



Who Has Your Back?

CENSORSHIP EDITION 2019



Who Has Your Back? Censorship Edition 2019

Authors: Andrew Crocker, Gennie Gebhart, Aaron Mackey, Kurt Opsahl, Hayley Tsukayama, Jamie Lee Williams, Jillian C. York

With assistance from: Caitlyn Crites, Hugh D’Andrade, Lena Gunn, Jason Kelley, Laura Schatzkin

A publication of the Electronic Frontier Foundation, 2019.

“Who Has Your Back? 2019” is released under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

View this report online: <https://www.eff.org/who-has-your-back-2019>



Who Has Your Back?

CENSORSHIP EDITION 2019

Andrew Crocker, Senior Staff Attorney
Gennie Gebhart, Associate Director of Research
Aaron Mackey, Staff Attorney
Kurt Opsahl, Deputy Executive Director and General Counsel
Hayley Tsukayama, Legislative Activist
Jamie Lee Williams, Staff Attorney
Jillian C. York, Director for International Freedom of Expression

JUNE 2019

| | Legal Requests | Platform Policy Requests | Notice | Appeals Mechanisms | Appeals Transparency | Santa Clara Principles |
|-------------------|----------------|--------------------------|--------|--------------------|----------------------|------------------------|
| Apple App Store | ★ | ★ | ★ | ★ | ★ | ★ |
| Dailymotion | ★ | ★ | ★ | ★ | ★ | ★ |
| Facebook | ★ | ★ | ★ | ★ | ★ | ★ |
| GitHub | ★ | ★ | ★ | ★ | ★ | ★ |
| Google Play Store | ★ | ★ | ★ | ★ | ★ | ★ |
| Instagram | ★ | ★ | ★ | ★ | ★ | ★ |
| LinkedIn | ★ | ★ | ★ | ★ | ★ | ★ |
| Medium | ★ | ★ | ★ | ★ | ★ | ★ |
| Pinterest | ★ | ★ | ★ | ★ | ★ | ★ |
| Reddit | ★ | ★ | ★ | ★ | ★ | ★ |
| Snap | ★ | ★ | ★ | ★ | ★ | ★ |
| Tumblr | ★ | ★ | ★ | ★ | ★ | ★ |
| Twitter | ★ | ★ | ★ | ★ | ★ | ★ |
| Vimeo | ★ | ★ | ★ | ★ | ★ | ★ |
| Wordpress.com | ★ | ★ | ★ | ★ | ★ | ★ |
| YouTube | ★ | ★ | ★ | ★ | ★ | ★ |

Contents

| | |
|---|----|
| Executive Summary | 6 |
| Introduction | 7 |
| Scope | 8 |
| Major Findings and Trends | 9 |
| Overview of Criteria | 10 |
| Transparent About Legal Takedown Requests | 11 |
| Transparent About Platform Policy Takedown Requests | 11 |
| Provides Meaningful Notice | 12 |
| Appeals Mechanisms | 12 |
| Appeals Transparency | 13 |
| Santa Clara Principles | 13 |
| Company Reports | 14 |
| Apple App Store | 14 |
| Dailymotion | 16 |
| Facebook | 17 |
| GitHub | 19 |
| Google Play Store | 21 |
| Instagram | 23 |
| LinkedIn | 25 |
| Medium | 27 |
| Pinterest | 29 |
| Reddit | 30 |
| Snap | 32 |
| Tumblr | 33 |
| Twitter | 35 |
| Vimeo | 37 |
| WordPress.com | 38 |
| YouTube | 39 |

Executive Summary

Over the past year, governments have made unprecedented demands for online platforms to police speech, and many companies are rushing to comply. But in their response to calls to remove objectionable content, social media companies and platforms have all too often censored valuable speech. While it is reasonable for companies to moderate some content, no one wins when companies and governments can censor legitimate online speech without transparency, notice, or due process.

This year's Who Has Your Back report examines major tech companies' content moderation policies in the midst of massive government pressure to censor. We assess companies' policies in six categories:

- Transparency in reporting government takedown requests based on **legal requests**
- Transparency in reporting government takedown requests alleging **platform policy violations**
- Providing **meaningful notice** to users of every content takedown and account suspension
- Providing users with an **appeals process** to dispute takedowns and suspensions
- Transparency regarding **the number of appeals**
- Public support of the **Santa Clara Principles**

These categories build on last year's first-ever censorship edition of Who Has Your Back¹ in an effort to foster improved content moderation best practices across the industry. Even with stricter criteria, we are pleased to see several companies improving from last year to this year.

Only one company—**Reddit**—earned stars in all six of these categories. And two companies—**Apple** and **GitHub**—earned stars in five out of six categories, both falling short only on appeals transparency. We are pleased to report that, of the 16 companies we assess, twelve publicly endorse the Santa Clara Principles on Transparency and Accountability in Content Moderation,² indicating increasing industry buy-in to these important standards.

Some content moderation best practices are seeing wider adoption than others. Although providers increasingly offer users the ability to appeal content moderation decisions, they do not as consistently provide users with clear notice and transparency regarding their appeals processes. According to the policies of several providers, users have the ability to appeal all content removals, but they may not receive notification that their content has been removed in the first place. This creates a critical gap in information and context for users trying to navigate takedown and suspension decisions—and for advocates striving to better understand opaque content moderation processes. We will continue encouraging more consistent adoption of the best practices identified in this report, and closing such critical information gaps, moving forward.

¹ Nate Cardozo, Andrew Crocker, Gennie Gebhart, Jennifer Lynch, Kurt Opsahl, and Jillian C. York, "Who Has Your Back? Censorship Edition 2018," *Electronic Frontier Foundation*, <https://www.eff.org/who-has-your-back-2018>.

² The Santa Clara Principles on Transparency and Accountability in Content Moderation, <https://santaclaraprinciples.org/>.

Introduction

In the aftermath of horrific violence in New Zealand and Sri Lanka and viral disinformation campaigns about everything from vaccines to elections, governments have made unprecedented demands for online platforms to police speech. And companies are rushing to comply. Facebook CEO Mark Zuckerberg even published an op-ed³ imploring governments for more regulation “governing the distribution of harmful content.”

But in their response to calls to remove objectionable content, social media companies and platforms have all too often censored valuable speech. Marginalized groups are particularly impacted by this increased content policing, which impairs their ability to use social media to organize, call attention to oppression, and even communicate with loved ones during emergencies. And the processes used by tech companies to moderate content are often tremendously opaque. While it is reasonable for companies to moderate some content, no one wins when companies and governments can censor legitimate online speech without transparency, notice, or due process.

This year’s Who Has Your Back report assesses company policies in these areas in the midst of significant government pressure to censor. Along with increased government action to mandate certain kinds of content moderation, some companies reported an uptick in the number of government requests for platforms to take down content based on claims of legal violations. At Twitter, for example, such requests increased 84 percent and affected more than twice as many accounts from 2017 to 2018.⁴

After the attacks in Christchurch—which left 51 people dead, injured more than 40 others, and were livestreamed on Facebook—New Zealand released the Christchurch Call, a plan to combat terrorism and violent extremism online. While the plan has valuable components addressing the need for governments to deal with the root causes of extremism, it also asks governments to consider developing industry standards and regulations, and asks companies to employ upload filters to detect and block extremist content.

