Account Deactivation and Content Removal:
Guiding Principles and Practices for Companies and Users

Erica Newland, Caroline Nolan, Cynthia Wong, and Jillian York[†]

[†] Erica Newland is a Policy Analyst at the Center for Democracy & Technology. Caroline Nolan is a Senior Project
Manager at the Berkman Center for Internet & Society. Cynthia Wong is Director of the Center for Democracy &
Technology's Project on Global Internet Freedom. Jillian York is Director for International Freedom of
Expression at the Electronic Frontier Foundation. The authors wish to thank their colleagues at the Global
Network Initiative (GNI) for their feedback and suggestions for improvement. They are also grateful to
participants in the GNI's shared learning workstream on account deactivation and content removal for their inputs
and review.

Executive Summary

From the role of Facebook during protests in the Middle East and North Africa,[1] to the use of YouTube, Twitter, and other tools in the wake of earthquakes in Haiti and Japan and the wildfires in Russia,[2] platforms that host user-generated content (UGC) are increasingly being used by a range of civic actors in innovative ways: to amplify their voices, organize campaigns and emergency services, and advocate around issues of common concern. However, while these platforms and services may be perceived as public, their users are subject to the rules and controls that private companies create and implement. Intentionally or not, private entities assume a primary role in providing and controlling access to the 'networked public sphere.' This 'networked public sphere' has supplanted, in part, the traditional town square by providing an open and dynamic online space for social and political debate and activism where citizens around the world increasingly exercise their rights to free expression, opinion, assembly, and association. Platform operators are often faced with challenging decisions regarding content removal or account deactivation, which, even when valid or reasonable, may have costly implications for the rights of users, especially activists.

This report explores these dilemmas, and recommends principles, strategies, and tools that both UGC platforms and users can adopt to mitigate the negative effects of account deactivation and content removal. We use select examples to highlight good company practices, including efforts to balance complex and often competing considerations—the enforcement of site guidelines, responses to government pressure, the free expression and privacy rights of users, and the potential risks faced by activists—in consistent, transparent, and accountable ways. Importantly, this report does not put forth a one-size-fits-all solution for the complex set of challenges raised by Terms of Use (ToU) enforcement. Platforms vary in terms of history, mission, content hosted, size, and user base, and no single set of practices will be an appropriate fit in every case. Moreover, while the examples in this report focus on platforms that host social media, the recommendations are broadly applicable to companies that host different types of user-generated content.

Most critically, this report demonstrates the ways in which companies can have a significant impact on user rights *and* user satisfaction by being clearer and more consistent in how they implement ToU and interact with users. While the costs associated with creating channels for customer support, responses to user inquiries, appeals processes, and similar mechanisms should not be underestimated, positive outcomes often rely on proactive and transparent communications with users from the outset and at each stage of interaction between the company and a user. Given the growth of digital activism and the explosion of user-generated content,

---

[1] Jillian York, "How are protestors in Egypt using social media?," *Jilliancyork.com*, January 27, 2011, http://jilliancyork.com/2011/01/27/how-are-protestors-in-egypt-using-social-media; Ethan Zuckerman, "What if Tunisia had a revolution, but nobody watched?," …*My Heart is In Accra,* January 12, 2011, http://www.ethanzuckerman.com/blog/2011/01/12/what-if-tunisia-had-a-revolution-but-nobody-watched.

[2] Gregory Asmolov, "Russia: Crowdsourcing assistance for victims of wildfires," *Global Voices*, August 3 2010, http://globalvoicesonline.org/2010/08/03/russia-crowdsourcing-assistance-for-victims-of-wildfires.

these challenges will only multiply with time, underscoring the value of developing and implementing effective systems to address them now.

This document draws on conversations emerging from an ongoing learning series hosted by the Global Network Initiative (GNI).[3] It builds on the work and input of diverse participants, including company representatives, socially responsible investors, academics, and advocates.[4]

Summary of Recommendations for Companies

A. Offer Clear, Consistent, and Transparent ToU and Guidelines

- Use clear and accessible language during the sign-up process, in terms of use, and in other forms of user guidelines.
- Translate ToU and user guidelines into languages in which services and features are offered.
- Clearly communicate to users when and how ToU vary across jurisdictions.

