the Global Network Initiative Embrace the Character of the Net?*, MIT Press, Cambridge, Mass, 2010, p. 87.

[30] *Ibid.*

[31] BSR, «Human Rights Impact Assessment: Facebook in Myanmar 1». Business for Social Responsibility, New York. November 5, 2018.

[32] *Ibid.*, p. 40-41.

pacts on the standard of living and access to culture in Myanmar.[33] BSR found that Facebook contributes to freedom of expression violations "when content is removed in error, but is only linked to violations resulting from government actions or demands for content removal." In contrast, BSR determined that Facebook neither causes nor contributes to privacy, nondiscrimination, or child rights violations in the country, "but is linked to them via its products and services."[34] Similarly, the assessment concluded that "Facebook itself does not cause or contribute to" either security or freedom of expression human rights violations, but rather is directly linked to these violations via "the actions of users that violate Facebook's Community Standards," and "government actions or demands for content removal," respectively.[35] In a blog post accompanying the release of the assessment, Alex Warofka, a Product Policy Manager at Facebook, conceded that the company had not "done enough to help prevent [the] platform from being used to foment division and incite offline violence."[36]

Less than two years later, on May 12, 2020, Facebook released the findings of three HRIAs evaluating the company's presence in Cambodia, Indonesia, and Sri Lanka.[37] BSR conducted Facebook's HRIA in Cambodia, while a new partner, Article One, conducted Facebook's HRIAs in Indonesia and Sri Lanka. Unlike 2018 when Facebook released BSR's 62-page Myanmar HRIA to the public in full, Facebook only published eight-page executive summaries of the three new HRIAs.[38] BSR's assessment of Facebook's human rights impact in Cambodia surfaced the same impacts identified in its 2018 Myanmar HRIA.[39] However, unlike the 2018 Myanmar assessment, the eight-page executive summary did not present determinations about whether Facebook caused, contributed to, or was directly linked to the impacts identified in the assessment.[40]

Article One's HRIAs both begin by acknowledging that Facebook "has bro-

---

[33] *Ibid.*

[34] BSR, "Human Rights Impact Assessment: Facebook in Myanmar 1," cit.

[35] *Ibid.*, pp. 35-36.

[36] A. Warofka, *An independent assessment of the human rights impact of Facebook in Myanmar*, About Facebook, 05/11/2018, retrieved from https://about.fb.com/news/2018/11/myanmar-hria/.

[37] M. Sissons; A. Warofka, *An update on Facebook's human rights work in Asia and around the world*, About Facebook, 05/11/2018, retrieved from https://about.fb.com/news/2020/05/human-rights-work-in-asia/.

[38] *Ibid.*

[39] BSR, «Human Rights impact assessment: Facebook in Cambodia. Executive Summary». Business for Social Responsibility, New York. 2019.

[40] *Ibid.* Under the heading «About This Executive Summary» on the first page of the report, BSR notes that its full assessment consists of "a longer report with additional explanatory detail, supporting evidence, and context."

ken down barriers across the world," "made neighbors of strangers," and "as its mission states, '*brought the world closer together.*'"[41] "At the same time," the assessments continue, Facebook's platforms have "been misused by bad actors to infringe on human rights," and Indonesia and Sri Lanka represent two "of the most critical countries when it comes to potential human rights infringements on Facebook's platforms."[42] "Recogniz[ing] the need to conduct country-level human rights impacts assessments," that prioritize "at-risk countries," the assessment explains, Facebook partnered with Article One "[t]o determine the degree to which Facebook's platform may or may not have contributed to adverse human rights impacts and to mitigate the risk of further impacts," in Indonesia and Sri Lanka.[43]

Article One's Indonesia HRIA identifies two primary risk categories with the potential to affect Indonesian society as a whole. Article One found that "[t]he proliferation of disinformation and misinformation – or 'hoaxes,' as they are referred to in Indonesia" on Facebook's platforms "have the ability to adversely impact human rights."[44] The assessment reported that hoaxes "have been used to target political figures and promote politically motivated disinformation in the hopes of impacting the outcomes of national and local elections," and in the case of kidnapping hoaxes "have stoked fear in communities."[45] The assessment also concluded that Facebook poses salient human rights risks to free expression and privacy. According to the report, "overly broad laws regarding blasphemy and defamation…have been used against individuals based on their social media usage."[46] The assessment reported that Indonesia's blasphemy law has "been used to chill free expression," and that blasphemy convictions in the country have "resulted directly from the posting of material on Facebook."[47] Concerning the right to privacy, Article One identified human rights risks relating to

[41] Article One, «Human Rights impact assessment: Facebook in Indonesia. Executive Summary». Article One, New York. 2018, «Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary». Article One, New York. 2018.

[42] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[43] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[44] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit., p. 4.

[45] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[46] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[47] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

"illegitimate government requests for user data; Facebook's involvement in the Cambridge Analytica scandal; and the practice of doxing which has been used against individuals critical of Islam."[48]

Taken together, Article One's Indonesia HRIA concluded that "the actions of Facebook users may infringe upon," the general Indonesian population's right to: "non-discrimination," "safety and security of person," "privacy," "freedom of religion," "free expression and to seek and impart information," and "take part in government."[49] Moreover, Article One found that Facebook's presence in Indonesia posed additional risks to vulnerable groups including women, LGBTQ+ individuals and children. In addition to the human rights impacts these groups may face as members of the general population, the report concluded that "[t]he actions of Facebook users may infringe on," their right to: "be treated with dignity," "health," "be free from slavery," "work…and to an adequate standard of living," "free assembly," "education…if the cyberbullying [of a child] affects school attendance and/or performance," "be free from attacks on a child's reputation," and "be protected from sexual exploitation and abuse."[50]

In Sri Lanka, Article One found that Facebook's most severe impacts were borne by "vulnerable groups, including religious minorities, women, LGBTQ communities, children and human rights defenders."[51] The assessment found that "the actions of Facebook users may infringe upon," the rights to: "be treated with dignity," "non-discrimination," "safety and security of person," "be free from slavery," "privacy," "free expression and to seek and impart information," "freedom of religion," "freedom of expression," and to "take part in government."[52]

Significantly, Article One's Sri Lanka assessment concluded: "that the Facebook platform *contributed to* spreading rumors and hate speech, which may have led to 'offline violence.'"[53] The assessment also suggested that "prior to the 2018 Internet disruption,[54] Facebook failed to effectively implement its Community Standards in

