



RANKING DIGITAL RIGHTS

CONSULTATION DRAFT

Human Rights Risk Scenarios: Targeted Advertising

Note: This document is the second in a series of three [documents](#) that are being shared for consultation. The purpose is to obtain expert and stakeholder feedback on the concepts, principles, and standards for company best practice that will in turn inform the development of new indicators for possible inclusion in future iterations of the Ranking Digital Rights (RDR) Corporate Accountability Index. This document should be read after the *Rationale for RDR's methodology expansion to address targeted advertising*, as it builds on concepts summarized in that overview. This document in turn should be read before the list of *Best Practices for Business and Human Rights: Targeted Advertising*, which are based on the *Risk Scenarios* outlined below. All of these documents can be downloaded from the RDR website at: <https://rankingdigitalrights.org/methodology-development/2021-revisions/#targeted-advertising>.

What are human rights risk scenarios?

This document presents 14 different human rights risk scenarios: short narratives linking company practices to violations of human rights enumerated in the Universal Declaration of Human Rights (UDHR). These scenarios are directly derived from news reports or published research, and illustrate the human rights harms related to privacy and expression that can result from targeted advertising business models and the company practices they incentivize. Mapping these scenarios enables us to identify recommended company practices that would either prevent or mitigate the risk and severity of these harms (as articulated in the *Best Practices* document), which are in turn used as the basis for developing research indicators to evaluate company disclosures of relevant policies and practices.

Each scenario relates to either the right to privacy (UDHR art. 12), the right to freedom of expression and information (UDHR art. 19), or both. Many also touch on other human rights, as further explained in the next section titled "Scope of Human Rights Harms." In some scenarios the company is the immediate cause of the violation, while in other cases the targeted advertising infrastructure facilitates human rights violations committed by other parties. In yet other scenarios, targeted advertising and the business models it fuels have a fundamental impact on the information environments of entire societies, in turn creating the enabling

conditions for human rights harms. The scenarios are therefore organized into four categories, some of which overlap and intersect with one another, as detailed in the section titled “Typology of Human Rights Harms.”

Taken as a whole, these risk scenarios highlight how the targeted advertising ecosystem can threaten human rights and provide the basis for “best practices” that would mitigate the risk and severity of these harms. *Best Practices* are presented in a companion document, to be read next. Please note that this is a draft document that will be subject to an iterative process of consultation, feedback, and revision.

Scope of Human Rights Harms

RDR’s current Index methodology focuses on identifying and mitigating individual human rights harms—specifically infringement of internet users’ freedom of expression and privacy—that directly result from using the products and services of internet, mobile, and telecommunications companies. This new methodology development workstream focused on targeted advertising builds on and expands the scope of the RDR Index in two ways: First, by considering human rights harms beyond freedom of expression and privacy; and second, by addressing group and indirect harms that affect entire groups of people or individuals who are not necessarily users of a company’s products and services.

Beyond UDHR Articles 12 and 19: RDR focuses on freedom of expression and privacy because the protection of these rights ensures the ability to exercise many other rights. If people’s expression and privacy rights are not protected and respected, they cannot use technology effectively to exercise and defend other political, religious, economic, and social rights. Indeed, as we consider how to expand the methodology to address harms associated with targeted advertising, in many of the risk scenarios below, companies’ failure to respect privacy and/or freedom of expression and information causes or contributes to the violation of other human rights, specifically: right to life, liberty and security of person (UDHR art. 3); non-discrimination (UDHR art. 7, art. 23); freedom of thought (UDHR art. 18); freedom of association (UDHR art. 20); right to take part in the government of one’s country, directly or through freely chosen representatives (UDHR art. 21).

Group and indirect harms: Company practices incentivized by targeted advertising business models can also endanger human rights indirectly, or contribute to the violation of the rights of entire communities or categories of people. Many of these harms are rooted in an initial violation of privacy that occurs in service of targeted advertising. In addition, overly broad or poorly implemented efforts to mitigate secondary harms and/or third-party harms can themselves lead to human rights harms, such as over-moderation of protected expression.