B. Respond when a Suspected Violation of ToU is Identified

- Require users who flag inappropriate content to specify the rule the content allegedly violates.
- Accompany warnings, content removal and blocking, and account deactivations with immediate notice, clear explanation of the violation, and descriptions of next steps.
- Provide for intermediate steps or an escalation process during the content or account review process.
- Provide clear channels through which users can contact the platform with complaints, questions, and issues.

C. Provide Opportunities for Recourse: Appeals, Due Process, and Data Export

- Develop an appeals process to mitigate the impact of mistakes or abuse by third parties.
- Respond to appeals promptly and communicate to users why the action was taken and what they should expect from the appeals process.
- Provide users with options to preserve and export data upon the deactivation of accounts.

D. Embed Human Rights Considerations into Company Practice and Platform Design

- Consider temporal, political, geographic, and other contexts when evaluating content.

---

[3] The Global Network Initiative is a multi-stakeholder group of companies, civil society organizations (including human rights and press freedom groups), investors, and academics that has created a collaborative approach to protect and advance freedom of expression and privacy in the ICT sector. *See* GNI, "Principles on Freedom of Expression and Privacy," http://www.globalnetworkinitiative.org/principles/index.php.

[4] *See* GNI, "GNI Convenes Stakeholders on Account Deactivation and Human Rights Issues," May 11, 2010, http://www.globalnetworkinitiative.org/newsandevents/Account_Deactivation_call.php.

- Provide relevant human rights education for staff, including review teams and platform managers.
- Engage academic institutions, human rights groups, and local NGOs regarding emerging issues, local contexts, and other factors that may influence user activity and content and inform company decision-making.
- Incorporate human rights considerations into both the policy development process and ongoing evaluations of existing policies.
- Develop specific guidance for human rights activists regarding good practice and relevant rules.

Summary of Recommendations for Users

- Educate yourself about the platforms that you engage with and use these services in a responsible manner.
- Use labels, tags, and other cues to provide contextual and other relevant information regarding your content.
- Where possible, keep copies of your materials or upload them to another service as a back-up.
- Engage with companies and other stakeholders by participating in networks, dialogues, and other efforts aimed at deepening understanding of emerging issues that may impact user content, rights, and activities.

Introduction

"The Internet is not a public sphere. It is a private sphere that tolerates public speech." – Clay Shirky

Online services such as social networking sites and blog-hosting sites are important tools for citizen journalists, political campaigners, human rights advocates, and the general public to express their points of view and to organize around common interests. However, while these platforms may be perceived as "public," their users are subject to the rules and controls that private companies create and implement, whether in the form of Terms of Use (ToU), user agreements, acceptable use policies, or other parameters. Users must operate within the confines of these guidelines or risk losing their accounts, their networks, their ability to post content, and maybe even content they have created and posted when service providers enforce those terms or exert other forms of control in response to potential threats, misuse, or ToU violations.

Content removal and account deactivations can occur for various reasons and across different contexts. They can be triggered by user misunderstandings of ToU, hacking incidents, illegitimate takedowns (resulting from organized user campaigns to "flag" content), and government or third party requests. Some national laws may either require or strongly incent content removal in certain circumstances, such as through notice-and-takedown safe harbors.

In some cases, account deactivation and content removal decisions can have severe and negative consequences for human rights.[1] Companies must therefore seek to balance complex and often competing considerations—enforcing site guidelines, maintaining liability protections, protecting the free expression and privacy rights of users—in consistent, transparent, and accountable ways. This report considers these dilemmas and offers a high level set of practices aimed at informing both company and user behavior on these platforms.

