



# RANKING DIGITAL RIGHTS

## **Submission to OHCHR on the application of the UN Guiding Principles on Business Human and Human Rights to tech companies**

This document is submitted in response to the High Commissioner's call for stakeholder input on the practical applications of the United Nations Guiding Principles on Business and Human Rights (UNGPs) to technology companies. Our submission focuses on the first theme of the consultation: human rights risks associated with tech company business models. Given our substantive expertise, we focus specifically on tech companies in the digital platform and telecommunications sectors, offering a far-from-exhaustive list of examples using the Shift Business Model Red Flags as an organizing principle. We are also including an unpublished book chapter (authored by RDR staff) that examines the influence of Meta's surveillance advertising business model on corporate decision-making, and the resulting human rights harms.<sup>1</sup> Our hope is that the present document and accompanying book chapter will provide the B-Tech Project team with illustrative examples of the connection between companies' business activities and human rights harms.

Ranking Digital Rights (RDR) is a non-profit research and advocacy program at New America that works to advance freedom of expression and privacy on the internet by establishing global standards and incentives for companies to respect and protect the human rights of internet users and their communities. We carry out this mission by researching and analyzing the commitments, policies, and practices of major global digital platforms and telecommunication firms based on international human rights standards.<sup>2</sup> In addition to our research, we also advocate for laws and public policies that safeguard the fundamental rights of online users and their communities.<sup>3</sup>

Our ranking methodology is based on international human rights standards and frameworks, including the Universal Declaration of Human Rights, the International Covenant on Civil and Political Rights, the European Convention on Human Rights, and the UNGPs. To date, RDR has produced five RDR Indexes (2015, 2017, 2018, 2019 and 2020) and offers the only year-on-year human rights ranking of these corporations. Research for the seventh iteration, now divided into the Big Tech and Telco Giants Scorecards, is currently underway. The Big Tech Scorecard will be released in April 2022.

---

<sup>1</sup> Maréchal, Nathalie, Alex Rochefort & Leandro Ucciferri. "The Business Model is the Message: Reconfiguring the Enabling Environment for Hate Speech on Social Media." Forthcoming.

<sup>2</sup> For more on our Rankings please visit <http://rankingdigitalrights.org>.

<sup>3</sup> Ranking Digital Rights. *Our Principles* (last accessed February 23, 2022), <https://rankingdigitalrights.org/about/principles/>.

The UNGPs offer a clear framework for corporate accountability by emphasizing the duty of the state to protect citizens from negative consequences of business activity, the responsibility of corporations to self-assess and prioritize transparency, and the importance of access to remedy through appropriate channels.<sup>4</sup> These concepts are clearly applicable to technology companies but must be carefully evaluated in order to take into account the specific human rights risks that digital platform and telco business models create, as well as the constantly evolving context within which policy must be created to regulate platforms.

The UNGPs establish the expectation that companies will do their own assessment through human rights due diligence in order to minimize human rights abuses caused by their policies, practices and products.<sup>5</sup> Benchmarking organizations including RDR echo this by highlighting the importance of human rights due diligence (HRDD) through impact assessments.<sup>6</sup> In the 2020 Index and in the upcoming 2022 Rankings, RDR's methodology includes a set of indicators about HRDD, which measure whether a company discloses that it conducts human rights impact assessments (HRIAs) specifically related to digital rights and government policies and regulations, policy enforcement, targeted advertising, and algorithmic systems.<sup>7</sup> In our 2020 evaluation, Facebook was the only one of the 26 ranked companies to receive any points on the indicator related to HRIAs of targeted advertising systems.<sup>8</sup> Companies fared slightly better on due diligence on algorithmic systems, where Telefónica distinguished itself by receiving all possible points. The other six companies that received credit for this indicator scored 25% or less.<sup>9</sup> The goal of these assessments is for companies to identify and address some of the risks associated with their activities before they become entrenched. These paltry scores strongly suggest that doing so is currently the exception, not the rule. It should come as no surprise that human rights harms associated with tech company business models are rampant.

## Applying the UNGPs to tech company business models

The Shift Business Model Red Flags project explores the potential for human rights risks that a company's business model may present.<sup>10</sup> Several of these red flags are prevalent and particularly worrisome throughout the ICT sector.<sup>11</sup> This submission explores how rights to personal data protection and non-discrimination are regularly threatened by elements of digital platforms' and telecommunications companies' business models including the surveillance

---

<sup>4</sup> United Nations Office of the High Commissioner for Human Rights. *Guiding Principles on Business and Human Rights* (2011), [https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr\\_en.pdf](https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_en.pdf).