With other freedom of expression advocates, we at the Electronic Frontier Foundation have raised concerns⁵ that the plan could lead to blunt measures that undermine free speech. Nonetheless, eighteen countries, as well as providers such as Google, Facebook, Twitter, and Microsoft, signed on to the call. The United States declined, citing First Amendment concerns.

Other countries took action before the Christchurch Call was unveiled. Australia passed legislation that would penalize companies for failing to quickly remove videos containing “abhorrent violent content” from social media platforms, with fines as high as 10% of annual

³ Mark Zuckerberg, “The Internet needs new rules. Let’s start in these four areas,” *The Washington Post*, 30 March 2019, <https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/>.

⁴ Twitter transparency report: removal requests, <https://transparency.twitter.com/en/removal-requests.html>.

⁵ Jillian C. York, “The Christchurch Call: The Good, the Not-So-Good, and the Ugly,” *Electronic Frontier Foundation Deeplinks*, 16 May 2019, <https://www.eff.org/deeplinks/2019/05/christchurch-call-good-not-so-good-and-ugly>.

revenue and potential jail time for executives.⁶ European Union lawmakers approved a plan requiring platforms to remove terrorist content within one hour of being notified about it by authorities.⁷ The United Kingdom proposed the creation of a regulatory body to enforce rules against online misinformation, hate speech, and cyberbullying.⁸ And in 2017, Germany passed the “Network Enforcement Law,” which has already led to the deletion of legitimate expressions of opinion.⁹

In authorizing new regulations that carry potential punishments in the billions of dollars and even possible jail time, governments seem to be sending a clear message to platforms: police your users, or else. This could easily inspire platforms—which already make too many unacceptable content moderation mistakes—to over-censor and effectively silence people for whom the Internet is an irreplaceable forum to express ideas, connect with others, and find support.

Scope

This report provides objective measurements for analyzing the content moderation policies of major technology companies. We focus on a handful of specific, measurable criteria that reflect attainable best practices.

We assess those criteria for 16 of the biggest online platforms that publicly host a large amount of user-generated content. The group of companies and platforms we assess does not include infrastructure providers (e.g. Cloudflare), file hosting services (e.g. Dropbox, Google Drive), communications providers (e.g., Gmail, Outlook), or search engines (e.g. Bing, Google).

The scope of this report does not include several types of censorship. We do not cover removals of child exploitation imagery, or intellectual property removals, restrictions, and reporting. Further, for the two “app stores” we evaluate this year, we limit the scope of our review to developer accounts and the apps themselves.

As tech companies face more pressure to take down content, the line between government censorship and platform censorship is increasingly hard to draw. With this in mind, this report does not just assess companies’ reporting and handling of explicit government takedown requests. We also look more comprehensively at whether notice and appeals processes apply to *all* content takedowns and account suspensions, regardless of whether they are driven by government pressure, company content rules, or some combination of the above.

⁶ Paul Karp, “Australia passes social media law penalising platforms for violent content,” *The Guardian*, 3 April 2019, <https://www.theguardian.com/media/2019/apr/04/australia-passes-social-media-law-penalising-platforms-for-violent-content>.

⁷ Colin Lecher, “Aggressive new terrorist content regulation passes EU vote,” *The Verge*, 17 April 2019, <https://www.theverge.com/2019/4/17/18412278/eu-terrorist-content-law-parliament-takedown>.

⁸ Chris Fox, “Websites to be fined over ‘online harms’ under new proposals,” *BBC News*, 8 April 2019, <https://www.bbc.com/news/technology-47826946>.

⁹ Bernhard Rohleder, “Germany set out to delete hate speech online. Instead, it made things worse,” *The Washington Post*, 20 February 2018, https://www.washingtonpost.com/news/theworldpost/wp/2018/02/20/netzdg/?utm_term=.5fef3daf9328

Major Findings and Trends

Our major findings include:

- Only one company—**Reddit**—received credit in all six categories.
- Two companies—the **Apple App Store** and **GitHub**—received credit in five out of six categories, falling short only on appeals transparency.
- Of the 16 companies we assess, twelve publicly endorse the **Santa Clara Principles**, indicating increasing industry buy-in to these important standards.
- Although providers are increasingly offering users the ability to appeal content moderation decisions, **notice policies and appeals transparency are not keeping up.**

We are pleased to announce that **Reddit** earned stars in every category we evaluated in this year's report, and the **Apple App Store** and **GitHub** earned stars for all but one of this year's categories. These companies stand out as examples of strong content moderation policies across the board.

Notably, two of this year's highest-scoring companies have unique models of user-generated content and communities. Reddit's sub-reddit model relies on moderators to create and enforce community norms, setting the context for the site-wide corporate policies this report assesses. And GitHub is one of the largest sites hosting user-generated code and the communities that form around it. Both companies have managed to meet industry best practices while also adapting those best practices to the community models and types of content they host. Both Reddit and GitHub also employ small policy and content moderation teams relative to some of the other companies we assess this year. If they can achieve this outstanding level of transparency and accountability with regard to content takedowns and account suspensions, it's reasonable to expect other companies to also meet that standard.

Even with stricter criteria this year, several companies improved their scores from the 2018 report to the 2019 report. In particular, **Facebook** and **Reddit** both committed to more comprehensive notice policies. In another example of improvement, **Pinterest** made its transparency reporting more detailed. It provides a good example for other companies, like Twitter and LinkedIn, that do provide data on government takedown requests, but not at the level of detail this report requires. **Snap** has also made notable improvements. Despite the fact that its primarily ephemeral content complicates the logistics of implementing some processes like appeals, it has improved its transparency reporting and indicated a commitment to content moderation best practices going forward.

We are also pleased to see over half the companies in this year's report publicly supporting the Santa Clara Principles on Transparency and Accountability in Content Moderation. These principles outline a set of minimum content moderation policy standards in three areas: transparency, notice, and appeals. This year's Who Has Your Back criteria roughly mirror these

areas, and reflect continued corporate progress toward fulfilling the spirit of the principles. The principles were released by EFF in conjunction with the ACLU of Northern California, Center for Democracy & Technology, New America's Open Technology Institute, and a group of academic experts and advocates. Now, one year later, we are glad to see industry actors joining civil society and academia in indicating their support for the principles moving forward.

Out of this year's six star categories, notice posed one of the biggest challenges, with only four of the companies assessed in this year's report meeting our criteria for meaningful notice. Executing a comprehensive notice policy requires a large, ongoing, potentially unpredictable commitment of resources. However, notice remains a critical component of accountable content moderation. Meaningful notice can build trust between users and platforms by ensuring that users affirmatively know when and why a platform has removed their content. Without timely, informative, on-the-record notice, users are left grasping for evidence with which to appeal to companies directly. Further, notice of legal takedown requests in particular can give users the information they need to draw public attention to government targeting and censorship.

Many companies in this year's report do provide another critical component of accountable content moderation: an appeals process. While we applaud the eleven companies that allow users to appeal content takedowns and account suspensions, notice is still a missing piece for many. This leads to a challenge for users: how can you utilize an appeals process if you are not notified that there is something to appeal? Without the knowledge that an enforcement decision has been made and why, users are not in a strong position to take advantage of mechanisms to challenge it.