This report examines several areas where UGC platforms can mitigate the human rights risks of account deactivation and content removal decisions. It also highlights steps users can take to protect their content from removal, work within platform guidelines, and employ strategies to avoid account deactivation. Recommendations fall in the following areas:

A.) Offer Clear, Consistent, and Transparent ToU and Guidelines
B.) Respond when a Suspected Violation of the ToU is Identified
C.) Provide Opportunities for Recourse: Appeals, Due Process, and Data Export
D.) Embed Human Rights Considerations into Company Practice and Platform Design
E.) Recommendations for Users

---

[1] *See, e.g.,* Jennifer Preston, "Ethical Quandary for Social Sites," *New York Times*, March 17, 2011, at B2, http://www.nytimes.com/2011/03/28/business/media/28social.html; Adrian Chen, "Why Facebook Should Do More to Help Egypt's Protesters," *Gawker*, February 5, 2011, http://gawker.com/#!5752904/why-facebook-should-do-more-to-help-egypts-protesters.

Each section describes the issues raised, provides illustrative examples, examines existing good practice, and presents recommendations to service providers. While these recommendations are designed to apply to diverse companies across the UGC platform landscape, smaller companies may face resource or capacity constraints that hamper, for example, the implementation of sophisticated systems for handling appeals or evaluating content. However, actively engaging with networks of users, activists, companies, and other stakeholders can lead to the development and adoption of innovative, collaborative, and less resource-intensive models for handling account deactivation and content removal issues. Finally, all platforms, regardless of their size, should be transparent and consistent in their interactions with users and in how they establish, communicate, and implement the "ground rules."

As the 'networked public sphere' has become the primary locus of modern public discourse, many platforms have sought to develop systems and adopt practices in response to the needs of their diverse user bases. Throughout this document, examples are drawn from some of the most popular global online platforms because their services are most likely to be used for activism or political campaigns.[2] While examples have been selected as a means of highlighting good practices, even successful policies should be continually evaluated, improved upon, and refined in response to changes in user activity and contexts. Moreover, even though many platforms have adopted certain good policies or practices, no platform is without room for improvement. Finally, while we acknowledge that national laws may require or strongly incent company action against certain content or user behavior, we focus our report here on areas where companies have discretion in implementing *how* they enforce their terms of use.

Our recommendations draw on a number of efforts focused on establishing guiding principles and rights-sensitive practices for companies, including the Global Network Initiative's Principles on Free Expression and Privacy;[3] the Protect, Respect, and Remedy framework created by John Ruggie, the United Nations Special Representative of the Secretary General on transnational corporations and other business enterprises;[4] and the OECD's "Guidelines for Multinational Enterprises—Recommendations for Responsible Business Conduct in a Global Context."[5] This document also draws conclusions coming out of an ongoing learning series hosted by the Global Network Initiative (GNI) and builds on the work and input of participants in the series, including company representatives, socially responsible investors, academics, and advocates.[6]

---

[2] For additional information regarding the innovative use of common Web 2.0 tools by activists, see Ethan Zuckerman, "The Cute Cat Theory Talk at ETech," *…My Heart is in Accra*, March 3, 2008, http://www.ethanzuckerman.com/blog/2008/03/08/the-cute-cat-theory-talk-at-etech.

[3] The Global Network Initiative is a multi-stakeholder group of companies, civil society organizations (including human rights and press freedom groups), investors, and academics that has created a collaborative approach to protect and advance freedom of expression and privacy in the ICT sector. *See* GNI, "Principles on Freedom of Expression and Privacy," http://www.globalnetworkinitiative.org/principles/index.php.

[4] *See* John Ruggie, "Protect, Respect and Remedy: a Framework for Business and Human Rights," Human Rights Council, A/HRC/8/5, April 7, 2008, http://www.reports-and-materials.org/Ruggie-report-7-Apr-2008.pdf.

[5] *See* OECD, "OECD's Guidelines for Multinational Enterprises – Recommendations for Responsible Business Conduct in a Global Context," May 25, 2011, http://www.oecd.org/dataoecd/43/29/48004323.pdf.

[6] *See* GNI, "GNI Convenes Stakeholders on Account Deactivation and Human Rights Issues," May 11, 2010, http://www.globalnetworkinitiative.org/newsandevents/Account_Deactivation_call.php; GNI, "Account

Our aim is to identify realistic and concrete approaches that are rights-sensitive while also flexible enough to be practically implemented across diverse platforms and contexts.