<sup>5</sup> United Nations Office of the High Commissioner for Human Rights. "Addressing Business Model Related Human Rights Risks: A B-Tech Foundational Paper" (2020), [https://www.ohchr.org/Documents/Issues/Business/B-Tech/B\\_Tech\\_Foundational\\_Paper.pdf](https://www.ohchr.org/Documents/Issues/Business/B-Tech/B_Tech_Foundational_Paper.pdf).

<sup>6</sup> Ranking Digital Rights. "2020 indicators, G4" (2019), <https://rankingdigitalrights.org/2020-indicators/#G4>.

<sup>7</sup> Ibid

<sup>8</sup> Ranking Digital Rights. "2020 indicators, G4c" (2019), <https://rankingdigitalrights.org/index2020/indicators/G4c>.

<sup>9</sup> Ranking Digital Rights. "2020 indicators, G4d" (2019), <https://rankingdigitalrights.org/index2020/indicators/G4d>. These companies were: Facebook, Microsoft, Verizon Media, Apple, Deutsche Telekom, and AT&T.

<sup>10</sup> Shift Project. "Business Model Red Flags" (2021), <https://shiftproject.org/resource/business-model-red-flags/red-flags-about/#chapter>.

<sup>11</sup> Shift Project. "Business Model Red Flags: Menu of Red Flags" (2021) <https://shiftproject.org/resource/business-model-red-flags/menu-of-red-flags/#chapter>.

advertising industry’s symbiotic relationship with these companies, problematic content moderation algorithms priorities, and the creation of products that can be easily abused. For ease of reference, this section follows Shift’s typology of Business Model Red Flags. Note that this list isn’t exhaustive, and merely reflects RDR’s analysis of the risks with which we have the greatest familiarity.

### ***Red Flag #5 Algorithmic Discrimination***

Online platforms introduce risks like “algorithmic decision-making to profile, and make predictions about, people in ways that can result in discrimination or other human rights harms” when collecting personal information like demographic data from users.<sup>12</sup> This risk is inherent to surveillance advertising systems that display ad content based on demographic and behavioral information about individual internet users, whether that information is collected or inferred based on other data points. A 2016 study from *ProPublica* demonstrated how advertisers can intentionally discriminate against users based on personal data like race.<sup>13</sup> Facebook’s due diligence around ad-targeting tools clearly failed to take into consideration the possibility that bias could be introduced by advertisers.<sup>14</sup> This dynamic also operates in the context of algorithmic systems for recommending non-ad (user-generated) content and for moderating content that may violate platform rules. For example, Facebook algorithms struggle to distinguish between hate speech and posts calling out racial abuse, frequently sending Black users to “Facebook jail” for discussing harms. The algorithm’s “race-blind” approach shuts down dialogue about personal experiences with many users of color citing this as motivation to leave the platform.<sup>15</sup> A 2020 audit of Facebook’s content moderation and “race-blind” algorithms found that content criticizing groups like white men was taken down with much greater frequency than any related to marginalized groups. The “worst of the worst” project was then created to prioritize removing content disparaging groups who have been historically marginalized and are more likely to be the victims of hate speech and offline harms.<sup>16</sup>

### ***Red Flag #6 Potential for online and offline harm: Facebook in Myanmar and Ethiopia***

Silicon Valley CEOs and leaders in the tech industry often describe online connection as an unequivocal good. Their failure to acknowledge the risks of groups inciting hatred and violence and the specific role their companies’ business models play in amplifying this kind of content

---

<sup>12</sup> Shift Project. “Business Model Red Flags: Red Flag #5” (2021), <https://shiftproject.org/resource/redflag-05/>.

<sup>13</sup> Angwin, Julie & Terry Perris Jr. *ProPublica*. “Facebook Lets Advertisers Exclude Users by Race” (2016), <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>.

<sup>14</sup> Maréchal, Nathalie & Ellery Roberts Biddle. Ranking Digital Rights. *It’s Not Just the Content, It’s the Business Model: Democracy’s Online Speech Challenge* (2020), [https://d1y8sb8igg2f8e.cloudfront.net/documents/REAL\\_FINAL-Its\\_Not\\_Just\\_the\\_Content\\_Its\\_the\\_Business\\_Model.pdf](https://d1y8sb8igg2f8e.cloudfront.net/documents/REAL_FINAL-Its_Not_Just_the_Content_Its_the_Business_Model.pdf).

<sup>15</sup> Dvoskin, Elizabeth, Nitasha Tiku, & Craig Timberg. *The Washington Post*. “Facebook’s race-blind practices around hate speech came at the expense of Black users, new documents show” (2021), <https://www.washingtonpost.com/technology/2021/11/21/facebook-algorithm-biased-race/>.