[48] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[49] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[50] Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit., "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[51] Article One, "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[52] *Ibid.*

[53] *Ibid.*

[54] *Ibid.*

Sri Lanka by failing to take down content that violated these Standards in a timely manner."[55] Article One reported that "the majority of civil society organizations engaged by Article One stated that they had tried to engage Facebook regarding the misuse of its platform dating back to 2009."[56] Despite direct calls to the company and public reports linking Facebook to offline violence, "Article One repeatedly heard allegations that Facebook did not respond to these calls promptly, and that the company was largely unresponsive until the government shut down of social media in March 2018."[57] Further, in both assessments, Article One found that Facebook's response to impacts "was slow and, at times, insufficient, which potentially exacerbated the impacts" identified.[58] Referring to the UNGPs and Facebook's responsibility to exercise due diligence to mitigate adverse human rights impacts related to its operations, Article One's Sri Lanka assessment reasoned that "Facebook's lack of formal human rights due diligence in Sri Lanka prior to this HRIA and the limited cultural and language expertise among Facebook staff at the time of the May 2018 Kandy incident may have contributed to offline harm stemming from online engagement."[59] Both assessments conclude that human rights impacts may have been "exacerbated by now phased out algorithms designed to drive engagement on the platform, regardless of the veracity or intention of the content."[60]

E. Analysis

We recount Facebook's adoption and implementation of HRIAs to frame our discussion of the three trends that surfaced in our review of HRIAs in the ICT sector. To summarize, our conclusions are the following:

1. ICT companies are more likely to be transparent about their human rights due diligence when the human rights risks related to their operations have a clear bad actor, such as a state.
2. The lack of consensus regarding the extent to which the ICT sector should be held responsible for the promotion and protection of freedom of opinion and expression significantly hinders the effectiveness of HRIAs.

[55] *Ibid.*

[56] *Ibid.*

[57] *Ibid.*

[58] *Ibid.*

[59] *Ibid.*

[60] *Ibid.*

3. There is a general lack of transparency or consistency regarding processes, methodology, and stakeholder involvement in HRIAs.

We propose that these trends reflect the triumph of the voluntary, soft law model over the hard law model discussed in our previous paper. We discuss these three trends, their origins in the voluntary soft law framework, and their implications on HRIAs in the ICT sector moving forward in the remainder of the paper.

1. Transparency and Bad Actors

Our analysis begins with a discussion of our observation that ICT companies may be more transparent about their human rights due diligence when the human rights risks related to their operations have a clear bad actor, such as a State. Consider, for instance, BSR HRIA evaluating Facebook's presence in Myanmar. BSR concluded that "Facebook itself does not *cause* or *contribute* to" either security or freedom of expression human rights violations, but rather is directly *linked* to these violations via "the actions of users that violate Facebook's Community Standards," and "government actions or demands for content removal," respectively. Similarly, when presenting impacts identified in its HRIAs evaluating Facebook's presence in Indonesia and Cambodia, Article One concluded that "the actions of Facebook users," rather than the actions of Facebook itself, "may infringe upon" those rights.

These language choices are likely deliberate and follow closely Ruggie's voluntary framework. In that sense, the language of connection implied by the idea of *linkage* signals towards a problem in which some other actor is responsible, in this case, users and the state. This points to a rather important weakness in the use of the tool in the ICT sector because usage of these tools seems to be a fundamental consideration to study and bear in mind when evaluating how these tools *impact* human rights. This weakness comes out, for instance, in the way BSR Vice President Dunstan Allison-Hope discusses some of the challenges of the tool for companies in the technology industry, where "the interplay between the design of the product by the technology company and how it is used in real life, whether by individuals, enterprise customers, or governments" is very important.[61] He writes

[61] D. Allison-Hope, *Human rights assessments in the decisive decade: Applying the UNGPs in the technology sector*, BSR, 11/02/2020, retrieved from https://www.bsr.org/en/our-insights/blog-view/human-rights-assessments-in-the-decisive-decade-ungp-challenges-technology.

that BSR "often finish[es] a human rights assessment for a technology company and wryly conclude[s] that we should be undertaking the assessment for the [consumers] using the technology, not just the company designing it."[62]

This perspective is evident in the language BSR used in its evaluation of the human rights impacts of Facebook's presence in Myanmar. For example, the report found that "[*a*] *minority of users* is seeking to use Facebook as a platform to undermine democracy and incite offline violence."[63] Likewise, the assessment found that "[t]he prevalence of hate speech, disinformation, and *bad actors on Facebook* has had a negative impact on freedom of expression, assembly, and association on Myanmar's most vulnerable users."[64] Attributing the human rights impacts to a minority of users and bad actors external to Facebook allows BSR to conclude that these external actors, and not Facebook, are responsible for the human rights impacts related to the company's business operations in Myanmar. This is significant because the determination of the company's relationship to human rights impacts informs the "appropriate action" the company should take to address those impacts,[65] and while identifying another party responsible for a known impact should not necessarily result in the determination that the company is only "directly linked" to the impact,[66] that appears to be the case in the Myanmar assessment.[67]

The reluctance to identify the ICT company being assessed as a bad actor is also evident in relation to a challenge Allison-Hope labels the "substitutability problem", which stands as the idea that if companies do not develop a product and sell it to a "nefarious actor" some other competitor will.[68] Google's only publicly available HRIA, an assessment conducted by BSR, illustrates this challenge.[69]

---

62 *Ibid.*

63 BSR, "Human Rights Impact Assessment: Facebook in Myanmar 1," cit.

64 *Ibid.* (emphasis added).

65 J. Ruggie, "Guiding Principles on Business and Human Rights. Implementing the United Nations «Protect, Respect and Remedy» Framework," cit., pp. 20-21 (Principle 19(b) (Stating, «Appropriate action will vary according to: (i) Whether the business enterprise causes or contributes to an adverse impact, or whether it is involved solely because the impact is directly linked to its operations, products or services by a business relationship».).

66 OHCHR, «OHCHR response to request from BankTrack for advice regarding the application of the UN Guiding Principles on Business and Human Rights in the context of the banking sector.» UN High Commissioner for Human Rights, Geneva. June 12, 2017.