Proposed Good Practices

A. Offer Clear, Consistent, and Transparent ToU and Guidelines

Before users first interact with a UGC platform—upload their pictures to Flickr, add a video to YouTube, or 'friend' their contacts on Facebook—they typically must register with the service in some way. They are often required to indicate that they understand the intended use of the site by creating a profile, for example, or by clicking on a box that signals that they agree to abide by the ToU and any other policies that govern their accounts and activities. Although these policies can cover a range of issues and vary widely across platforms, they often address acceptable user behavior on the site, the company's privacy, marketing, and other policies, and copyright notices.

Users often do not read ToU when setting up their accounts, in part because such documents tend to be written in dense, legalistic terms or may not be available in users' primary languages.[7] However, as users begin to post content or engage with others via the platform, they often seek out guidance on acceptable use. By providing transparent, accessible, and practical user information and educational tools, platforms can help users avoid unintended rule violations while strengthening user trust.

Use clear and accessible language during the sign-up process, in ToU, and in other forms of user guidelines
The need for clarity begins with the user sign-up process. Especially when a platform establishes policies that may not comport with users' expectations, such as requirements that users sign up under their legal names, the platform should take special care to educate users about these policies during the signup process and to indicate that violation of these policies can result in account deactivation or content removal; they should not relegate a discussion of such policies to the middle of a long or complex ToU.

As users sign up for a platform, the ToU establishes what types of behaviors are appropriate and inappropriate. Service providers should strive to craft ToU that are as easy to understand as

---

Deactivation and Content Removal: Takeaways from Oct. 28 Call," Nov. 16, 2010, http://www.globalnetworkinitiative.org/newsandevents/Account_Deactivation_and_Content_Removal.php; GNI, "GNI Account Deactivation and Content Removal Call," July 26, 2011, http://www.globalnetworkinitiative.org/newsandevents/GNI_Account_Deactivation_and_Content_Removal_Call.php.

[7] Yannis Bakos, Florencia Marotta-Wurgler & David R. Trossen, "Does Anyone Read the Fine Print? Testing a Law and Economics Approach to Standard Form Contracts," NYU Law and Economics Working Paper No. 09-04, NET Institute (October 6, 2009), available at http://ssrn.com/abstract=1443256. *See also* Nathaniel Good et al., "Stopping Spyware at the Gate: A User Study of Privacy, Notice and Spyware," Symposium on Usable Privacy and Security, July 6-8, 2005, available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.6435&rep=rep1&type=pdf.

possible. ToU should explicitly outline what constitutes acceptable and unacceptable uses of their platforms; how the service provider will respond to a violation of its rules; how users can find out more information about account deactivations and content removal; and how users can appeal account deactivations or content removal.

Complex language in the ToU may leave the average user unsure of what she is signing up for, what violations may result in content removal or account deactivation, or what options for recourse exist. Categories of controversial speech may also vary across cultures and contexts— for example, understandings of what constitutes hate speech or pornography can differ widely. Therefore, providing more detailed guidance about what constitutes prohibited behavior is preferable. For example, a platform might clarify that it prohibits "racial or ethnic slurs, profanity, sexual harassment, and bullying" instead of merely "inappropriate language or behavior."[8] Where possible, service providers should post such standards in multiple places and through multiple channels; they should also ensure that their public representations about their policies in fact align with the language in their ToU and with their actual enforcement practices.

Many platforms provide separate, more informally written user guidelines, FAQs, or help centers that detail key points in the ToU. As companies seek to clarify and communicate the rules, this is an area where they might engage community members in the process, via message boards, tips, and other information.