Typology of Human Rights Harms

Direct harms to users—those that directly result from using a company’s products and services—are already addressed by the RDR Index and are therefore not included here. The scenarios are organized into four categories, some of which overlap and intersect with one another. In each case, company practice has a negative impact on freedom of expression and information and/or on privacy. In many cases this initial violation further affects other human rights. For each scenario, the document specifies which rights are affected and points to the relevant article in the UDHR.

A. Secondary harms: User information collected for targeted advertising purposes enables companies and advertisers to segment audiences in a very granular manner, tailoring messages to very specific attributes including preferences, habits, or traits. As this data is shared across the targeted advertising ecosystem, this in turn enables discrimination against internet users on the basis of protected traits and even the targeting of specific individuals.

B. Group harms: Targeted advertising and the design and engineering choices it incentivizes impact societies’ information environments, creating enabling conditions for human rights harms.

C. Third-party harms: The targeted advertising infrastructure is abused by actors who are outside of the company’s control to perpetrate or encourage human rights abuses.

D. Expression: Algorithmic content curation that is optimized for user engagement has the effect of prioritizing controversial and inflammatory content, including content that is not protected under international human rights law. Overly broad content moderation can result in the removal or de-prioritization of protected expression, while the spread of harmful content can have severe consequences for human rights. Importantly, our initial definition of “the user” did not include advertisers. As a result, the existing methodology does not address harms related to the removal of paid content.

Risk scenarios

A. Secondary harms

Scenario A1: Company A places employment ads on the platform of Company B, using targeting parameters that reflect Company A's current workforce, which mostly consists of women. As a result, individuals whom Company B does not classify as women do not know about the opportunities.

References:

- [Facebook Lets Advertisers Exclude Users by Race](#)
- [Facebook Is Letting Job Advertisers Target Only Men](#)
- [Facebook Promises to Bar Advertisers From Targeting Ads by Race or Ethnicity. Again.](#)
- [How Facebook knows which apps you use, and why this matters](#)
- [Friction-free racism](#)

Human Rights Risk: Privacy (UDHR art. 12), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Advertising networks, services that display targeted ads.

Scenario A2: Company A uses an algorithm on its platform to generate “affinity groups” that advertisers can use to target specific audiences. The algorithm recognizes certain patterns in user profiles, and determines that people who express hate toward Ethnic Group A are a valuable audience for advertisers. A hate group associated with Ethnic Group B uses this affinity group to spread hate speech targeting Group A to the users who are most likely to engage with that content.

References:

- [Facebook Enabled Advertisers to Reach ‘Jew Haters’](#)
- [How the online business model encourages prejudice](#)
- [What we learned from collecting 100,000 targeted Facebook ads](#)
- [Facebook will remove 5,000 ad targeting categories to prevent discrimination](#)

Human Rights Risk: Privacy (UDHR art. 12), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Advertising networks, services that display targeted ads.

Scenario A3: A telecommunications company sells browsing and location history of its users to a data broker who then makes the data available to various other entities through a series of transactions. One such entity helps customers target specific individuals for personalized advertising and location tracking. This service is used by Person A to harass, gaslight, and stalk their former spouse.

References:

- [For \\$29, This Man Will Help Manipulate Your Loved Ones With Targeted Facebook and Browser Links](#)
- [I gave a bounty hunter \\$300. The he located our phone.](#)

Human Rights Risk: Privacy (UDHR art. 12), right to life, liberty and security of person (UDHR art. 3), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Advertising networks, services that display targeted ads.

Scenario A4: A platform uses its deep insight into its users' traits, behaviors, habits, and other characteristics to nudge users in Country A to vote in a national election, but does not do so in Country B, in order to measure how "effective" such nudges are.

