



## Transparency and Accountability Standards for Online Platforms

In response to the Digital Services Act (DSA) open consultation, [Ranking Digital Rights](#) (RDR) is submitting a set of materials (*provided below*) that address the DSA's focus on harmonizing responsibilities and oversight of online platforms and ISPs in the EU.<sup>1</sup>

RDR works to promote freedom of expression and privacy on the internet by creating global transparency and accountability standards for companies to respect and protect human rights. We do this by producing the [Ranking Digital Rights Corporate Accountability Index](#), which ranks the world's most powerful digital platforms and telecommunications companies on how transparent they are about policies and practices affecting fundamental freedom of expression, information, and privacy rights. The RDR Index methodology is based on international human rights standards and frameworks, including the Universal Declaration of Human Rights,<sup>2</sup> the International Covenant on Civil and Political Rights,<sup>3</sup> the European Convention on Human Rights, and the U.N. Guiding Principles on Business and Human Rights.<sup>4</sup> To date, RDR has produced four RDR Indexes (2015, 2017, 2018, and 2019) and offers the only year-on-year ranking of these platforms.

### Materials for consideration

Online platforms exercise tremendous influence on politics, societies, and the enjoyment of fundamental human rights, notably through their governance of online content. Despite measurable progress over the years, companies continue to lack transparency about the rules, processes and outcomes that shape platform users' experiences and information diets. The materials provided below outline key transparency standards that will enable citizens, regulators, and civil society organizations to hold digital platforms accountable for how they govern, moderate, and otherwise shape content. These policies and practices must protect fundamental freedom of expression, information, and privacy rights, and support human rights and democracy: transparency and the accountability it enables are necessary for ensuring that outcome.

---

<sup>1</sup> <https://ec.europa.eu/digital-single-market/en/news/consultation-digital-services-act-package>

<sup>2</sup> "Universal Declaration of Human Rights," <https://www.un.org/en/universal-declaration-human-rights/>

<sup>3</sup> "International Covenant on Civil and Political Rights," *UN Human Rights Office of the High Commissioner*, <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>.

<sup>4</sup> "Guiding Principles on Business and Human Rights," *UN Human Rights Office of the High Commissioner*, [https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR\\_EN.pdf](https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf).

- The [2020 RDR Corporate Accountability Index methodology](#), which includes new indicators on algorithmic systems and targeted advertising. The methodology reflects more than a year of stakeholder engagement, research, and pilot testing to develop global accountability and transparency standards for how tech companies can demonstrate respect for human rights as they develop and deploy these new technologies. RDR is currently applying the methodology to 26 companies, and the results will be published in February 2021. The 2019 RDR Index is available at <https://rankingdigitalrights.org/index2019/>.
- [“Transparency and accountability standards for targeted advertising and algorithmic systems: Pilot study and lessons learned,”](#) *Ranking Digital Rights*, March 2020. This study presents findings from evaluating a select group of US platforms and European telecommunications companies against draft indicators on targeted advertising and algorithmic systems. The final 2020 RDR Index methodology was refined based on findings from this report.
- **Human rights risk scenarios:** a compilation of risk scenarios describing possible violations to human rights—as enumerated in the Universal Declaration of Human Rights (UDHR)—that could occur as a direct or indirect result of companies’ targeted advertising policies and practices, and as a result of companies’ development and use of algorithmic systems, respectively:
  - [“Human rights risk scenarios: Targeted advertising,”](#) *Ranking Digital Rights* (2019): Scenarios illustrate the human rights harms related to privacy and expression that can result from targeted advertising business models and the company practices they incentivize.
  - [“Human rights risk scenarios: Algorithms, machine learning and automated decision-making,”](#) *Ranking Digital Rights* (2019): Scenarios illustrate the human rights harms related to privacy and expression that can result from companies’ use of algorithms, machine learning, and automated decision-making.
- [“It’s the Business Model: How Big Tech’s Profit Machine is Distorting the Public Sphere and Threatening Democracy,”](#) a two-part series looking at the root cause of the proliferation of disinformation and other types of harmful content online: the targeted advertising business model. While these reports address the U.S. policy context in general and discussions about removing intermediary liability safeguards for online platforms, per Section 230 of the Communications Decency Act, in particular—they are relevant to ongoing discussions about the regulation of digital platforms in the EU.

## Transparency standards: Key highlights

In the following section, we highlight (select) indicators from the [2020 RDR Corporate Accountability Index methodology](#) that provide transparency and accountability benchmarks for content moderation and curation for online platforms.