67 This topic is discussed at length in the following section.

68 D. Allison-Hope, "Human rights assessments in the decisive decade: Applying the UNGPs in the technology sector," cit. («"If we don't sell this product to a nefarious actor, then someone else will," is a common refrain in the technology industry—and while not unique to the technology industry, the problem does take on special significance»).

69 BSR, «Google celebrity recognition API human rights assessment: Executive summary.» Business for Social Responsibility, New York. October 2019.

While it is not entirely clear whether any human rights impacts, actual or potential, were identified by the assessment,[70] BSR notes that "[u]nderpinning all these issues is the essential point that *M&E companies* have a crucial role to play in addressing potential adverse human rights impacts—while there are actions Google can take to avoid, prevent, or mitigate impacts, *much depends on the customers* also using the API responsibly."[71] Further, in a section titled "System Wide Change" BSR writes "Google is not the only provider of celebrity recognition products—if Google chooses to reduce human rights risk by not selling the API to certain customers, other providers without the same controls may step in, and the same human rights violations may occur anyway".[72] As Allison-Hope explains in the blog post, "the overall realization of human rights is not improved if a responsible company decides not to provide service to a nefarious actor and a different company chooses to do so anyway."[73]

We observed further evidence that companies may be more transparent about their human rights due diligence where the human rights risks related to their operations have a clear bad actor in the context of the Global Network Initiative (GNI). GNI was launched in 2008 to address how technology companies can best respect the freedom of expression and privacy rights of users in the face of daily requests from governments asking these companies to censor content, restrict access to communications, and provide access to user data.[74] GNI members agree to follow the GNI Principles and follow an Implementation Guideline. As part of their membership, companies are assessed every two years.[75] Even though these assessments are not public, a summary report is published by GNI.[76] In contrast to HRIAs where ICT companies under evaluation are frequently the subject of scrutiny, GNI membership requires the ICT companies to report on instances of state efforts to interfere with human rights. Within this system, ICT companies

[70] *Ibid.*, p. 2 (Under the heading «Potential Human Rights Issues and Impacts for Celebrity Recognition Product» the assessment states «BSR and Google Cloud AI identified a range of potential issues, challenges, and dilemmas relating to the field of celebrity recognition products, which BSR further analyzed for potential human rights impacts. These were the most significant: There is no consensus on the definition of "celebrity."…Meaningful consent needs to be addressed…Celebrities can be vulnerable…Human rights risks will vary by content». However, it is important to note that the four-page executive summary may not adequately reflect the findings of the full assessment conducts.).

[71] *Ibid.*, p. 3 (emphasis added).

[72] BSR, "Google celebrity recognition API human rights assessment: Executive summary," cit.

[73] D. Allison-Hope, "Human rights assessments in the decisive decade: Applying the UNGPs in the technology sector," cit.

[74] GNI, «The GNI Principles at Work.» Global Network Initiative, Washington D.C. 2019.

[75] *Ibid.*

[76] R. F. Jørgensen; C. B. Veiberg; N. ten Oever, "Exploring the role of HRIA in the information and communication technologies (ICT) sector," cit., p. 5.

are not at risk of being identified as bad actors—rather, it is more likely that their compliance with the GNI Principles can help them document their compliance with national law and procedure. In doing so, they can "know and show" that they respect human rights without having to take responsibility for human rights violations related to their operations. The state-centered focus of GNI seems to be less risky for companies than independent HRIAs. As Jorgensen et al. point out "the guidelines developed by the GNI are oriented towards government interference, whereas corporate practices outside the government-company binary are left unaddressed (Jørgensen, 2018). As such, the risk assessments conducted pursuant to the GNI guidelines tend not to scrutinize the full range of business practices and procedures for their potential human rights impacts…"[77]

2. The Uncertain Scope of Liability and Responsibility

In this section, we discuss the lack of understanding and consensus around the extent to which ICT companies should be held responsible for human rights impacts within the international community itself, an obstacle we link to the problem of causality. We begin with a discussion of how this deficit complicates our ability to evaluate the actions of ICT companies concerning human rights impacts. We then consider this observation in the context of the different approaches BSR and Article One take to determine whether Facebook *caused*, *contributed to,* or is *directly linked* to various human rights impacts their respective HRIAs identified. We argue that in the absence of greater consensus on this issue, and thus greater consensus on how to evaluate the actions of companies regarding these impacts, companies are unlikely to feel pressured to take any action at all.

In 2016, UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, David Kaye, presented the first report in a series intended to "provide guidance on how private actors should protect and promote freedom of expression in a digital age."[78] Recognizing that "[t]he contemporary exercise of freedom of opinion and expression owes much of its strength to private industry, which wields enormous power over digital space," the report raised a question that the international community has yet to resolve:

[77] *Ibid.*, p. 6.

[78] D. Kaye, «Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression.» Office of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, New York. A/HRC/32/38. May 11, 2016.

"to what extent should the [ICT sector] be responsible for the promotion and protection of freedom of opinion and expressions?"[79]

Private actors in the ICT sector "[e]nable connection to the Internet," "[h]ost information," "[f]acilitate aggregating, sharing and searching for information," "[p]roduce and regulate access to one's own content," "[c]onnect users and communities," and "[c]ollect, repurpose and sell data."[80] In doing so, these companies act as a "gateway for information and an intermediary for expression"[81] with outsized control and influence over the flow of information in modern society. This power leaves them vulnerable to state "pressures to conduct their businesses in ways that interfere with freedom of expression."[82] Even in the absence of such pressure, ICT companies independently possess the power to "advance or restrict rights," and pose a threat to human rights if they fail "to ensure the promotion and protection of rights in their pursuit of commercial interests."[83]

Moreover, in the past several years, we have seen how impacts on freedom of expression and access to information and privacy have the potential to impact a variety of other internationally recognized human rights. As evidenced by the current global pandemic, "violations of free expression and information rights can cause or contribute to the violation of other rights, such as the rights to life, liberty, and security of person."[84] Violations of privacy rights may "set off a chain reaction for the violation of other rights," such as non-discrimination, freedom of thought, freedom of association, and the right to take part in government.[85] In addition to these, the assessments we reviewed found human rights impacts on the right to assembly and association, children's rights, the standard of living, and access to culture. These threats are further complicated by the rapidly evolving nature of the technology industry and the constant innovation of products,

---

79 *Ibid.*

80 *Ibid.*, p. 6-8.

81 *Ibid.*, p. 3.