Another missing piece of the puzzle is transparency about appeals. Only one company—Reddit, this year's sole all-star—provided transparency about the total number of appeals it had received. Further, Reddit published the aggregate outcomes of appeals, reporting the percentage of appeals resolved in favor of or against the appeal. Without this kind of information, it is impossible to begin to understand the context of an individual appeal, or to interpret how a company deals with appeals more broadly. Of course, there are no ideal numbers here. Receiving zero total appeals does not necessarily indicate a company employs perfect content moderation, and a large number of appeals does not mean it is making mistakes. Taking steps to provide transparency around appeals can, however, give the public a window into what is currently an opaque process at most companies.

Transparency, notice, appeals, and other areas of content moderation policy and ethics are complex, and cannot fit neatly into just six criteria. This report strives to build on last year's initial censorship edition of Who Has You Back as we continue to push companies toward wider adoption of content moderation best practices.

Overview of Criteria

Only publicly available statements can qualify for credit in this report. Positions, practices, or policies that are conveyed privately or in internal corporate standards, regardless of how laudable, are not factored into our decisions to award companies credit in any category.

Requiring public documentation serves several purposes. First, it ensures that companies cannot quietly change an internal practice in the future in response to government pressure, but must also change their publicly posted policies—which observers can note and document. Second, by asking companies to put their positions in writing, we can examine each policy closely and prompt a larger public conversation about what standards tech companies should strive for. Third, it helps companies review one another’s policies around content moderation, which can serve as a guide for startups and others looking for examples of best practices.

In this report, we strive to offer ambitious but practical standards. To that end, we only include criteria that at least one major company has already adopted. This ensures that we are highlighting existing and achievable, rather than theoretical, best practices.

We analyzed six criteria for this report:

Transparent About Legal Takedown Requests

To earn a star in this category, the service provider must regularly publish records of government requests for takedowns based on claims of legal violations, for instance, in its transparency report. This should include, at a minimum, the information necessary to determine:

- the number of requests received,
- the country from which the request originated,
- the number of requests acted upon and/or the number of posts removed or restricted or accounts suspended, and
- for service providers reporting on multiple products/platforms, the product/platform associated with the requested content or account.

Takedown requests include requests to restrict public access to a post, including geographic limitation, and account suspensions that limit access to posts for a period of time.

Reporting must distinguish legal takedown requests from platform policy takedown requests.

A request is categorized as a “government request” if it is provided through official channels (such as an order issued by a competent judicial authority); if the requestor identifies themselves as a government official or relies upon their governmental position or authority; or if the provider otherwise is aware a government is being represented in the request.

If the service provider is restricted by applicable law from disclosing the request, it may delay including the request until that restriction is lifted and still get credit in this category.

Transparent About Platform Policy Takedown Requests

To earn a star in this category, the service provider must regularly publish records of content or account restrictions based upon identifiable government allegations of violations of the provider’s policies, such as Terms of Service or Community Standards, regardless of whether

the request came through channels for government requests or through customer service channels. This includes government requests alleging facts that lead to a content or account restriction based on a provider's policies.

The provider's reporting should include, at a minimum, the information necessary to determine:

- the number of requests received,
- the country from which the request originated, and
- the number of requests acted upon and/or the number of posts removed or restricted or the number of accounts suspended, and
- for service providers reporting on multiple products/platforms, the product/platform associated with the requested content or account.

Reporting must distinguish legal takedown requests from platform policy takedown requests.

A request is identifiable from a government if it is provided through official channels (such as an order issued by a competent judicial authority); if the requestor identifies themselves as a government official or relies upon their governmental position or authority; or if the provider otherwise is aware a government is being represented in the request.

Provides Meaningful Notice

To earn a star in this category, the service provider must publicly commit to provide meaningful notice to users of every removal and suspension, unless prohibited by law, in very narrow and defined emergency situations,¹⁰ or if doing so would be futile or ineffective.¹¹

For legal takedowns and suspensions, the notice must (1) identify the specific content that allegedly violates the law, and (2) inform the user that it was a legal takedown request. If the takedown is a "geoblock"—that is, a content restriction limited to the jurisdiction where the provider is legally required to restrict it—then the users must also be notified of the geographic scope of the takedown.

For policy takedowns and suspensions, this notice must (1) identify the specific content that allegedly violates a provider policy, and (2) include the specific provider policy the content allegedly violates.

Appeals Mechanisms

To earn a star in this category, the service provider must publicly commit to provide users with an appeals process and be transparent about the results of that process.

¹⁰The exceptions should not be broader than the emergency exceptions provided in the Electronic Communications Privacy Act, 18 USC § 2702 (b)(8): "if the provider, in good faith, believes that an emergency involving danger of death or serious physical injury to any person requires disclosure without delay of communications relating to the emergency[.]"

¹¹An example of a futile scenario would be if a user's account has been compromised or their mobile device stolen, and informing the "user" would concurrently—or only—inform the attacker.

This appeals process must provide users with effective mechanisms to appeal all provider-policy based content and account restriction decisions, including *during* temporary suspensions. Upon a successful appeal, the account or material must be reinstated promptly.

Appeals Transparency

The provider must also regularly publish records of appeals and their aggregate outcomes, for instance in a transparency report. This should include, at a minimum, the information necessary to determine the total number of appeals filed.

Santa Clara Principles

To earn a star in this category, the service provider must publicly support the Santa Clara Principles. This does not require the service provider to meet the principles, but rather to indicate their endorsement of them. While the previous four criteria reflect a service provider's current implementation, this one indicates their commitment to support content moderation best practices moving forward.

(All criteria exclude spam, phishing, and child exploitation imagery.)

Company Reports

Apple App Store



Transparent About Legal Takedown Requests. Apple has publicly committed to reporting government takedowns in its future transparency reports:

Starting with the Transparency Report period July 1 - December 31, 2018, Apple will report on Government requests to take down Apps from the App Store in instances related to alleged violations of legal and/or policy provisions. Apple's Transparency Report will include two new tables: "Worldwide Government App Store Takedown Requests - Legal Violations" and "Worldwide Government App Store Takedown Requests - Platform Policy Violations."

Transparent About Platform Policy Takedown Requests. Apple has publicly committed to reporting government takedowns in its future transparency reports:

Starting with the Transparency Report period July 1 - December 31, 2018, Apple will report on Government requests to take down Apps from the App Store in instances related to alleged violations of legal and/or policy provisions. Apple's Transparency Report will include two new tables: "Worldwide Government App Store Takedown Requests - Legal Violations" and "Worldwide Government App Store Takedown Requests - Platform Policy Violations."

Provides Meaningful Notice. The Apple App Store publicly commits to notifying users of every app removal with the reason for removal and the scope of geoblocking, if applicable:

Apple sometimes receives notices that require us to remove content on the App Store. We may also remove content for the reasons set forth in the App Review Guidelines or any of our agreements with you. Apple will notify you when, where, and why an app is removed from sale, with the exception of situations in which notification would be futile or ineffective, could cause potential danger of serious physical injury, could compromise Apple's ability to detect developer violations, or in instances related to violations for spam, phishing, and child exploitation imagery.

For account suspensions, Apple does not lock or disable accounts except in relation to security issues.

Allows Appeals. The Apple App Store allows users to appeal app removals, as well as app rejections.