**Transparency standards for ad content and ad targeting:** The ability for advertisers or other third parties to target users with personalized content—based on their browsing behaviors, location information, and other data and characteristics that have been inferred about them<sup>5</sup>—can significantly shape (and in some cases, distort) a user’s online experience and information diet. Personalization, which can affect both paid and unpaid content, can amplify offline social inequities and can be overtly discriminatory. It can also result in so-called “filter bubbles” as well as amplify problematic content, including content intended to mislead or to spread falsehoods.

Therefore, online platforms that enable advertisers and other third parties to target their users with personalized ads or content should:

- Disclose evidence that they conduct regular, comprehensive, and credible due diligence, such as through robust human rights impact assessments, to identify how all aspects of their targeted advertising policies and practices affect users’ fundamental rights to freedom of expression and information, to privacy, and to non-discrimination, and to mitigate any risks posed by those impacts ([Indicator G4c](#)).
- Ensure ad content and ad targeting policies are easy to find and understand ([Indicators F1b](#) and [F1c](#)), and provide prior notice of changes to these terms ([Indicators F2b](#) and [F2c](#)).
- Clearly disclose ad content and ad targeting rules ([Indicators F3b](#) and [F3c](#)), including which types of targeting parameters—such as audience categories, age, location, or other characteristics—are prohibited. Companies should also disclose their processes for identifying breaches to these rules.
- Provide evidence of enforcement of ad content and ad targeting rules ([Indicator F4d](#)) by publishing data on the number and type of ads removed as a result of breaches to ad content policies and by disclosing the rule(s) violated. Companies should also publish this data at least once a year and in a structured data file.
- Clearly disclose that targeted advertising is off by default: users should only be shown personalized ads if they explicitly opt in ([Indicator P7, Element 6](#)).

---

<sup>5</sup> For more about data inference policies, see Section 6.2 of “2020 Pilot Study and Lessons Learned,” *Ranking Digital Rights*, March 16, 2020, <https://rankingdigitalrights.org/wp-content/uploads/2020/03/pilot-report-2020.pdf>.

**Transparency standards for user content moderation:** Digital platforms can and should set rules prohibiting certain content or activities, such as toxic speech or malicious behavior. However, when companies develop and enforce rules about what people can say and do on the internet—or whether they can access a service at all—they must do so in a way that is transparent and accountable in order to ensure that freedom of expression and information rights are being respected.

Online platforms therefore should:

- Disclose evidence of conducting regular, comprehensive, and credible due diligence, such as through robust human rights impact assessments, to identify how their processes for policy enforcement affect users' fundamental rights to freedom of expression and information, to privacy, and to non-discrimination, and to mitigate any risks posed by those impacts ([Indicator G4b](#)).
- Ensure terms of service are easy to find and understand ([Indicator F1a](#)), and provide prior notice of changes to these terms ([Indicator F2a](#)).
- Clearly disclose policies describing what types of content and activities are not permitted on their platforms and services and how they enforce these rules ([Indicator F3a](#)). Companies should also disclose their processes for identifying breaches to targeting rules.
- Provide evidence of enforcement of their terms of service ([Indicators F4a and Fb](#)) by publishing data on content and accounts removed as a result of breaches to platform rules, and which rule(s) violated. Companies should also publish this data regularly and in a structured data format.

**Transparency regarding the development and use of algorithmic systems for content moderation and governance:** Using algorithmic systems to moderate and govern the dissemination of user content can have adverse effects on fundamental human rights, specifically, the rights to free expression, access to information, privacy, and non-discrimination.<sup>6</sup> Algorithmic content curation, recommendation, and ranking systems play a critical role in shaping what types of content and information users can see and access online. In addition, systems that are optimized for user engagement can have the effect of prioritizing controversial and inflammatory content, including content that is not protected under international human rights law. Over time, reliance on algorithmic curation and recommendation systems that are optimized for engagement can alter the news and information ecosystems of entire communities or countries. These systems can be manipulated to spread disinformation and otherwise distort

---

<sup>6</sup> “Consultation draft: Human Rights Risk Scenarios: Algorithms, Machine Learning and Automated Decision-Making,” *Ranking Digital Rights* (2019), [https://rankingdigitalrights.org/wp-content/uploads/2019/07/Human-Rights-Risk-Scenarios\\_-\\_algorithms-machine-learning-automated-decision-making.pdf](https://rankingdigitalrights.org/wp-content/uploads/2019/07/Human-Rights-Risk-Scenarios_-_algorithms-machine-learning-automated-decision-making.pdf).