82 *Ibid.*

83 D. Kaye ("Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression," cit.), 3 ("States place undeniable pressures on the private information and communication technology sector that often lead to serious restrictions on the freedom of expression"). We should note that this is a rather recent development, especially if we consider ICT as a sector. For instance, in 2011, Martinuzzi et al. could say that the "most important social issues of the ICT sector are poor working conditions outside Europe." *See* A. Martinuzzi; R. Kudlak; C. Faber; A. Wiman («CSR Activities and Impacts of the ICT Sector.» Research Institute for Managing Sustainability, Vienna University, Vienna. 2011.), 1. This no longer seems to be the case.

84 N. Maréchal; R. MacKinnon; J. Dheere, «Getting to the source of infodemics: It's the business model, RANKING DIGITAL RIGHTS 24.» Ranking Digital Rights. May 2020.

85 *Ibid.*

services, and processes that determine the flow of information in the digital age.

Kaye's report concludes that "[a]s the project of exploring the [ICT sector's] responsibilities moves forward," businesses "should be evaluated on the steps they take both to promote and undermine freedom of expression, even in hostile environments unfriendly to human rights."[86]

However, this seemingly straightforward guidance ignores the reality that some human rights risks associated with ICT sector business operations are better understood than others. More specifically, it is oblivious to the fact that the international community itself lacks the understanding and consensus necessary to evaluate whether a business' actions constitute a step to promote or a step to undermine freedom of expression. And we would like to suggest something more: that whether a certain action goal is to promote or undermine one of the rights generally affected by technological companies depends on the eye of the beholder. In that sense, the lack of consensus we are highlighting is the outcome of extensive disagreements as to the scope and limits of freedom of expression and the viability and desirability of potential remedies to harms produced through the exercise of this right.

To illustrate, consider BSR's HRIA of Facebook's presence in Myanmar. As already noted, BSR determined that Facebook did not cause and does not cause any of the actual or potential adverse human rights impacts identified. Actually, the assessment found the company contributed to only one adverse human rights impact related to the company's operations, violations of freedom of expression, and even so, the company only contributed to this impact when it removed content from the platform in error. While BSR notes that the *cause*, *contribute to*, and directly *linked* designations come from Principle 19 of the UNGPs, the assessment itself does not define these terms. It does, however, detail the implications associated with each.

If the company is determined to have *caused* the impact identified, "the company should take the necessary steps to cease or prevent the impact." If the company is determined to have *contributed* to the impact, "the company should take the necessary steps to cease or prevent its contribution and use its leverage to mitigate any remaining impact to the greatest extent possible." If the company is determined to be *linked* to the impact through its products, services, operations, or business relation-

[86] D. Kaye, "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression," cit., p. 21.

ships, "the company should determine action, based on factors such as the extent of leverage over the entity concerned, how crucial the relationship is to the enterprise, the severity of the abuse, and whether terminating the relationship with the entity itself would have adverse human rights consequences."[87] These terms are elusive in general, but they are so particularly concerning the kind of actions that HRIAs usually scrutinize. In that sense, to be able to answer the relationship between a company's communication platform and an outbreak of ethnic violence depends to an extent on being able to determine what caused that outbreak of violence in the first place, a problem likely to have multiple causes. So, if we know, for example, that a particular region of the world suffers from ethnic violence, could efficient communication systems that allow for the spread of messages that feed into that violence cause, contribute, or be linked to the violence once that occurs? We believe there is reasonable space for disagreement in answering this question, and this is one of the challenges HRIAs in the ICT sector face.

Myanmar's Facebook HRIA underscores this point. The assessment provides "that Internet companies will often be linked to human rights impacts that they do not cause or contribute to. For example, Internet companies may be linked to hate speech, child sexual abuse material, and hacking that takes place over their platforms, even though they do not cause or contribute to these adverse human rights impacts themselves." The assessment then notes: "When a company is linked to human rights impacts, the UNGPs expect companies to take action, though the nature of the action will be very different than had the company caused or contributed to these impacts."[88] At the same time, the assessment concedes that "the actual relationship between content posted on Facebook and offline harm is not fully understood."[89] As previously noted, BSR's HRIA evaluating Facebook's presence in Cambodia does not present any determinations about the company's connection to the impacts identified.

In contrast, Article One's assessments suggest that the company may have contributed to some of the adverse human rights impacts identified but do not actually make that determination explicit. For example, the Sri Lanka assessment states "that the Facebook platform *contributed to* spreading rumors and hate speech, which may have led to 'offline violence.'"[90] In both assessments, Article

[87] BSR, "Human Rights Impact Assessment: Facebook in Myanmar 1," cit., pp. 33-34.

[88] BSR, "Human Rights Impact Assessment: Facebook in Myanmar 1," cit.

[89] *Ibid.*, p. 24.

[90] Article One, "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

One found that Facebook's response to impacts "was slow and, at times, insufficient, which *potentially exacerbated* the impacts" identified.[91] Referring to the UNGPs and Facebook's responsibility to exercise due diligence to mitigate adverse human rights impacts related to its operations, Article One's Sri Lanka assessment reasoned that "Facebook's lack of formal human rights due diligence in Sri Lanka prior to this HRIA and the limited cultural and language expertise among Facebook staff at the time of the May 2018 Kandy incident *may have contributed to offline harm* stemming from online engagement."[92] Further, both assessments conclude that impacts may have been "exacerbated by now phased out algorithms designed to drive engagement on the platform, regardless of the veracity or intention of the content."[93] Thus, while Article One's assessments suggest that Facebook may have contributed to some of the adverse human rights impacts identified, they do not make decisive determinations in this regard, at least not in the executive summaries made publicly available.

The difference between BSR and Action One impact assessments is difficult to evaluate. What led researchers in one case to a conclusion of "linkage" while in another case researchers reached a conclusion, albeit tentative, of "contribution"? Again, a fundamental piece of the puzzle is missing in this scenario: sufficient knowledge and understanding in terms of the social effects of these technologies. This is an old problem. In 2006, Kuhndt, von Geibler and Herrndorf identified this challenge when assessing the ICT contributions to the then millennium development goals, for the "existing singular approaches fail to account for causal chains linking companies' activities to societal outcomes."[94] The recommendations made at a time when what was sought was to uncover the beneficial impacts of the ICT sector should inform a quest of negative impacts: "The ICT sector should aim to provide information on all steps in the causal chain to demonstrate the existence of benefits from ICT and identify the most critical obstacles towards an extension of the opportunities that have been realized to date."[95]

---

[91] *Ibid.*; Article One, "Human Rights impact assessment: Facebook in Indonesia. Executive Summary," cit.