<sup>16</sup> Dvoskin, Elizabeth, Nitasha Tiku, & Heather Kelly. *The Washington Post*. “Facebook to start policing anti-Black hate speech more aggressively than anti-White comments, documents show” (2020), <https://www.washingtonpost.com/technology/2020/12/03/facebook-hate-speech/>.

has devastating human rights impacts. Technology and ICT sectors are more likely to see risks like “providing online platforms for individuals where use of the platform can lead to harm to human rights.”<sup>17</sup> Inflammatory and harmful content thrives on online platforms due to the very specific structure of their business model. Metrics like Average Revenue Per User (ARPU) and engagement measure success for investors and advertisers respectively. Content that generates strong reactions and that users spend time engaging with is boosted and spread more widely, often exacerbating potential harms.<sup>18</sup> Discussing the Facebook Papers before the Senate Subcommittee on Consumer Protection in October, whistleblower Frances Haugen described consequences of this system saying, “It is causing teenagers to be exposed to more anorexia content. It is pulling families apart. And in places like Ethiopia, it's literally fanning ethnic violence’.”<sup>19</sup>

Social media platforms have long struggled with trolling and hateful rhetoric, seemingly torn between a simplistic understanding of free speech and protecting human rights both online and offline.. And in many cases, companies have simply neglected to put in place the appropriate mechanisms to enforce their own rules for user content. As a result, unchecked hate speech devolving into genocidal rhetoric in online communities has deadly real world consequences. In Myanmar, Facebook infamously failed to consider the importance of hiring for skills in native languages or regional familiarity in order to properly address nationalistic posts from political and religious leaders calling for violence against the Rohingya population.<sup>20</sup> Consequently, a class action lawsuit for 150 billion pounds moving forward in the U.S. and the UK on behalf of the Rohingyas could hold Facebook responsible.<sup>21</sup> Similar problems were evident in Ethiopia, and it became clear through documents that Frances Haugen produced to the U.S. Congress that Facebook was aware of how its platform has been used to amplify violent rhetoric in the Tigray conflict.<sup>22</sup> A study from Amnesty International explored the “significant rise in other recent social media posts inciting violence and using ethnic slurs against Tigrayans, some of which have gone unchecked. According to analysis from a local human rights organization, influential people such as journalists and political figures were among those sharing the posts, and social media platforms were slow to remove them.”<sup>23</sup> These are some of the worst case examples of how company decisions like failure to create nuanced content moderation, centralizing resources in powerful markets, and ignoring due diligence responsibilities can create an environment for

---

<sup>17</sup> Shift Project. “Business Model Red Flags: Red Flag #6” (2021), <https://shiftproject.org/resource/redflag-06/>.

<sup>18</sup> Allyn, Bobby. *NPR*. “Here are 4 key points from the Facebook whistleblower’s testimony on Capitol Hill” (2021), <https://www.npr.org/2021/10/05/1043377310/facebook-whistleblower-frances-haugen-congress>.

<sup>19</sup> *Ibid*

<sup>20</sup> Maréchal, Rochefort & Ucciferri, forthcoming.

<sup>21</sup> Milmo, Dan. *The Guardian*. “Rohingya sue Facebook for £150bn over Myanmar genocide” (2021), <https://www.theguardian.com/technology/2021/dec/06/rohingya-sue-facebook-myanmar-genocide-us-uk-legal-action-social-media-violence>.

<sup>22</sup> MacKintosh, Eliza. *CNN*. “Facebook knew it was being used to incite violence in Ethiopia. It did little to stop the spread, documents show” (2021), <https://www.cnn.com/2021/10/25/business/ethiopia-violence-facebook-papers-cmd-intl/index.html>.

<sup>23</sup> Amnesty International. “Ethiopia: Sweeping emergency powers and alarming rise in online hate speech as Tigray conflict escalates” (2021) <https://www.amnesty.org/en/latest/news/2021/11/ethiopia-sweeping-emergency-powers-and-alarming-rise-in-online-hate-speech-as-tigray-conflict-escalates/>.

violence to grow.<sup>24</sup> There are many more, almost certainly including some that have not been publicly reported.

The harms carried out on these social media platforms and the subsequent real world dangers demand deep examination into their root causes. This goes beyond hiring a few more content moderators who speak relevant native languages. Risk and impact assessments that carefully examine local contexts are key. Additionally, stakeholder engagement with diverse groups like ethnic and religious minorities in order to learn how their services may be negatively impacting users in different communities, and identify ways to mitigate these harms. Finally, transparency about content moderation should be expected as a first step toward accountability. By making this type of information available, governments and the public can have a more nuanced understanding of the inner workings of the digital ecosystem. When governments regularly demand removals that limit the expression of activists or minority groups, that should be disclosed. If there is a large gap between the company's policy on hate speech and what is actually removed from the site, that should be disclosed.<sup>25</sup> All of these efforts towards greater transparency are important steps to realizing corporate responsibility to respect human rights.