References:

- [Everything we know about Facebook's secret mood manipulation experiment](#)
- [The secret experiment behind Facebook's "I Voted sticker](#)

Human Rights Risk: Privacy (UDHR art. 12), freedom of expression and information (UDHR art. 19), right to take part in the government of one's country, directly or through freely chosen representatives (UDHR art. 21).

Types of companies/services: Advertising networks, services that display targeted ads.

B. Group harms

Scenario B1: Users seeking to learn more about their country's history visit a video-sharing platform. The recommendation algorithm suggests a series of videos that display marks of high engagement (views, comments, etc.) but are inaccurate or misleading and promote hateful views of a religious minority group. Some of these users

then go on to adopt and express hateful views against members of the religious minority views, to discriminate against them in the workplace and in daily life, or even to engage in physical violence against them.

References:

- [This is exactly how social media algorithms work today](#)
- [The algorithmic rise of the “alt-right”](#)
- [Alternative Influence: Broadcasting the Reactionary Right on YouTube.](#)
- [YouTube’s Alex Jones problem](#)
- [Facebook’s failure to enforce its own rules](#)
- [Facebook posts “substantively contributed” to Myanmar genocide, UN investigators say](#)
- [Where countries are tinderboxes, and Facebook is a match](#)

Human Rights Risk: Freedom of expression and information (UDHR art. 19), right to life, liberty and security of person (UDHR art. 3), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Social networking sites, services that host user-generated content.

Scenario B2: In a country whose telecommunications providers are widely believed to share user information with the government, and where same-sex sexual activity is illegal, a young person refrains from visiting sites that provide sexual health information for LGBTQ youth.

References:

- [Internet surveillance, regulation, and chilling effects online: A comparative case study](#)
- [Whose speech is chilled by surveillance?](#)
- [Open Letter to Google: Our Concerns over “Project Dragonfly”](#)

Human Rights Risk: Privacy (UDHR art. 12), freedom of expression and information (UDHR art. 19), freedom of thought (UDHR art. 18), freedom of association (UDHR art. 20).

Types of companies/services: Services that collect and process user information.

Scenario B3: A person with low media literacy incorrectly believes that the content they see on a social media platform is vetted by professional journalists and editors. Moreover, they find it difficult to distinguish between news articles, user-generated content, and advertising. They base their decision to vote for a particular candidate on the basis of inaccurate information, including inaccurate advertising.

References:

- [Facebook and Google make lies as pretty as truth](#)
- [Blurred Lines: Advertising or Content? – An FTC Workshop on Native Advertising](#)

Human Rights Risk: Freedom of expression and information (UDHR art. 19), right to take part in the government of one's country, directly or through freely chosen representatives (UDHR art. 21).

Types of companies/services: Services that display targeted ads.

Scenario B4: Company A, which depends on advertising revenue, provides access to a subset of internet services (including its own platform and websites) to its users at no financial cost (“zero-rating”), thus incentivizing users to favor those services over competitors and ensuring that the company is better able to track users’ online activity and serve them more “relevant” ads. *De facto* limited access to a broad choice of online information sources creates enabling conditions for human rights violations.

References:

- [Free Basics in real life](#)
- [Zero rating: a global threat to the open internet](#)
- [In Sri Lanka, Facebook’s dominance has cost lives](#)
- [Facebook’s Free Basics Is an African Dictator’s Dream](#)
- [Facebook’s Free Basics Doesn’t Connect You to the Global Internet – But it Does Collect Your Data](#)

Human Rights Risk: Privacy (UDHR art. 12), freedom of expression and information (UDHR art. 19), freedom of association (UDHR art. 20).

Types of companies/services: Services that display ads alongside user-generated content, services that collect and process user information, telecommunications companies

C. Third-party harms

Scenario C1: Elites from Ethnic Group A use a platform's targeted advertising capabilities to promote disinformation and hate speech about the minority Ethnic Group B as part of an ethnic cleansing campaign. False information and incitement to violence against Group B's members spread across the platform, inspiring acts of violence against Ethnic Group B.