[92] Article One, "Human Rights impact assessment: Facebook in Sri Lanka. Executive Summary," cit.

[93] *Ibid.*

[94] M. Kuhndt; J. von Geibler; M. Herrndorf, «Assessing the ICT sector contribution to the millennium development goals: status quo analysis of sustainability information for the ICT sector.» Wuppertal Report. 3. August 2006. P. 20.

[95] M. Kuhndt; J. von Geibler; M. Herrndorf ("Assessing the ICT sector contribution to the millennium development goals," cit.), 20. Interestingly, Kuhndt et al. identify six indicators that should be developed in order to benchmark potential impacts: activity, access, user readiness, use, cause and effect and societal outcome. But, as they soon recognize, the indicators on cause and effect "depicting the concrete contribution of ICT provision and usage to environmental, social or economic outcomes are less developed than the other types of indicators." *See* M. Kuhndt; J. von Geibler; M. Herrndorf (*Ibid.*), 23

Some difficulties in this regard are related to companies' lack of transparency. For instance, one area that seems to be particularly difficult to scrutinize is the use of algorithms to determine what users see. But even in a case of full algorithm transparency establishing a direct link between a company and the most severe adverse human rights impacts—like security, freedom of expression, privacy, and nondiscrimination—might be very difficult, if not impossible.

Consider, for instance, the following human rights risks to security identified in BSR's Myanmar HRIA: "Misinformation and disinformation that is intended to incite or exacerbate violence or coordinate harm may be posted on the Facebook platform, but not discovered and removed."[96] It is very unlikely that assessors could determine with any certainty the degree to which Facebook's operating systems contributed to the offline harms this risk refers to. While there is a chance for adverse human rights impacts to be observed or documented, the connection between Facebook's algorithms and automated content review systems and offsite violence may be much more difficult to observe. This is further exacerbated by the operating systems themselves. Because no two users have the same experience on the platform, it is much more difficult to monitor the platform for abuse. A recent Ranking Digital Rights report on this issue states "Disinformation, misinformation, hate speech, and scams of all sorts are powerful precisely because digital platforms' automated content optimization systems aim them at just the people who are most vulnerable to these messages, while hiding them from other users who would otherwise be in a position to flag them."[97] In the absence of greater understanding and consensus around what constitutes steps to promote and what constitutes steps to undermine freedom of expression, HRIAs may struggle to influence the action of companies at all.

This point has been made before. Jørgensen et al. said that "it is challenging to identify the adverse human rights impacts in the sector due to the diverse range of activities, the physical and virtual footprints of the actors that comprise the industry, as well as the size and global character of affected rights-holders."[98] These features of the ICT explain why uncertainty in terms of evidence, causality and—thus—assessment is so difficult to achieve.

---

[96] BSR, "Human Rights Impact Assessment: Facebook in Myanmar 1," cit., p. 35.

[97] N. Maréchal; R. MacKinnon; J. Dheere, "Getting to the source of infodemics: It's the business model, RANKING DIGITAL RIGHTS 24," cit., p. 20.

[98] R. F. Jørgensen; C. B. Veiberg; N. ten Oever, "Exploring the role of HRIA in the information and communication technologies (ICT) sector," cit., p. 2.

> "Effectively, the role that many ICT companies play as gatekeepers and intermediaries in the online ecosystem implies that the way they prioritize, curate, remove, process, and share content affects what information users may communicate, access and view through their services and platforms … Likewise, self-regulatory measures to counter, for example, hate speech, affect the ways in which users encounter content and expression on sensitive topics … it remains an open question how freedom of expression concerns raised by corporate policy, design and engineering choices should be reconciled with the freedom of private entities to design and customize their platforms as they choose."[99]

This open question is not only because of insufficient research and difficult causal calls. It is also the outcome of profound disagreement regarding what constitutes a wrong, whether the wrong deserves a remedy and who should be in charge of enforcing it. Consider the following example. Suppose that in a given country, a right-wing, anti-immigration group grows and becomes more visible. It issues statements that are clearly racist, calling immigrants derogatory names, blaming them for rising crime, and so on. Its discourse is extreme and could be classified under different conceptions as *hate speech*. What should we do about it? This question has no easy answer. In the United States, for instance, hateful groups are allowed to exercise broad freedom of expression that allows for all sorts of hateful comments as long as these are not clearly related to an "imminent danger or great substantive evil."[100] Yet, in Germany, such a case would have a very different outcome, for hate speech is much more directly punished. Even if we would somewhat agree that the speech produces a wrong that renders it illegal, because under the Inter-American standards of human rights it is directly connected to an actual and likely exercise of violence, we would not agree on what remedy is most effective. For instance: is censoring or punishing this speech the best way of combating its evilness, or should we focus on producing the reasons that render this speech morally and politically wrong, objectionable, and reproachable? If reproach is called for, who should do it? The state or citizens?

The disagreement that exists around the scope and reach of freedom of expression protections, in particular, makes HRIAs very unlikely to be a useful tool

[99] *Ibid.*, p. 3.

[100] 7th Cir., *Collin v. Smith*, F.2d 578:1202 (1978). Retrieved from: https://supreme.justia.com/cases/federal/us/432/43/.

in assessing these sorts of wrongs that pit important values (such as equality or anti-discrimination) against different conceptions of free speech. Depending on the one we choose or are bound by, the chips would fall in very different places.

3. HRIAs Procedural Transparency (And Lack Thereof)

The final section of our analysis focuses on the general lack of transparency and consistency regarding process, methodology, and stakeholder engagement concerning the implementation of HRIAs in the ICT sector. We begin this discussion by considering the importance of transparency, external participation, and independent monitoring in assuring effective human rights due diligence. Next, we present findings from a 2015 study examining the different ways companies conduct human rights due diligence and consider the implications of these findings in the context of the ICT sector.