Transparent About Appeals. While Apple has publicly committed to begin transparency reporting on appeals in the future, it has committed to reporting only on appeals received with regard to government-requested takedowns:

In addition to reporting on Government requests to take down Apps from the App Store in instances related to alleged violations of legal and/or policy provisions, starting with the Transparency Report period July 1 - December 31, 2019, Apple will report on appeals we receive pursuant to such Government requests.

Santa Clara Principles. Apple publicly supports the Santa Clara Principles:

Apple supports the spirit of the Santa Clara Principles as a starting point for further conversation about content moderation in general, and refinements that will benefit both users and platforms.

References and useful links:

Transparency report:

<https://www.apple.com/legal/transparency/>

App store information for developers:

<https://developer.apple.com/support/app-store/>

If your Apple ID is locked or disabled:

<https://support.apple.com/en-ca/HT204106>

App review information:

<https://developer.apple.com/app-store/review/>

Dailymotion



Transparent About Legal Takedown Requests. Dailymotion does not publish a transparency report.

Transparent About Platform Policy Takedown Requests. Dailymotion does not publish a transparency report.

Provides Meaningful Notice. Dailymotion does not publicly commit to providing meaningful notice to users of every removal and suspension.

Allows Appeals. Dailymotion does not have a published policy or process for users to appeal takedowns and suspensions.

Transparent About Appeals. Dailymotion does not report the number or results of appeals.

Santa Clara Principles. Dailymotion has not publicly supported the Santa Clara Principles.

References and useful links:

Terms of Use:

<https://www.dailymotion.com/legal/termsofsales>

Facebook



Transparent About Legal Takedown Requests. While Facebook produces a transparency report on government legal takedown requests that breaks requests down by country and reports the number of pieces of content removed, it does not report the total number of government takedown requests received.

Transparent About Platform Policy Takedown Requests. While Facebook produces a Community Standards enforcement report that details enforcement actions, it does not specify government requests in this category and only details enforcement on 9 categories of standards violations: Adult nudity and sexual activity, bullying and harassment, child nudity and sexual exploitation of children, fake accounts, hate speech, regulated goods (drugs and firearms), spam, terrorist propaganda, and violence and graphic content.

Provides Meaningful Notice. Facebook commits to providing users notice for both legal takedowns and Community Guidelines violations.

For legal takedowns:

We provide notice to people when we restrict something they posted based on a report of an alleged violation of local law, and we also tell people when they try view something that is restricted in their country. We provide this notice except where legally prohibited or when technical constraints prevent us from doing so.

For Community Guidelines violations:

Let's say someone publishes a post which we decide to remove from Facebook for going against our Community Standards. The person who posted it is notified, and given the option to request a review or accept the decision.

And:

When we take action on a piece of content, we notify the person who posted it and offer them the ability to tell us if they think we made a mistake.

Allows Appeals. While Facebook allows users to appeal takedowns, it does not commit to allowing appeals for all Community Standards violation types:

Today, we offer appeals for the vast majority of violation types. We don't offer appeals for violations with extreme safety concerns, such as child exploitation imagery.

Because Facebook's policy does not further explain the full scope of "extreme safety concerns" that fall outside the scope of "the vast majority of violation types," it does not earn a star in this category.

Transparent About Appeals. While Facebook reports the number of appeals and rate of content restoration in its Community Guidelines report, that reporting is limited to nine categories of standards violations: adult nudity and sexual activity, bullying and harassment, child nudity and sexual exploitation of children, fake accounts, hate speech, regulated goods (drugs and firearms), spam, terrorist propaganda, and violence and graphic content. Reporting also does not include appeals metrics for accounts, pages, groups and events.

Santa Clara Principles. Facebook publicly supports the Santa Clara Principles:

We support the spirit of the Santa Clara Principles on Transparency and Accountability in Content Moderation and, informed by the DTAG report's findings on the challenges of content moderation at scale, are committed to continuing to share more about how we enforce our Community Standards in the future.

References and useful links:

Content Restrictions Based on Local Law transparency report:

<https://transparency.facebook.com/content-restrictions>

Community Standards Enforcement Report:

<https://transparency.facebook.com/community-standards-enforcement>

Understanding the Community Standards Enforcement Report:

<https://transparency.facebook.com/community-standards-enforcement/guide>

Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process:

<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>

Exploring feedback from data and governance experts: A research-based response to the Data Transparency Advisory Group Report:

<https://research.fb.com/exploring-feedback-from-data-and-governance-experts-a-research-based-response-to-the-data-transparency-advisory-group-report/>

GitHub



Transparent About Legal Takedown Requests. GitHub publishes a transparency report that includes the total number of government takedown requests, breaks them down by country, and reports the total number and types of content removed as a result:

In 2018, GitHub received nine requests—all from Russia—resulting in nine projects (all or part of three repositories, five gists, and one GitHub Pages site) being blocked in Russia.

In addition, GitHub commits to publicly post takedown notices from governments “to document their potential to chill speech.”

Transparent About Platform Policy Takedown Requests. GitHub has not received any platform policy takedown requests from governments, and reports that fact in its transparency report.

In a submission to UN Special Rapporteur on the right to freedom of opinion and expression David Kaye, GitHub also provides additional content for the type of requests it has received from governments to date:

When States request content removals, they invariably claim that the content violates a State law, rather than one of GitHub’s Terms of Service. Our Terms of Service prohibit unlawful content, so if a State actor were to report unlawful content as a violation of our Terms of Service, we would process that as a government takedown request. That includes confirming that the request is coming from a genuine State official and posting the government’s request. Conversely, when a non-State actor reports unlawful content as violating our Terms of Service, we do not feel compelled to take it down in the same way that we do when the takedown is demanded by a State official. GitHub does not receive other kinds of content-related requests from States.

Provides Meaningful Notice. GitHub publicly commits to notifying users of every content restriction and account suspension with the reason for removal. From its submission to UN Special Rapporteur on the right to freedom of opinion and expression David Kaye:

We notify users about content restrictions, takedowns, and account suspensions, and, when we determine a need to remove content, we provide reasons for our decision, with the ability for users to contact us to appeal the decision.

GitHub also makes specific commitments around notifying users of legal takedown requests from governments, including notifying users of geoblocking:

When we receive a notice from an official government agency that identifies illegal content and specifies the source of the illegality, we

- notify the affected users of the specific content that allegedly violates the law and that this is a legal takedown request
- allow the affected users to dispute the decision as part of that notification
- limit the geographic scope of the takedown when possible and include that as part of the notification
- post the official request that led to the takedown in this repository.

Allows Appeals. GitHub allows users to appeal content restrictions and account suspensions. From its submission to UN Special Rapporteur on the right to freedom of opinion and expression David Kaye:

We notify users about content restrictions, takedowns, and account suspensions, and, when we determine a need to remove content, we provide reasons for our decision, with the ability for users to contact us to appeal the decision.

Transparent About Appeals. GitHub does not report the number or results of appeals.

Santa Clara Principles. GitHub publicly supports the Santa Clara Principles in its transparency report:

...transparency reporting has broadened as people pay more attention to companies' practices on information disclosure and removal. One recent example is the Santa Clara Principles on Transparency and Accountability of Content Moderation Practices. We support the spirit of those principles and are working to align our practices with them with as much as possible. Through our transparency reports, we're continuing to shed light on our own practices, while also hoping to contribute to broader discourse on platform governance.