YouTube provides tools that are instructive in this regard. Its "Community Guidelines" include a number of rules, written in plain language, that advise users on how to "respect the community," how not to "cross the line," and how guidelines are enforced.[9] Videos featured on the same page further explain and clarify the terms through which users might be flagged for potential use violations. A "Tips" section provides additional detail and draws on familiar or widely accessible examples. For example, a section titled "Shocking and Disgusting" suggests that:

> [i]f a video is particularly graphic or disturbing, it should be balanced with additional context and information. For instance, including a clip from a slaughterhouse in a video on factory farming may be appropriate. However, stringing together unrelated and gruesome clips of animals being slaughtered in a video may be considered gratuitous if its purpose is to shock rather than illustrate.

Other factors that may impact content takedown decisions for categories such as "Sex and Nudity," "Hate Speech," and "Dangerous Illegal Acts" are also described.[10] Clear parameters are also set regarding copyright, privacy, harassment, impersonation, and threats.

---

[8] For example, Facebook's Statement of Rights and Responsibilities clearly explains that it prohibits content that is "hateful, threatening, or pornographic; incites violence; or contains nudity or graphic or gratuitous violence." *See* Facebook, "Statement of Rights and Responsibilities," http://www.facebook.com/terms.php (accessed October 4, 2010).

[9] YouTube, "YouTube Community Guidelines," http://www.youtube.com/t/community_guidelines.

[10] *Ibid.*

Twitter's help section similarly includes a thorough set of explanations for users regarding how they can avoid account deactivations. In a section that is somewhat counter intuitively titled "Report a violation,"[11] Twitter offers a set of "Twitter Rules" that clearly lays out the types of behaviors that can result in an account suspension.[12] Twitter also describes some of the factors that it considers when deciding whether or not an account has violated its ToU and includes two sets of "best practices" that detail how users can avoid violating the ToU when using automated-tweeting tools and when adding new followers.[13] The "Report a Violation" section also features pages that detail Twitter's policies with regard to specific activities that violate its ToU—such as name squatting, impersonation, and the posting of child pornography—and how Twitter responds when a user is identified as potentially engaging in one of these specific activities.[14]

In early 2011, Facebook adopted a similar approach and began providing clearer guidelines on the types of material and behavior that can result in content removal and account deactivation. Designed to supplement its pre-existing Statement of Rights and Responsibilities (the equivalent of Facebook's ToU), Facebook's Community Standards page includes a preamble that articulates the type of environment and community the site is aiming to cultivate; the rest of the document clearly describes rules aimed at protecting users from negative, offensive, or unsafe content (e.g., threats, hate speech, theft, vandalism, harassment, bullying, etc.).[15]

Translate ToU and user guidelines into languages in which services and features are offered
Service providers should endeavor to translate their ToU and other user guidelines into the languages in which they offer their platforms.[16] In some instances, as platforms translate their services into new languages, they have neglected to translate their ToU and guidelines in parallel. As a result, users who have been unable to review the platform's rules in their own language may be kicked off for a violation of these rules. In cases where creating a binding ToU in all languages of use would require significant resource investment, some platforms provide non-binding translations of their ToU in the languages in which they offer their service, with the original ToU serving as the official, binding version.

---

[11] Report a Violation, http://support.twitter.com/groups/33-report-a-violation#topic_121 (last visited June 14, 2011).

[12] Twitter, "The Twitter Rules," http://support.twitter.com/articles/18311.

[13] Twitter, "The Twitter Rules," http://support.twitter.com/articles/18311; Twitter, "Twitter Automation Rules and Best Practices," http://support.twitter.com/articles/76915; Twitter, "Following Rules and Best Practices," http://support.twitter.com/articles/68916.

[14] *See, e.g.,* Twitter, "Name Squatting Policy," http://support.twitter.com/articles/18370; Twitter, "Impersonation Policy," http://support.twitter.com/articles/18366; Twitter, "Child Pornography Policy," http://support.twitter.com/articles/37370.

[15] Facebook, "Facebook Community Standards," http://www.facebook.com/communitystandards. In addition, Facebook has recently launched a beta version of an interactive privacy policy, also referred to as their "data use policy." The policy offers explanations of the company's data collection and use practices that are more understandable and comprehensive than the traditional, text-based, legalistic privacy policy. Facebook Data Use Policy, https://www.facebook.com/about/privacy/ (last visited July 7, 2011). *See also*, "A Privacy Policy Re-Imagined for Users Like You," *Facebook.com*, February 25, 2011, https://www.facebook.com/notes/facebook-site-governance/a-privacy-policy-re-imagined-for-users-like-you/10150434660350301.