In order to understand the significance of transparency and external participation in the human rights due diligence process, it is essential to understand the "origins of the debate around corporate human rights performance," and the "relatively low level of consensus around shared values and appropriate standards of good practice in HRIA."[101] Because "the failure and inadequacies of existing international mechanisms for effectively tackling … human rights issues are well catalogued," and the use of "various self-regulatory processes," have also been ineffective, "there is likely to be intense scrutiny of companies utilizing HRIAs in order to meet their HRDD obligations." [J. Harrison.[102] Relevant stakeholders will want to be reassured that an HRIA represents a robust and meaningful process that will identify the most significant human rights risks and concerns, even where the identification of such issues might be seen as causing problems for the companies concerned."] Because HRIAs are likely to be met with skepticism," it becomes even more imperative to examine the degree to which there are shared understandings in relation to the HRIA process, and that these shared understandings can act as a firm foundation for future practice in the field."[103]

In this regard, transparency is essential because "assessors conducting due dili-

[101] J. Harrison, «Establishing A Meaningful Human Rights Due Diligence Process For Corporations: Learning From Experience Of Human Rights Impact Assessment,» Impact Assessment and Project Appraisal, vol. 31, 2, 2013, Retrieved from http://www.tandfonline.com/doi/abs/10.1080/14615517.2013.774718.

[102] *Ibid.*

[103] *Ibid.*

gence processes need to learn from each other about how those processes should function."[104] Transparency enables businesses to answer questions like: "what indicators should be utilized…? What evidence is required to support conclusions about impacts concerning particular aspects of corporate activity…? What form should recommendations take to ensure that they are likely to be acted upon by those who will enact change?"[105] Transparency is also necessary for external actors to assess and evaluate a business's performance. Furthermore, "[t]here is a need to distinguish carefully between a public and verifiable human rights impact assessment process and a private human rights risk assessment conducted by companies purely for their own internal use."[106] While these may be helpful, they "should not be seen as a form of due diligence that would meet the requirements of the Ruggie Framework without, at least, the additional transparency requirements set out above being adopted."[107] Moreover, while "[t]here will be occasions where there is a need to keep some information secret and not to publish it as part of the reporting process… the situations in which this will be necessary should be relatively rare, and the information that needs to be kept secret should not be extensive."[108]

Likewise, external participation is essential to HRIAs both because "stakeholders are vital to the credibility of HRIAs," and because "affected individuals and communities are often in the best position to inform the HRIA process about the actual and potential human rights impacts of corporate actors upon them."[109] Because "external civil society pressure" is usually the force that brings global attention to human rights issues relating to a business's operations, "[i]t is therefore vital that these stakeholders are also able to engage with the HRIA as verifiers of the credibility of its methodology and the reliability of its results."[110]

Recognizing the "limited information available" about how companies conduct human rights due diligence in practice, in 2015, researchers at the British Institute of International and Comparative Law (BILCL) partnered with the Business Ethics & Anti-Corruption Group of the global law firm Norton Rose Fulbright

[104] *Ibid.*, p. 112.

[105] *Ibid.*

[106] *Ibid.*

[107] *Ibid.*

[108] *Ibid.*, p. 113.

[109] *Ibid.*

[110] *Ibid.*

LLP to study "to what extent, if at all, companies are currently undertaking" human rights due diligence. [Mccorquodale 2017. This article presents findings from a study informed by "combined desk-based literature, policy and legal research," "qualitative insights obtained empirically through surveys and interviews." The survey was completed by 152 respondents between June and July 2015 and was followed by "semi-structured interviews" with 14 people "working in companies at a senior level and with knowledge of [human rights due diligence] practices within their company, as well as a roundtable discussion." Interviews were conducted between October 2015 and January 2016.] The survey found that almost half of respondents had never undertaken any human rights due diligence process or conducted an HRIA.[111] Of these respondents, 72.34 percent "indicated that they had considered human rights as part of other processes, predominantly workplace health and safety, labour rights, equality and non-discrimination, and community, indigenous or land rights."[112] The study found that respondents whose companies had conducted human rights due diligence were significantly more likely to identify actual or potential human rights than those whose companies had not "but considered human rights as part of other due diligence processes".[113] While "77.14 percent of companies using dedicated [human rights due diligence] identify adverse impacts, 80.77 percent of companies using non-specific [human rights due diligence] *do not* identify adverse impacts."[114]

The study cites various examples of regulatory frameworks mandating corporate due diligence related to human rights but ultimately concludes that compliance with these regulatory requirements is unlikely to actually satisfy a company's responsibility to conduct human rights due diligence.[115] While various sector-specific regulations require due diligence "for activities which may implicitly have human rights impacts," compliance with these types of regulations alone will miss key components of human rights due diligence. To illustrate, consider the European Union Transparency Directive which mandates country-by-country reporting on material payments made to governments by large extractive and

[111] R. McCorquodale, «Human Rights Due Diligence In Law And Practice: Good Practices And Challenges For Business Enterprises,» Business and Human Rights Journal, vol. 195, 2017, p. 205.

[112] *Ibid.*

[113] *Ibid.*

[114] *Ibid.*, p. 206.

[115] United Kingdom's Modern Slavery Act 2015 (discussed in Paper 1), California Transparency Supply Chains Act 2010 (CTSC),

logging companies.[116] While this reporting "may improve human rights conditions on the grounds indirectly, this type of regulation does not require the identification or addressing of human rights impacts *within the company's own operations* or within its value chain."[117] Furthermore, "most of this legislation and regulation is not in express terms of human rights and generally requires reporting by companies without expressly requiring companies to address and remediate their human rights impacts."[118]

As stated at the onset of this paper, research on how HRIAs are being implemented in the ICT sector is necessarily limited to what companies themselves make public. Facebook has made four HRIAs publicly available, though three out of those four HRIAs were released in an eight-page executive summary format, and these executive summaries did not make determinations about Facebook's relationship to the impacts those HRIAs identified. Google has made one HRIA publicly available, though as previously mentioned, we are limited to a four-page summary of the assessment, and it does not evaluate Google itself. Telefonica's website indicates that it "rel[ies] on external experts for [HRIAs]," and that BSR conducted one such assessment for the company in 2013 and Business and Human Rights (BHR) conducted another one in 2018.[119] Telia has made three HRIAs publicly available in full: a twenty-page summary project report on the company's divestment plan for the Eurasia business region, and two full forty-page HRIAs on the company's presence in Lithuania and Sweden. Microsoft stands out in this regard. The company publishes an annual human rights report that provides details on the most salient human rights risks; however, these reports do not make determinations about the company's connection to the impacts.