References and useful links:

Transparency report:

<https://github.blog/2019-01-23-2018-transparency-report/>

Government takedowns repository:

<https://github.com/github/gov-takedowns/>

Submission to UN Special Rapporteur on the right to freedom of opinion and expression David Kaye:

<https://www.ohchr.org/Documents/Issues/Opinion/ContentRegulation/Github.pdf>

GitHub and the United Nations free expression expert's content moderation report:

<https://github.blog/2018-05-30-github-and-un-content-moderation-report/>

GitHub contributes to UN free speech expert's report on content moderation:

<https://github.blog/2018-05-30-github-and-un-content-moderation-report/>

Google Play Store



Transparent About Legal Takedown Requests. Google publishes a transparency report that includes all government takedown requests. The transparency report states:

Governments contact Google with content removal requests for a number of reasons. Government bodies may claim that content violates a local law, and include court orders that are often not directed at Google with their requests. Both types of requests are counted in this report.

Information for each country includes the total number of takedown requests, the total number of items requested for removal, and the percentage of requests in which some content was removed. Publicly downloadable spreadsheets accompanying the report also categorize the reasons behind requests, as well as how many requests Google has received about Google Play specifically.

Transparent About Platform Policy Takedown Requests. Google publishes a transparency report that includes all government takedown requests. The transparency report states:

We also include government requests to review content to determine if it violates our own product community guidelines and content policies.

Information for each country includes the total number of takedown requests, the total number of items requested for removal, and the percentage of requests in which some content was removed. Publicly downloadable spreadsheets accompanying the report also categorize the reasons behind requests, as well as how many requests Google has received about Google Play specifically.

Provides Meaningful Notice. While Google Play publicly commits to notifying users of app removals and “warnings,” it does not commit to notifying users of legal takedown requests or specifying the scope of geoblocking when applicable.

Allows Appeals. Google Play allows users to appeal app rejections, removals, and suspensions, as well as account termination.

Transparency About Appeals. Google Play does not report the number or results of appeals.

Santa Clara Principles. Google publicly supports the Santa Clara Principles:

Google supports the spirit of the Santa Clara Principles as an effort to help shape how companies across our industry can think about transparency for action taken on content.

References and useful links:

Government requests to remove content transparency report:

<https://transparencyreport.google.com/government-removals/overview>

Download transparency report data:

https://support.google.com/transparencyreport/answer/7347561?hl=en&ref_topic=7294962

Fair warnings:

https://support.google.com/googleplay/android-developer/answer/2985876?hl=en&ref_topic=3453554

Enforcement Process:

<https://play.google.com/about/enforcement/enforcement-process/>

My app has been removed from Google Play:

<https://support.google.com/googleplay/android-developer/answer/2477981?hl=en>

Contact Google Play about an account termination or app renewal:

https://support.google.com/googleplay/android-developer/troubleshooter/2993242?visit_id=636948403288317340-3458399755&rd=1

Appeal an app removal from Google Play:

https://support.google.com/googleplay/android-developer/troubleshooter/2993242?visit_id=636948403288317340-3458399755&rd=1

Appeal a Google developer account termination:

<https://support.google.com/googleplay/android-developer/contact/accountappeals>

Understanding Google Play developer Account terminations:

https://support.google.com/googleplay/android-developer/answer/2491922?hl=en&ref_topic=3453554

Transparency report FAQ:

https://support.google.com/transparencyreport/answer/7347744?hl=en&ref_topic=7295796#

Instagram



Transparent About Legal Takedown Requests. While Instagram's parent company Facebook produces a transparency report on government legal takedown requests that breaks requests down by country and reports the number of pieces of content removed, it does not report the total number of government takedown requests received.

Transparent About Platform Policy Takedown Requests. While Instagram's parent company Facebook produces a Community Standards enforcement report that details enforcement actions, it does not specify government requests in this category and only details enforcement on 9 categories of standards violations: Adult nudity and sexual activity, bullying and harassment, child nudity and sexual exploitation of children, fake accounts, hate speech, regulated goods (drugs and firearms), spam, terrorist propaganda, and violence and graphic content.

Provides Meaningful Notice. While Instagram commits to notifying users of Community Guidelines takedowns and account suspensions, it does not commit to providing the reason for the takedown:

If we remove something that goes against our Community Guidelines, we'll tell the person who posted it, but we'll never reveal any information about the person who reported it.

For disabled accounts:

If your Instagram account was disabled, you'll see a message telling you when you try to log in.

For legal takedowns, Instagram's parent company Facebook states in its transparency report:

We provide notice to people when we restrict something they posted based on a report of an alleged violation of local law, and we also tell people when they try view something that is restricted in their country. We provide this notice except where legally prohibited or when technical constraints prevent us from doing so.

Allows Appeals. While Facebook allows users to appeal takedowns, it does not commit to allowing appeals for all Community Standards violation types:

Today, we offer appeals for the vast majority of violation types. We don't offer appeals for violations with extreme safety concerns, such as child exploitation imagery.

Because Facebook's policy does not further explain the full scope of "extreme safety concerns" that fall outside the scope of "the vast majority of violation types," it does not earn a star in this category.

For disabled accounts, Instagram does not clearly commit to allowing appeals for all provider policy-based suspensions (emphasis added):

If you think your account was disabled by mistake, you **may** be able to appeal the decision by opening the app, entering your username and password and following the on-screen instructions.

Transparent About Appeals. While Instagram's parent company Facebook reports the number of appeals and rate of content restoration in its Community Guidelines report, that reporting is limited to nine categories of standards violations: adult nudity and sexual activity, bullying and harassment, child nudity and sexual exploitation of children, fake accounts, hate speech, regulated goods (drugs and firearms), spam, terrorist propaganda, and violence and graphic content. Reporting also does not include appeals metrics for accounts, pages, groups and events.

Santa Clara Principles. Instagram's parent company Facebook publicly supports the Santa Clara Principles:

We support the spirit of the Santa Clara Principles on Transparency and Accountability in Content Moderation and, informed by the DTAG report's findings on the challenges of content moderation at scale, are committed to continuing to share more about how we enforce our Community Standards in the future.

References and useful links:

Content Restrictions Based on Local Law transparency report:

<https://transparency.facebook.com/content-restrictions>

Community Standards Enforcement Report:

<https://transparency.facebook.com/community-standards-enforcement>

Understanding the Community Standards Enforcement Report:

<https://transparency.facebook.com/community-standards-enforcement/guide>

Learn how to address abuse

<https://help.instagram.com/527320407282978>

Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process:

<https://newsroom.fb.com/news/2018/04/comprehensive-community-standards/>

What can I do if my account as been disabled?

<https://help.instagram.com/366993040048856>

Exploring feedback from data and governance experts: A research-based response to the Data Transparency Advisory Group Report:

<https://research.fb.com/exploring-feedback-from-data-and-governance-experts-a-research-based-response-to-the-data-transparency-advisory-group-report/>

LinkedIn



Transparent About Legal Takedown Requests. While LinkedIn's transparency report includes the total number of government takedown requests, breaks them down by country, and reports the number of requests on which it takes action, it does not distinguish between legal takedown requests and platform policy takedown requests.