### ***Red Flag #9 Products that cause harms when misused: Location tracking***

A major concern, as technology seems to develop at a far more rapid pace than regulation and certainly than due diligence, is products that cause harm when they are misused. While tools developed by technology companies, particularly features of smartphones like Apple's Find My Friends or AirTag, have practical safety uses, they are easy to abuse. This type of location tracking technology is almost too good to be true for domestic abusers or stalkers. AirTags are small devices that can be easily concealed in a car or bag, making all the protections in place to disable them useless as a victim would have no way of knowing they are a target.<sup>26</sup> Sometimes framed as outlier behavior that only a minority of users would engage in, risks like stalking, domestic violence, wrongful arrests, and activist suppression are major violations of human rights, and tools that can facilitate them must be carefully considered through due diligence before they enter the market.

### ***Red Flag #16: Data misuse: government surveillance***

User rights to privacy are under constant assault thanks to the surveillance advertising business model that is prevalent across the digital platforms sector, and increasingly the telecommunications sector as well. Technology companies live off of advertising dollars and are therefore incentivized to collect personal information on users to inform targeted advertising. The existence of these data troves increases the risk of a company misusing user data, whether

---

<sup>24</sup> Maréchal, Rochefort & Ucciferri, forthcoming.

<sup>25</sup> Ullman, Ilana, Laura Reed, and Rebecca MacKinnon. Ranking Digital Rights. "Submission to the UN Special Rapporteur for Freedom of Expression and Opinion David Kaye: Content Regulation in the Digital Age" (2017), <https://rankingdigitalrights.org/wp-content/uploads/2018/01/RDR-2018-David-Kaye-Submission.pdf>.

<sup>26</sup> Cerullo, Megan. *CBS News*. "Apple AirTags are being used to stalk people" (2022), <https://www.cbsnews.com/news/apple-airtags-stalk-strangers/>.

they choose to or are compelled to.<sup>27</sup> Google pioneered the practice of “transparency reporting” in 2010 when it started reporting the number of government demands it received, and complied with, from governments around the world. In 2013, Edward Snowden famously revealed the extent of the US government’s legal and extralegal access to the global communications infrastructure, both with and without the knowledge of the companies involved. Since then, other governments have only redoubled their efforts to obtain similar access to digital communications, with grave consequences for human rights, the rule of law and democracy itself.

As for telecommunications companies, they have access to broad streams of user information from call records, text messages, internet browsing history, location data tied to mobile phone use, and financial information for billing purposes.<sup>28</sup> They are also increasingly moving into the targeted advertising space, as our 2020 research demonstrates.<sup>29</sup> In 2018, Verizon, AT&T, Sprint, and T-Mobile pledged to stop selling location data to third party location aggregators accused of misusing location tracking capabilities, only to turn around and keep doing it.<sup>30</sup> Recently, Norwegian telecommunications company Telenor has been criticized for the impending transfer of user data in Myanmar to the repressive military government. Claiming that refusal would put employees based in the country in danger, Telenor plans to sell its operations in Myanmar entirely, including the information of almost 18 million users.<sup>31</sup> Whether telecommunications companies choose to sell personal user information or are in precarious geopolitical situations and have to, in order to benefit from modern communication infrastructure, users have no choice but to hand this information over.

---

<sup>27</sup> Shift Project. “Business Model Red Flags: Red Flag #16” (2021), <https://shiftproject.org/resource/redflag-16/>.

<sup>28</sup> Rogoff, Zak. Ranking Digital Rights. “How are US telcos playing the targeted advertising game?” (2020), <https://rankingdigitalrights.org/2020/10/29/how-us-telcos-play-targeted-advertising-game/>.

<sup>29</sup> Brouillette, Amy. Ranking Digital Rights. “Key findings: Companies are improving in principle, but failing in practice” (2020), <https://rankingdigitalrights.org/index2020/key-findings>.

<sup>30</sup> Cox, Joseph. Vice Motherboard. “I Gave a Bounty Hunter \$300. Then He Located Our Phone” (2019), <https://www.vice.com/en/article/nepxbz/i-gave-a-bounty-hunter-300-dollars-located-phone-microbilt-zumigo-tmobile>.

<sup>31</sup> Fouche, Gwladys. *Reuters*. “UPDATE 3-Norway can't stop transfer of Telenor data to Myanmar rulers-minister” (2022), <https://www.reuters.com/article/myanmar-telenor-Norway-idCNL8N2UR4HI>.