[16] We recognize that in some cases, a company's decision to translate their platform's ToU into another language can strengthen another country's claim of jurisdiction over that platform's service, which may be undesirable if that country's government has a poor human rights record. Platforms should evaluate this risk when making decisions about whether or not to translate the ToU.

A number of companies have experimented with crowd-sourcing to further translation efforts and reduce the cost of language scaling. For example, Facebook's Community Standards have been translated into nearly 30 languages and have been opened up to community translation for the rest of the languages that Facebook supports. The Community Standards page is supplemented by a help section that has been translated into dozens of languages.

Clearly communicate to users when and how ToU vary across jurisdictions
Jurisdictional variations can also play a role in how ToU are applied. In cases in which platforms have developed local ToU in order to comply with the respective legal standards (related to hate speech, for example), those distinctions should be clearly communicated to users. For example, in its instructions to users regarding how to appropriately categorize and label photos, Flickr notes that "[i]f your login ID is based in Singapore, Hong Kong, India, Korea or with Maktoob.com you will only be able to view safe content based on your local Terms of Service."[17] Users in these countries have access to only that content which is deemed safe by Flickr staff in adherence to local standards.[18]

B. Respond when a Suspected Violation of ToU is Identified

As noted in the previous section, platforms should communicate with users proactively in order to reduce inadvertent violations of existing rules. When a suspected violation of the ToU has been identified, platforms should have strategies for mitigating the harm that may result from false positives.

UGC platforms have adopted a range of tools and processes for identifying and reporting ToU violations, including user-driven flagging mechanisms, automated responses based on pre-determined criteria, human review, or some combination of mechanisms (we explore these methods further in Section D). These processes naturally vary depending on the form of content allowed on a platform, the substantive ToU, and even legal requirements. Importantly, these processes are never perfect: automated moderation and abuse prevention processes, even when coupled with human review, can be subject to error; community reporting systems are also vulnerable to user abuse.

When a ToU violation is detected or reported, the platform should clearly communicate the nature of the violation to the user so that the user has an opportunity to adjust her behavior (or challenge the rule violation if she believes a rule has been misapplied). The means through which companies convey warnings or notices of violations to users should be transparent, timely, and defined from the outset—for example, such communication may come in the form of a flagged account, an interstitial pop-up window, or an email notice. Where content is removed or blocked or an account deactivated, the platform should communicate the reason for the adverse action

---

[17] Flickr, "Help/Filters," http://www.flickr.com/help/with/other/ (last visited July 6, 2011).

[18] Jillian C. York, "Policing Content in the Quasi-Public Sphere" at 5, *OpenNet Initiative*, http://opennet.net/sites/opennet.net/files/PolicingContent.pdf.

and explain any processes for appeals or further action that either the platform or the user may take.

Require users who flag inappropriate content to specify the rule the content allegedly violates
Many platforms offer mechanisms through which community members can flag content or user activity that may be in violation of the platform's ToU. These mechanisms should be designed to gather as much information as possible about the nature of the alleged violation and to educate users about when it is inappropriate to submit a complaint about a piece of content or another user's activity. Well-designed complaint mechanisms can both discourage users from flagging material that is not in fact a violation of the platform's ToU and provide context for staff members who ultimately review the flagged content. For example, platforms that require users to indicate why they entered a complaint will have more useful information for evaluation: a gulf between the reported reason the content was flagged and the nature of the content would suggest abuse of the system. Service providers might ultimately give less credence to complaints filed by users who consistently flag material that does not actually violate the ToU or to users whose stated reasons for flagging material do not align with the nature of the material.

