

---

# Human Rights for Small and Medium Sized Technology Companies: Freedom of Expression and Content Moderation

---

June 2020

---

For platforms that host user-generated content, it is important to develop detailed, digestible, and transparent content policies and procedures. We recognize that smaller companies are often constrained in terms of resources and personnel. Below are eight important points platforms should consider when developing and implementing user content moderation policies. They have been broken down according to whether they are essential for any company to have or whether they are expected of advanced, larger companies.

---

## Essential

- 1. Develop a clear set of content policies:** These content policies should outline what categories of content are permitted on your platform, and what categories of content are prohibited. Additionally, you should develop clear policies for responding to external requests. In order to ensure that your policies and processes comply with any specific national or regional regulations related to online content governance, you should consult with legal experts. All of your content policies should be publicly available, easily accessible on your platform, and should be easy to understand.
- 2. Explain how your content policies will be enforced:** All of your users should have access to information which explains how your company enforces its content policies. These explanations should, for example, outline in which scenarios, if any, you remove or geo-block content, and in which scenarios you temporarily suspend or permanently delete accounts on your service for violating your content policies. It should also include to what extent automated processes are used to filter, flag, or remove content. This information should be publicly available online and should be easy to understand. Where possible, use case studies or examples to illustrate how different types of content or accounts may be permitted or prohibited on your service.
- 3. Give impacted users appropriate notice:** You should establish a procedure for providing adequate notice (when legal) to any impacted users when you decide to remove or restrict a user's content, or suspend or delete their

---

accounts. This notice should clearly explain which piece of content violated your rules (including by providing a URL to the content or an excerpt), the specific policy the content or account violated, how the content was detected and removed (e.g. flagged by an external party, detected using automated tools), and whether or not the user can appeal the decision. Notices should be available in a durable form that is accessible even if a user has their account suspended or deleted. If a user flags content for violating your policies, this user should be able to monitor whether or not the content they reported has been removed, preferably through a log of activity or content history. These users should also have the opportunity to appeal moderation decisions.

4. **Establish a meaningful appeals process:** In order to provide transparency and accountability to your users, you should establish a thorough and timely appeals process for content moderation and account suspension or deletion cases. An appeals process should be available to users who have had their content removed or restricted or their account suspended or deleted, as well as users who have flagged content or accounts for violating the platform's policies. During the appeals process, the content or account in question should be reviewed by a person or group of individuals who were not involved in the initial moderation decision. The user submitting the appeal should also have the opportunity to present additional information to substantiate their appeals case. After the appeals process has concluded, the user should receive a clear and detailed notification explaining the results of the appeals process and how this decision was made.
5. **Publish a transparency report outlining the scope and scale of your content moderation efforts:** Your transparency report should detail the number of posts you removed and restricted and the number of accounts you suspended or removed for violating your content policies. This data should be broken down into different categories, including by product, the content policy violated, and the detection method, such as automated tools, reports from a user, or reports from governments. In addition, your transparency report should include data on the volume and outcome of appeals you have received. There are a

---

number of guidelines on how to develop granular and meaningful transparency reports, including the [Santa Clara Principles on Transparency and Accountability in Content Moderation](#),<sup>1</sup> the Ranking Digital Rights [Corporate Accountability Index](#),<sup>2</sup> and New America's [Transparency Reporting Toolkit on Content Takedowns](#).<sup>3</sup> All transparency reports should be published on a regular schedule, in an openly licensed, machine-readable format.

- 6. Consult with stakeholders:** In order to ensure that your content policies and procedures are transparent, rights-protective, and responsive to the needs and rights of your diverse user base, your company should conduct regular consultations with representatives from different communities, sectors, and organizations. In particular, these consultations should be held when you are considering introducing a new content policy or procedure, or making changes to your existing content policies and procedures.

---

## Advanced

- 7. Outline how your company responds to external requests for content moderation:** If you are a company that hosts user-generated content, it is likely that you will receive external requests from government agencies, users, and other parties to remove or restrict content from your services. This is a procedure that significantly impacts the free expression rights of your users, and it is therefore important that you develop a comprehensive and rights-respecting process for handling such requests. You should establish a process for tracking requests and their status, a process for reviewing and classifying requests based on their legal validity, and procedures to ensure that data about these requests is kept secure. In addition, you should establish clear parameters around what types of requests will be accepted and which will be rejected, as well as when impacted users will be notified. The Global Network Initiative Principles on Freedom of Expression and Privacy, together with their more detailed Implementation Guidelines offer some further guidance on how to respond to government requests. Further, individuals responsible for enforcement should be equipped and empowered to reject any illegitimate or disproportionate requests.

---

**8. Provide updates on changes in your content policies:**

As the internet and the landscape of online speech changes, so should your content policies. To ensure that your users are aware of any changes to your content policies, and subsequently can understand how this impacts their use of your service, you should always provide updates when your content policies undergo significant changes. For example, you can do this through blog posts or in-app notifications. In addition, you should create an archive of your content policies, so that users and researchers can track how your content policies have changed over time.

---

**Endnotes**

1. <https://santaclaraprinciples.org/>
2. <https://rankingdigitalrights.org/index2019/>
3. <https://www.newamerica.org/oti/reports/transparency-reporting-toolkit-content-takedown-reporting/>

Second Home  
68 Hanbury St  
London E1 5JL

+44 203 818 3258