Thus, there is a general lack of transparency around which HRIAs ICT companies release to the public, in what format, and when. This lack of transparency is problematic because it hinders the potential for assessors to learn from other HRIAs. It also makes it difficult to evaluate a business's performance and assess whether it is carrying out its responsibility to respect human rights or not. This lack of transparency extends to the stakeholders involved in the process. None of

[116] *Id.* Similar regulations include: Section 1501 of the US Dodd-Frank Wall Street Reform and Consumer Protection Act (requirements); Illegal Logging Prohibition Act in Australia (requirements)

[117] R. McCorquodale, "Human rights due diligence in law and practice: Good practices and challenges for business enterprises," cit., p. 202.

[118] R. McCorquodale, "Human rights due diligence in law and practice: Good practices and challenges for business enterprises," cit.

[119] Retrieved from: https://www.telefonica.com/en/web/responsible-business/human-rights

the HRIAs reviewed disclose which stakeholders were consulted in the process which may jeopardize the credibility of the observations and recommendations made. Moreover, because of the technical considerations at issue, the need for external consultation with engineers and other experts with the knowledge to advise assessors may make or break the validity of the findings in an ICT sector HRIA. Perhaps most importantly, as the HRIAs released by Facebook tend to show, it is not entirely clear why certain HRIAs are conducted and released to the public. We have no insight into what process the companies are following and when we can expect to receive communications about updates on the development that will enable us to track the progress these companies are making regarding their responsibility to respect human rights. In a sense, this sporadic release of seemingly randomly chosen assessment targets may ensure that the public cannot, in fact, make informed judgments about what human rights impacts the ICT sector is taking accountability for and what steps they are taking in response to that recognition of accountability.

F. Conclusion: The Challenges Ahead

The trends that surfaced in our review are interconnected and likely the direct result of a framework that imposes no legally enforceable duties on businesses regarding their expectations to protect human rights. The system is further hindered by what appears to be an excess of the discretion it grants to businesses in an attempt to promote compliance. As a result, we have limited visibility into the work businesses are actually doing to comply with their responsibility to respect human rights. This leaves the international community and rights-holders alike unable to assess whether the framework itself is adequate to address the human rights violations related to businesses worldwide. As the demand for accountability for human rights risks associated with businesses continues to grow, we need a way to evaluate the strengths and weaknesses of the current framework.

This problem is likely to have many causes. We already pointed to one: the voluntary nature of Ruggie's framework means that it is for companies to decide the scope and reach of their human rights programs, including their HRIAs.[120] As Bittle and Snider pointed out, it is impossible to ignore the "structural con-

[120] R. Álvarez Ugarte; L. Krauer, "El sector TICs y derechos humanos: hacia un marco conceptual de las evaluaciones de impacto en DDHH," cit.

tradiction between the corporation's legal obligation to maximize profits for its shareholders and its non-mandatory human rights obligations."[121] For them, this "is a major weakness in Ruggie's work: the corporation is legally committed to upholding the 'laws of the market, not … human rights standards.'"[122] In that sense, this *weakness* is the outcome of Ruggie's framework with regard to the decade-long fight over corporations' responsibility for human rights abuses at the heart of the UN, a point we discussed in our previous paper.[123] But this point should not be denied the relevance it deserves for having already been made: on the contrary, we still believe that as long as HRIAs keep being closer to the world of corporate social responsibility than human rights law, these inherent limits are not only to be expected but are, to an extent, unavoidable.

This does not mean that HRIAs and the corporate practice built around them in the ICT sector is not to have a regulatory impact, especially if we define it as the capacity of a practice to generate rule-like criteria that actors within the field routinely obey.[124] This definition would suggest that certain voluntary codes of conduct *may* reach a degree of acceptability within an industry that plays an effective regulatory function, even if it is not in the form that regulation usually happens. But that future is still distant, and we believe that with the information available to us now, it is difficult to predict whether that stage will be reached in the current path of events. We believe that the future hangs on one particular issue: the way dominant Internet companies are to be regulated. It is no longer a question of *if*; the one on *when* also appears close enough. But it is a matter of what seems to be more relevant and uncertain.

Our informed guess is that the *form* this regulation will take is most likely a mixture of different regulatory approaches. There is no pure form waiting at the end of the road, neither the state-based legal model nor the international treaty; neither corporate self-regulation nor no regulation at all. Following Marsden et al., we believe these companies will be regulated in different ways for different purposes, and a mixture of self-regulation, audited self-regulation, formal self-regulation, or

[121] S. Bittle; L. Snider, «Examining The Ruggie Report: Can Voluntary Guidelines Tame Global Capitalism?», Critical Criminology, vol. 21, 2, 2013, p. 188, Retrieved from https://doi.org/10.1007/s10612-013-9177-4.

[122] *Ibid.*

[123] R. Álvarez Ugarte; L. Krauer, "El sector TICs y derechos humanos: hacia un marco conceptual de las evaluaciones de impacto en DDHH," cit.

[124] This broad take on "regulation" is somewhat accepted within the field. *See* R. Baldwin; M. Cave; M. Lodge (eds.) (*The Oxford Handbook Of Regulation*, Oxford University Press, Oxford, 2010 Retrieved from http://doi.wiley.com/10.1111/j.1467-9299.2012.02049.x), 2-3

co-regulation will end up controlling the regulatory landscape.[125] Perhaps in that amalgamation of rules and practices, HRIAs will have a role to play.

But before moving forward in exploring possible futures, it is necessary to distinguish between companies that are non-controversially regulated by nation-states such as e.g. telecommunication companies, and those that are either not regulated by nation-states or are regulated in a way that is inconsistent, problematic, or controversial. Platforms such as Facebook or companies such as Google, whose HRIAs we have looked at more extensively in the previous sections, fall clearly in the latter category. These companies are regulated not as a specific industrial sector (as, e.g., telecommunication companies with telecommunications frameworks) but as a byproduct of concerns not directly linked to the technology itself. This is what happens, for instance, with hate speech laws in Europe or the approaches to disinformation in Germany and France that reach Internet companies in the same (or similar) way they reach media outlets. Sometimes, these laws apply specifically to Internet companies because legislators perceive that their reach or impact is especially troubling (as in the examples on hate speech and disinformation). Sometimes, regulations seek to exempt these companies from a more general liability regime, such as Section 230 of the *Communications Decency Act* of 1996 and the legislations around the world that followed that model. These are statutory regulations that do not affect the full operation of the companies of the second group but do reach some of its products and activities. It is important to keep in mind these two kinds of companies for uncertainty only haunts (or primarily haunts) companies in the second group. The following, then, is developed with that specific uncertainty in mind.