Transparent About Platform Policy Takedown Requests. While LinkedIn's transparency report includes the total number of government takedown requests, breaks them down by country, and reports the number of requests on which it takes action, it does not distinguish between legal takedown requests and platform policy takedown requests.

Provides Meaningful Notice. LinkedIn does not publicly commit to providing meaningful notice to users of every removal and suspension.

LinkedIn's commitment to transparency regarding content blocked from the site states (emphasis added):

When we block content that you have authored due to the local legal requirements of the country from which you access LinkedIn, **we will attempt to provide you with a notification that your content has been blocked.** LinkedIn would provide this notice to the primary email address that you gave to LinkedIn or through a message on the site. In some cases, local legal requirements may prevent us from providing you with a notification that your content has been blocked.

However, that same page also states (emphasis added):

If your content or the content you attempt to access has been blocked by LinkedIn in all locations because we believe the content is illegal or violates the terms of our User Agreement and/or Professional Community Guidelines, **you may not receive a notification that this content was removed.**

Allows Appeals. LinkedIn allows users to appeal takedowns and suspensions:

If your account has been restricted or content removed and you believe the action was in error, you can appeal your case and we'll review your account. To begin the appeal process, you can log into your account and follow the onscreen messaging or reply to the message you received that provided notice of the content removal

Transparency About Appeals. LinkedIn does not report the number or results of appeals.

Santa Clara Principles. LinkedIn publicly supports the Santa Clara Principles:

We're evaluating additional ways we can expand our transparency reporting, and collaborating with other companies and groups on shared transparency goals, like those outlined in the Santa Clara Principles. We support these industry initiatives and we're working hard to provide the right information and tools for our members.

References and useful links:

Transparency report:

<https://www.linkedin.com/legal/transparency>

LinkedIn's Commitment to Transparency Regarding Content Blocked From Our Site:

<https://www.linkedin.com/help/linkedin/answer/46925/linkedin-s-commitment-to-transparency-regarding-content-blocked-from-our-site>

Account/content restricted or removed:

<https://www.linkedin.com/help/linkedin/answer/82934>

Form to request the removal of a restriction on your account:

<https://www.linkedin.com/help/linkedin/ask/hr>

Transparency report: Second half of 2018 blog post:

<https://blog.linkedin.com/2019/april-/18/transparency-report--second-half-of-2018>

Medium



Transparent About Legal Takedown Requests. Medium sends all takedown requests it receives to Lumen (formerly Chilling Effects), a database for collecting and documenting legal complaints and takedown requests for online content. Its rules state:

Medium submits to the Lumen database government requests to restrict access to content (redacted where appropriate to protect privacy or prevent harm to a person), regardless of what or whether action is taken on the request. This includes government requests to review content to determine if it violates these Rules or other Medium content policies.

Each record specifies the country from which the request originated and the URL in question, as well as an explanation for the request, the law or regulation that motivated it (if applicable), and the government agency that made the request. This level of detail is sufficient to distinguish legal requests from platform policy-based requests.

Transparent About Platform Policy Takedown Requests. Medium sends all takedown requests it receives to Lumen (formerly Chilling Effects), a database for collecting and documenting legal complaints and takedown requests for online content. Its rules state:

Medium submits to the Lumen database government requests to restrict access to content (redacted where appropriate to protect privacy or prevent harm to a person), regardless of what or whether action is taken on the request. This includes government requests to review content to determine if it violates these Rules or other Medium content policies.

Each record specifies the country from which the request originated and the URL in question, as well as an explanation for the request, the law or regulation that motivated it (if applicable), and the government agency that made the request. This level of detail is sufficient to distinguish legal requests from platform policy-based requests.

Provides Meaningful Notice. While Medium has a policy of advance notice before taking down content, as well as a policy of notice specifically for government takedown requests, it does not publicly commit to specifying in that notice the content in question and the reason for taking it down. Medium also does not commit to providing notice for account suspensions.

Regarding advance notice before disabling content, Medium's rules state:

Before disabling content associated with your account, we will give you advance notice, unless we believe your account is automated or operating in bad faith, or that notifying you is likely to cause, maintain or exacerbate harm to someone.

Regarding notice for government takedown requests, Medium's rules state:

If Medium receives a request from a government actor to restrict access to content associated with your account, we will notify you unless we are prohibited by law or believe doing so may endanger others.

Generally regarding accounting account suspensions and content restrictions, Medium's rules also state:

If it looks like you've violated our rules, we may send you an email and ask you to explain what you're up to and why. Context is important, and we want to understand the big picture. If you don't adequately explain yourself or fix the problem, we may suspend your account or remove your content. We strive to be fair, but we reserve the right to suspend accounts or remove content, without notice, for any reason, particularly to protect our services, infrastructure, users, or community.

Allows Appeals. Medium allows users to appeal takedowns and suspensions:

If you believe your content or account have been restricted or disabled in error, or believe there is relevant context we were not aware of in reaching our determination, you can write to us at yourfriends@medium.com. We will consider all good faith efforts to appeal.

Transparency About Appeals. Medium does not report the number or results of appeals.

Santa Clara Principles. Medium publicly supports the Santa Clara Principles:

Medium is committed to providing a transparent, open platform for expression and therefore supports the goals and spirit of The Santa Clara Principles on Transparency and Accountability in Content Moderation as a starting point for further discussion.

References and useful links:

Medium Rules:

<https://medium.com/policy/medium-rules-30e5502c4eb4>

Lumen:

<https://lumendatabase.org>

Pinterest



Transparent About Legal Takedown Requests. Pinterest publishes a transparency report that includes the total number of government takedown requests broken down by country and type (Community Guidelines violation or legal removal), as well as whether Pinterest complied.

Transparent About Platform Policy Takedown Requests. Pinterest publishes a transparency report that includes the total number of government takedown requests broken down by country and type (Community Guidelines violation or legal removal), as well as whether Pinterest complied.

Provides Meaningful Notice. Pinterest does not make a clear public commitment to providing meaningful notice to users of every removal and suspension. It's Terms of Service state:

Pinterest may terminate or suspend your right to access or use Pinterest for any reason on appropriate notice. We may terminate or suspend your access immediately and without notice if we have a good reason, including any violation of our Community Guidelines.

Allows Appeals. Pinterest allows users to appeal takedowns and suspensions through the "Appeals" section of its contact form, which has options to "Appeal a policy violation removal" and "Appeal account suspension".

Transparency About Appeals. Pinterest does not report the number or results of appeals.

Santa Clara Principles. Pinterest has not publicly supported the Santa Clara Principles.

References and useful links:

Transparency report:

<https://help.pinterest.com/en/articles/transparency-report>

Terms of Service:

<https://policy.pinterest.com/en/terms-of-service>

Contact form for appeals (must be logged in):

<https://help.pinterest.com/en/contact>

Reddit



Transparent About Legal Takedown Requests. Reddit publishes a transparency report that breaks down all government takedown requests by country (or, where applicable, government entity), as well as noting whether it complied and why. Reasons for removal include “legal reasons.”

Transparent About Platform Policy Takedown Requests. Reddit publishes a transparency report that breaks down all government takedown requests by country (or, where applicable, government entity), as well as noting whether it complied and why. Reasons for removal include “Content Policy Violations.”