References:

- [What happens when the government uses Facebook as a weapon?](#)
- [How Duterte used Facebook to fuel the Philippine drug war](#)
- [Soldiers in Facebook's war on fake news are feeling overrun](#)
- [A genocide incited on Facebook, with posts from Myanmar's military](#)

Human Rights Risk: Freedom of expression and information (UDHR art. 19), right to life, liberty and security of person (UDHR art. 3), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Advertising networks, services that display ads.

Scenario C2: In the run-up to an election, voters from Ethnic Group A are shown ads with an incorrect election date, while voters from Ethnic Group B are shown ads with the correct date. As a result, members of Ethnic Group B turn out to vote at much higher rates than Group A members, and a political party that champions Ethnic Group B's interests wins the election. Once in power, this party enacts policies that discriminate against Group A.

References:

- [Voter suppression and racial targeting: In Facebook and Twitter's words](#)
- [We asked for examples of election misinformation. You delivered.](#)

Human Rights Risk: Privacy (UDHR art. 12), freedom of expression and information (UDHR art. 19), right to take part in the government of one's country, directly or through freely chosen representatives (UDHR art. 21).

Types of companies/services: Advertising networks, services that display targeted ads.

Scenario C3: A company collects vast amounts of user information, gaining insight into individual users' traits, habits, and behavior. The company provides access to this data to a commercial partner that in turn provides access to the data to a law enforcement agency, which uses the data to identify likely participants in a protest against police violence and places them under extra-legal surveillance.

References:

- [ACLU finds social media sites gave data to company tracking black protesters](#)
- [Facebook has banned developers from using data for surveillance](#)
- [Can Facebook and Twitter stop social media surveillance?](#)
- [Report: Police are now asking Google for data about all mobile devices close to certain crimes](#)

Human Rights Risk: Privacy (UDHR art. 12), freedom of expression and information (UDHR art. 19), non-discrimination (UDHR art. 7, art. 23), freedom of association (UDHR art. 20).

Types of companies/services: Services that collect user information.

Scenario C4: In a country where third-party election ads are prohibited by law, an industry lobbying group targets online ads in favor of its preferred candidate. Only the targeted audience can see the ads, and this opacity prevents the election authorities from enforcing the law. The industry lobbying group's preferred candidate wins the election and later repeals consumer protections opposed by the industry lobbying group.

References:

- [Computational Propaganda](#)
- [Challenging truth and trust: A global inventory of organized social media manipulation](#)
- [Machine politics: The rise of the internet and a new age of authoritarianism](#)
- [How to fix Facebook before it fixes us](#)
- [Weaponizing the digital influence machine: The political perils of online ad tech](#)

Human Rights Risk: Privacy (UDHR art. 12), freedom of expression and information (UDHR art. 19), right to take part in the government of one's country, directly or through freely chosen representatives (UDHR art. 21).

Types of companies/services: advertising networks, services that display ads, services that collect and process user information.

D. Expression Harms

Scenario D1: A non-profit that provides low-cost healthcare, including sexual and reproductive health care, tries to purchase ads on a popular online platform but the ads are rejected because the person or computer program carrying out content moderation incorrectly interprets the company's Community Guidelines to prohibit all references to sexual activity.

References:

- [Google bans Irish abortion referendum adverts](#)

Human Rights Risk: Freedom of expression and information (UDHR art. 19), freedom of association (UDHR art. 20), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Advertising networks, services that display targeted ads.

Scenario D2: An environmental non-profit purchases ads on a popular online platform that highlight how local industries harm native wildlife, but the ads are removed because of their "political" content.

References:

- [Facebook's new political ad policy is already sweeping up non-campaign posts](#)

Human Rights Risk: Freedom of expression and information (UDHR art. 19), freedom of association (art. 20), non-discrimination (UDHR art. 7, art. 23).

Types of companies/services: Advertising networks, services that display ads