What will the future regulation for these companies look like? We outline thus four possible scenarios, for different issues, following Marsden et al. model.

1. Corporate social responsibility will likely remain a fundamental part of the approach in the ICT sector. Companies generally enjoy greater leeway, in the sense that is a practice that is a hundred percent under their control, and actions that fall under the broad CSR label usually yield rewards that can be translated in a way that follows corporations' monetary goals.

2. Non-audited self-regulation. Marsden et al. mention within this category (a) industries' codes of practice and (b) transparency reports.[126] The latter is a

[125] C. Marsden; T. Meyer; I. Brown, "Platform values and democratic elections," cit., pp. 12-13.

[126] *Ibid.*, p. 12.

common practice that, in the last few years, has consistently improved within the main players of the ICT sector, even though we believe that there is room for further improvement. The former, on the other hand, has not yet fully materialized but some initiatives may be considered as having an *industry* reach. This is what happens, for instance, with the Manila Principles on intermediary liability, and what may happen with the Santa Clara principles on content moderation. If enough players within the industry abide by these general statements of purpose that—on occasion—provide specific guidance and principles useful to guide corporate conduct, then something akin to (a) would emerge.

3. Audited self-regulation. This happens when a body exerts some sort of oversight function. Again, in a world that looks very little like the regulation usually advanced by the nation-state, this may look strange. Marsden et al., for instance, identify the 2018 European Code of Practice and the audit reports of GNI as falling within this category.[127] Indeed, the GNI does appear to be an external body designed to include different stakeholders that effectively assess the behaviors of members every two years. Even though its approach is limited to privacy and freedom of expression and fails to consider other human rights,[128] insofar as it exerts some sort of oversight function over member companies, GNI appears to be a kind of self-regulation that is audited by a third party.

4. Co-regulation. This is a strange animal: an industry-wide code of conduct that has been approved by a traditional regulator, whether it is a legislature or an administrative agency. We believe it does not fall within the Marsden et al. model, but there has been a practice that seems to be close to this form of regulation. We mean the practice of general principles like e.g., the Manila or Santa Clara declarations, which eventually inform national legislation. The connection might not be direct: for instance, legislators may simply follow these general principles because they believe that doing so would be good law-making, or they may do so because they were lobbied by companies or some other reason. Perhaps those principles are not even mentioned in the regulation at hand. But the correlation is a good antecedent to this type of regulation.

5. Statutory regulation. Many aspects of the Internet are regulated by statutory regulation, but this applies unevenly depending on the country doing

[127] *Ibid.*

[128] R. F. Jørgensen; C. B. Veiberg; N. ten Oever, "Exploring the role of HRIA in the information and communication technologies (ICT) sector," cit., p. 6.

the regulation. In that sense, as we pointed out in the first paper—and Ruggie's framework rightly acknowledges—not all countries enjoy *in practice* the same regulatory capacity they have *in theory*. This is because of a myriad of reasons, but market share seems to be a huge driver of the divide.

Issues & Regulation

	1	2	3	4	5
Disinformation	1	1	0	0	1
Hate speech / anti-discrimination	1	1	0	0	0
Content moderation	0	1	0	0	0
Defamation	0	0	0	0	1
Privacy - Data Protection	0	0	0	0	1
Privacy - "Revenge porn"	0	0	0	0	1
IP	0	0	0	0	1
Right to be forgotten	0	0	0	0	1

In the meantime, as we wait for the future to arrive, we believe companies could improve their approach to HRIAs, especially concerning three areas in which improvements are called for: more transparency, a more developed view of causality based on research, and industry-wide consistency.

First, companies should be more transparent about how they use HRIAs when they decide to engage in these sorts of analyses, the kind of internal work that happens before it is carried out, and their effects on company policies. We do not know if companies are following the detailed process described by Abrahams and Wyss[129] or less structured alternatives. In any case, more transparency is called for and it is lacking. The fact that some HRIAs have been revealed in full but most have been shared in the form of an executive summary should make the point by itself.

Second, it is fundamental that HRIAs explore in more detail the difficult question of causality. As we pointed out before, the language that Ruggie's framework lends to HRIAs is especially limited to assess the impact of some technologies on human rights, particularly in the ICT sector and especially regarding platforms and services whose usage is still not fully understood. The assessment that they

[129] D. Abrahams; Y. Wyss, «Guide to Human Rights Impact Assessment and Management.» International Business Leaders Forum, International Finance Corporation & el Global Compact de las Naciones Unidas, Washington D.C. 2010.

do not *cause* or *contribute* but are *linked to* harms to human rights is an assessment that could be made generally regarding any technology that is used by people as a tool to perform certain actions. In that sense, the HRIAs published by Facebook are very telling on this point: they reach the same conclusion, over and over, for they are based on a story of causation in which Facebook simply offers a tool that can be used *for good or evil*. To an extent, the usage people make of these sorts of technologies should not make the company in charge of the technology strictly liable. But should they be *responsible* for the technology they develop? This is a question with no obvious answer. On the one hand, we all accept that some degree of responsibility is called for, e.g., to deal with child pornography. But, on the other, we all disagree as to the kind of involvement companies should have regarding other issues, such as e.g. the building of backdoors on encryption technologies. It is unclear where the chips will fall regarding several of these issues, but there is no doubt that big technology companies are *taking responsibility* in ways that will clearly have an impact on the regulatory outcomes of the future. If they move in this direction, further research is needed to better understand the relationships of causality that lie within the harms that Ruggie's framework seeks to prevent.

Third, our survey of HRIAs suggests that this is a tool that companies adopt willingly, and nothing forces them to take this particular path. An exception to this general principle would be GNI — companies within the initiative seem to be under some degree of encouragement to make due diligence analyses and perhaps use this tool to assess the way their products affect human rights in different contexts. But GNI's role as a *self-regulator* in the Marsden et al. terms[130] is not yet fully evident. We believe it is a development that would strengthen HRIAs as a tool if it happens.

[130] C. Marsden; T. Meyer; I. Brown, "Platform values and democratic elections," cit., p. 12.