Provides Meaningful Notice. Reddit publicly commits to providing users meaningful notice of every content takedown and account suspension.

For takedowns in response to a government legal request:

In cases where content is removed, users will be notified by a marker where the post or comment previously existed (best viewed on the desktop site), noting the legal removal. Where appropriate, rather than removing a post outright, Reddit may block the post from being accessible in a particular country. Such restrictions will be similarly noticed to users subject to them, noting the specific jurisdiction of the restriction.

For takedowns in response to a Content Policy violation, Reddit also suspends the posting user’s account. Users receive notice via private message:

Suspensions are notified via a private message. ... Information about your suspension will be specified here.

Allows Appeals. Reddit allows users to appeal content takedown and accompanying account suspensions via an appeals form. Users may also use this appeals process in the case of subreddit bans. Reddit’s transparency report states:

Whether applied against an individual account or an entire subreddit, actions taken by Reddit in response to Content Policy violations may be appealed. The appeals are evaluated by Reddit employees, and are either granted (reinstating an account / subreddit) or denied.

Further, in the event a subreddit is quarantined (a form of content restriction unique to Reddit), its moderators may appeal.

Transparency About Appeals. Reddit reports the total number of appeals it receives, as well as the aggregate outcomes of those appeals (“granted” or “denied”).

Santa Clara Principles. Reddit publicly supports the Santa Clara Principles in its transparency report:

This year, we will be sharing additional information with you about copyright removals, restorations, and retractions, as well as removals for violations of Reddit’s Content Policy and subreddit rules. Not only does this additional information increase transparency for our users, but it helps bring Reddit into line with The Santa Clara Principles on Transparency and Accountability in Content Moderation, the goals and spirit of which we support as a starting point for further conversation.

References and useful links:

Transparency report:

<https://www.redditinc.com/policies/transparency-report-2018>

Content Policy:

<https://www.redditinc.com/policies/content-policy>

Legal restrictions on content:

<https://www.reddithelp.com/en/categories/rules-reporting/account-and-community-restrictions/legal-restrictions-content>

My account was suspended for violating Reddit’s Content Policy:

<https://www.reddithelp.com/en/categories/rules-reporting/account-and-community-restrictions/my-account-was-suspended-violating>

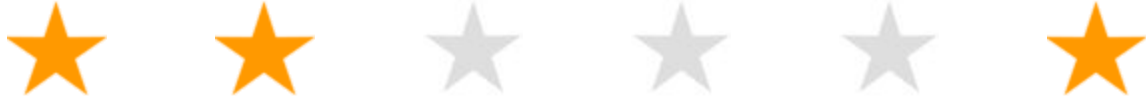
Appeals form (must be signed in):

<https://www.reddit.com/appeals>

Quarantined subreddits:

<https://www.reddithelp.com/en/categories/rules-reporting/account-and-community-restrictions/quarantined-subreddits>

Snap



Transparent About Legal Takedown Requests. Snap publishes a transparency report that breaks down all legal takedown requests by country, and notes the company's compliance rate.

Transparent About Platform Policy Takedown Requests. Snap publishes a transparency report that breaks down all platform policy takedown requests by country, as well as the number of posts taken down as a result.

Provides Meaningful Notice. Snap does not publicly commit to providing meaningful notice to users of every removal and suspension.

Allows Appeals. Snap does not have a published policy or process for users to appeal takedowns and suspensions.

Transparency About Appeals. Snap does not report the number or results of appeals.

Santa Clara Principles. Snap publicly supports the Santa Clara Principles in its transparency report:

At Snap, we support industry-wide efforts to improve content moderation reporting and transparency practices. In doing so, however, we recognize that technology platforms facilitate content creation, sharing, and retention in vastly different ways. As our platform evolves, so too, will Snap Transparency Reports, laying the groundwork to publish new categories of information to inform our community in the future. We support the spirit of the Santa Clara Principles on Transparency and Accountability in Content Moderation in creating a framework for best practices in content moderation.

References and useful links:

Transparency report:

<https://www.snap.com/en-US/privacy/transparency/>

Tumblr



Transparent About Legal Takedown Requests. While Tumblr’s parent company Oath publishes a transparency report that breaks down all government takedown requests by country, the number of “items specified” in those requests, and the company’s compliance rate, it does not report the product or platform associated with takedown requests or distinguish legal requests from platform policy-based requests.

Transparent About Platform Policy Takedown Requests. While Tumblr’s parent company Oath publishes a transparency report that breaks down all government takedown requests by country, the number of “items specified” in those requests, and the company’s compliance rate, it does not report the product or platform associated with takedown requests or distinguish legal requests from platform policy-based requests.

Provides Meaningful Notice. While Tumblr has a published policy of providing notice to users, it does not specify when users may or may not receive notice of government-ordered takedowns or suspensions. Its community guidelines state:

If we conclude that you are violating these guidelines, you may receive a notice via email. If you don't explain or correct your behavior, we may take action against your account.

Allows Appeals. Tumblr allows users to appeal content takedowns and account suspensions through Tumblr’s support interface. This interface allows users to choose a category in which their problem fits, including “Terminated blog” and “My blog was incorrectly marked as explicit.”

Transparency About Appeals. Tumblr does not report the number or results of appeals.

Santa Clara Principles. Tumblr publicly supports the Santa Clara Principles. The following language is not yet live on Tumblr’s site as of publication, but will be posted in the next several weeks on its staff blog:

Tumblr is committed to transparency, expression and our community. We support the spirit and goals of the Santa Clara Principles as a critical contribution to the discussion of how platforms can ensure that user rights are respected and valued. We look forward to continuing to engage with our users and with diverse communities and experts on these issues.

References and useful links:

Oath government removal requests report:

<https://transparency.oath.com/reports/government-removal-requests.html>

Oath transparency report FAQs and glossary:

<https://static.tumblr.com/zyubucd/gmnopeeat/combinedreport.pdf>

Community guidelines:

<https://www.tumblr.com/policy/en/community>

Support form:

<https://www.tumblr.com/support>

Staff blog:

<http://staff.tumblr.com>

Twitter



Transparent About Legal Takedown Requests. Twitter publishes a transparency report section on removal requests that includes all legal takedown requests from governments, specifying:

This section includes third-party requests that compel Twitter to remove content for legal reasons (“legal demands”) under our Country Withheld Content (“CWC”).

Twitter reports the number of legal requests per country, the type of legal request, its compliance rate, the number of accounts specified in requests, and the number of tweets and accounts ultimately withheld.

Transparent About Platform Policy Takedown Requests. While Twitter reports on platform policy takedown requests from governments in the rules enforcement section of its transparency report, it limits its reporting to six rules categories (abuse, child exploitation imagery, hateful conduct, private information, sensitive media, and violent threats) and does not break requests down by country or report whether the company acted on them.

Twitter does not include platform policy takedown requests from governments in its transparency reporting on removal requests:

[This report] does not include reports submitted by government officials to review content solely under the Twitter Rules.

Provides Meaningful Notice. While Twitter has a published policy of providing informative notice to users in the case of tweet removal and permanent account suspension, it does not commit to provide notice for legal takedown requests related to “terrorism”:

Twitter may notify you of the existence of a legal request pertaining to your account unless we are prohibited or the request falls into one of the exceptions to our user notice policy (e.g., emergencies regarding imminent threat to life, child sexual exploitation, terrorism).