the information ecosystem, which can in turn fuel human rights abuses. The development and testing of algorithmic systems can also pose significant risks to privacy, particularly when companies then use the information collected about users to develop, train, and test these systems without the data subject's informed consent.<sup>7</sup>

Online platforms that develop and deploy algorithms therefore should:

- Disclose a clear commitment to uphold international human rights standards in their development and deployment of algorithmic systems ([Indicator G1, Element 3](#)), in line with the Council of Europe's [Recommendation CM/Rec\(2020\)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems](#).
- Disclose evidence that they conduct regular, comprehensive, and credible due diligence, such as through robust human rights impact assessments, to identify how all aspects of its policies and practices related to the development and use of algorithmic systems affect users' fundamental rights to freedom of expression and information, to privacy, and to non-discrimination, and to mitigate any risks posed by those impacts ([Indicator G4d](#)).
- Publish policies that clearly describe the terms for how they use algorithmic systems across their services and platforms ([Indicator F1d](#)). Companies that use algorithmic systems with the potential to cause human rights harms should publish a clear and accessible policy stating the nature and functions of these systems.<sup>8</sup> This policy should be easy to find, presented in plain language, and contain options for users to manage settings.
- Publish information about whether they use algorithmic systems to curate, recommend, and rank content ([Indicator F12](#)). They should disclose how these systems work, what options users have to control how their information is used by these systems, and whether such systems are automatically on by default or users can opt-in to have their content automatically curated by the algorithmic system.

---

<sup>7</sup> Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY, USA: PublicAffairs; Nathalie Maréchal. Targeted Advertising Is Ruining the Internet and Breaking the World, [https://www.vice.com/en\\_us/article/xwjden/targeted-advertising-is-ruining-the-internet-and-breaking-the-world](https://www.vice.com/en_us/article/xwjden/targeted-advertising-is-ruining-the-internet-and-breaking-the-world), *Vice Motherboard*, November 16, 2018; "Human Rights Risk Scenarios: Algorithms, machine learning and automated decision-making," *Ranking Digital Rights*, July 2019, [https://rankingdigitalrights.org/wp-content/uploads/2019/07/Human-Rights-Risk-Scenarios\\_-algorithms-machine-learning-automated-decision-making.pdf](https://rankingdigitalrights.org/wp-content/uploads/2019/07/Human-Rights-Risk-Scenarios_-algorithms-machine-learning-automated-decision-making.pdf).

<sup>8</sup> "Addressing the impacts of Algorithms on Human Rights: Draft Recommendation of the Committee of Ministers to member States on the human rights impacts of algorithmic systems," Council of Europe, Committee of experts on human rights dimensions of automated data processing and different forms of artificial intelligence (2019), <https://rm.coe.int/draft-recommendation-of-the-committee-of-ministers-to-states-on-the-hu/168095eefc>.

- Clearly disclose algorithmic system development policies in a way that users can easily access and understand, so that users can make informed decisions about whether to use a company's products and services ([Indicator P1b](#)).
- Clearly disclose that they provide users with options to control how their data is used for the development of algorithmic systems ([Indicator P7, Element 7](#)).
- Clearly disclose whether they use user data to develop algorithmic systems by default, or if users must affirmatively consent to such use of their data ([Indicator P7, Element 8](#)).

### For more information

- Find out more about RDR and the RDR Corporate Accountability Index: [www.rankingdigitalrights.org](http://www.rankingdigitalrights.org).
- View and download the 2020 RDR Index methodology: <https://rankingdigitalrights.org/2020-indicators>.
- Read a summary of our research process: <https://rankingdigitalrights.org/2020-research-process>.
- Learn which companies we are ranking in the 2020 RDR Index, to be released in February 2021: <https://rankingdigitalrights.org/2020-companies>.
- Read about how we develop the RDR Index methodology: <https://rankingdigitalrights.org/methodology-development/>.

Thank you for your consideration of these comments and the attached resources. We look forward to the opportunity to further discuss these resources and recommendations at your convenience.

Best regards,

**Jessica Dheere**, Director, Ranking Digital Rights, [dheere@rankingdigitalrights.org](mailto:dheere@rankingdigitalrights.org)

**Amy Brouillette**, Research Director, Ranking Digital Rights, [brouillette@rankingdigitalrights.org](mailto:brouillette@rankingdigitalrights.org)

**Nathalie Maréchal**, Senior Policy Analyst, Ranking Digital Rights, [marechal@rankingdigitalrights.org](mailto:marechal@rankingdigitalrights.org)