Because Twitter does not commit to providing notice in cases related to “terrorism,” a class of content that is difficult to accurately identify and prone to mistakes, it does not earn a star in this category.

Allows Appeals. Twitter allows users to appeal tweet takedowns and account suspensions.

For tweet takedowns:

When we determine that a Tweet violated the Twitter Rules, we require the violator to delete it before they can Tweet again. We send an email notification to the violator identifying the Tweet(s) in violation and which policies have been violated. They will then need to go through the process of deleting the violating Tweet or appealing our review if they believe we made an error.

For permanent account suspensions:

Violators can appeal permanent suspensions if they believe we made an error. They can do this through the platform interface or by filing a report. Upon appeal, if we find that a suspension is valid, we respond to the appeal with information on the policy that the account has violated.

For other types of account locks and suspensions:

If you are unable to unsuspend your own account using the instructions above and you think that we made a mistake suspending or locking your account, you can appeal.

Further, Twitter provides step-by-step instructions for users whose accounts have been temporarily locked or limited, and allows appeals via a specific form for locked or suspended accounts.

Transparency About Appeals. Twitter does not report the number or results of appeals.

Santa Clara Principles. Twitter publicly supports the Santa Clara Principles in its transparency report:

We support the spirit of the Santa Clara Principles on Transparency and Accountability in Content Moderation, and are committed to sharing more detailed information about how we enforce the Twitter Rules in future reports.

References and useful links:

Removal requests report:

<https://transparency.twitter.com/en/removal-requests.html>

Twitter rules enforcement:

<https://transparency.twitter.com/en/twitter-rules-enforcement.html>

Our range of enforcement options:

<https://help.twitter.com/en/rules-and-policies/enforcement-options>

Legal requests FAQ:

<https://help.twitter.com/en/rules-and-policies/twitter-legal-faqs>

About suspended accounts:

<https://help.twitter.com/en/managing-your-account/suspended-twitter-accounts>

Form to appeal an account suspension or locked account:

<https://help.twitter.com/forms/general?subtopic=suspended>

Help with locked or limited accounts:

<https://help.twitter.com/en/managing-your-account/locked-and-limited-accounts>

Vimeo



Transparent About Legal Takedown Requests. Vimeo does not publish a transparency report.

Transparent About Platform Policy Takedown Requests. Vimeo does not publish a transparency report.

Provides Meaningful Notice. Vimeo does not publicly commit to providing meaningful notice to users of every removal and suspension.

Allows Appeals. Vimeo does not have a published policy or process for users to appeal takedowns and suspensions.

Transparent About Appeals. Vimeo does not report the number or results of appeals.

Santa Clara Principles. Vimeo has not publicly supported the Santa Clara Principles.

References and useful links:

Terms of service:

<https://vimeo.com/terms>

WordPress.com



Transparent About Legal Takedown Requests. WordPress.com’s parent company Automattic publishes a transparency report in which it reports the number of takedown requests per country, whether they were court orders or requests from government agencies/law enforcement, the number of sites specified in the requests, and if action was taken due to a policy violation or “solely in response to the demand.”

Transparent About Legal Takedown Requests. WordPress.com’s parent company Automattic publishes a transparency report in which it reports the number of takedown requests per country, whether they were court orders or requests from government agencies/law enforcement, the number of sites specified in the requests, and if action was taken due to a policy violation or “solely in response to the demand.”

Provides Meaningful Notice. While WordPress.com has a published policy of providing notice to users, it does not specify when users may or may not receive notice of government-ordered takedowns or suspensions:

Depending on the scenario, we will email you or add a warning notification in your dashboard. The notification will contain a link that you can use to contact us regarding the issue, and you can always contact us via email or via the form below for further explanation.

Allows Appeals. WordPress.com allows users to appeal takedowns, suspensions, or other errors:

We do make mistakes from time to time. If you feel that we’ve done anything in error, please contact us via the link on your dashboard or by using the form below. A real person will review your request and reply with our decision as soon as possible.

Transparency About Appeals. WordPress.com does not report the number or results of appeals.

Santa Clara Principles. WordPress.com has not publicly supported the Santa Clara Principles.

References and useful links:

Takedown demands report:

<https://transparency.automattic.com/takedown-demands/takedown-demands-2018-h2/>

Suspended content and sites:

<https://en.support.wordpress.com/suspended-blogs/>

YouTube



Transparent About Legal Takedown Requests. YouTube’s parent company Google publishes a transparency report that includes all government takedown requests. The transparency report states:

We receive content removal requests through a variety of avenues and from all levels of government — court orders, written requests from national and local government agencies, and requests from law enforcement professionals.

Information for each country includes the total number of takedown requests, the total number of items requested for removal, and the percentage of requests in which some content was removed. Google’s publicly downloadable spreadsheets also categorize the reasons behind requests, as well as how many requests it has received about YouTube specifically.

Transparent About Platform Policy Takedown Requests. YouTube’s parent company Google publishes a transparency report that includes all government takedown requests. The transparency report states:

We also include government requests to review content to determine if it violates our own product community guidelines and content policies.

Information for each country includes the total number of takedown requests, the total number of items requested for removal, and the percentage of requests in which some content was removed. Google’s publicly downloadable spreadsheets also categorize the reasons behind requests, as well as how many requests it has received about YouTube specifically.

Provides Meaningful Notice. While YouTube publicly commits to providing notice of “strikes”—which are when a review results in channel terminations or content takedowns based on Community Guidelines violations—it does not publicly commit to notifying users of legal takedowns.

For channel terminations:

When a channel is terminated, the channel owner receives an email explaining the reason for the termination.

For content takedowns due to Community Guidelines violations:

If a strike is issued, we’ll let you know by email, through notifications on mobile and desktop, and in your Channel Settings. We’ll also tell you:

- What content was removed
- Which policies it violated (for example sexual content or violence)
- How it affects your channel
- What you can do next

Allows Appeals. YouTube allows users to appeal takedowns and suspensions.

For content takedowns, users follow the process to appeal Community Guidelines actions.

For account suspensions, users can appeal through a dedicated form.

Transparency About Appeals. YouTube does not report the number or results of appeals.

Santa Clara Principles. YouTube's parent company Google publicly supports the Santa Clara Principles:

Google supports the spirit of the Santa Clara Principles as an effort to help shape how companies across our industry can think about transparency for action taken on content.

References and useful links:

Government requests to remove content transparency report:

<https://transparencyreport.google.com/government-removals/overview>

Download transparency report data:

https://support.google.com/transparencyreport/answer/7347561?hl=en&ref_topic=7294962

Legal complaints:

<https://support.google.com/youtube/answer/3001497?hl=en>

Channel terminations:

<https://support.google.com/youtube/answer/2802168?hl=en>

Community Guidelines strike basics:

<https://support.google.com/youtube/answer/2802032?hl=en>

Appeal Community Guidelines actions:

https://support.google.com/youtube/answer/185111?hl=en&ref_topic=2803138

“Unable to access a Google product” appeals form:

https://support.google.com/accounts/contact/suspended?p=youtube&visit_id=1-636610084672726027-1458380864&rd=1

Transparency report FAQ:

https://support.google.com/transparencyreport/answer/7347744?hl=en&ref_topic=7295796#