Facebook's model provides a useful example. Each Facebook profile provides a link that enables users to "Report/Block this person." If the user selects the report option, she is required to indicate a reason for the complaint, with options including (but not limited to) "this person is annoying me," "inappropriate profile photo," and "inappropriate profile information." When a user selects "this person is annoying me," Facebook suggests that she block the person, under the header "what you can do." No opportunity is offered for the user to submit a complaint. Facebook thereby establishes individual blocking as an appropriate strategy for handling "annoying" users who are otherwise not violating the ToU. Users who select "inappropriate profile photo" are also taken to another interstitial, where they are prompted to offer a reason why the photo is inappropriate. Potential reasons include (but are not limited to) "I don't like this photo of me," "nudity or pornography," and "graphic violence." Complaints are submitted when users select "nudity or pornography" or "graphic violence." If a user selects, "I don't like this photo of me," Facebook offers two options, again under the heading "what you can do." The user can either message the owner of the profile to ask that the photo be removed or she can block the user.[19] Both examples present models that benefit users and companies alike: users are empowered to shape their own activities and use environments and the number of complaints that companies have to process or evaluate is reduced.

Accompany warnings, content removal and blocking, and account deactivations with immediate notice, clear explanations of the violation, and descriptions of next steps
When user behavior or content has drawn scrutiny for ToU violations, communications to the user should strive to explain why the platform is issuing a warning or taking other adverse action against the user and what comes next. For example, such communication should include:

- a description of which specific rules the user's actions allegedly violate;
- a mechanism through which the user may request more information about the allegation (e.g., a web form or email address);

---

[19] *See also* Facebook, "Help Center – Privacy: How do I report abuse?," http://www.facebook.com/help/?page=843 (click on "Privacy: Report abuse").

- a mechanism through which the user may contact the company to challenge the allegation, in case the user believes a rule has been misapplied to her content; and,
- links to information explaining the service provider's process for responding to user communications, including timing and its processes for handling appeals.

When communicated along with a content takedown notice or an intermediate warning, such information educates the user, gives the user a meaningful opportunity to adjust her behavior, and helps prevent future violations. When content removal or account deactivation (or other adverse intermediate action) is imminent, communication with the user can allow her to mitigate the impact of adverse actions against her account and to appeal if she believes a mistake has been made.

The notice should be written in the language in which the user engages with the platform; for example, if the user engages with the French-language version of the platform, then the notice should be in French. Except where prohibited by law, a user should also have the opportunity to salvage her content or account information, even if her account is no longer publicly available.

Some platforms use automated filters to block and prevent certain links, images, videos, or words from being posted. When a filter has blocked posting of a piece of content or a message, the platform should immediately notify the user—either through an interstitial or through email—that the material did not successfully post. The notice should explain why the material was deemed inappropriate (for example, it used language that violates the ToU or contained a URL known to link to illicitly obtained copyrighted material), so that the user can adjust her behavior. As with content removal decisions, the notice should include information about how the user can appeal the action.

In some cases, companies will want to—or may be compelled by law to—take down certain categories of content immediately, without warning or use of escalation procedures. These exceptional categories might include content such as child abuse images or instances where rapid action is necessary to avoid immediate harm to third parties (such as victims of domestic violence). In implementing such policies, however, companies should still provide timely appeals processes as mistakes or abuse of flagging systems can occur.

Provide for intermediate steps or an escalation process during the content or account review process
Service providers should consider taking intermediate steps before removing content or deactivating accounts completely. These escalation processes should also include a user education component: platforms will more effectively shape user behavior by explaining to users precisely which actions are in violation of the ToU and which provision of the Terms these actions have violated. For example, when Blogger identifies a blog with adult content that has not been properly labeled as adult, the service sometimes inserts an unavoidable "mature content" interstitial between the referring URL and the allegedly mature content blog. The interstitial allows Blogger to warn users that certain content may be inappropriate for some users. In other instances, instead of removing a blog altogether, Blogger will "sandbox" it for a certain period of time, during which only the blog author can access the content.[20]

---

[20] Blogger, "Blogger Content Policy," http://www.blogger.com/content.g.

YouTube also employs an escalation model for enforcing its general policies. When content is removed for violating its Community Guidelines, the user who posted it receives a strike, which is conveyed via email and also pops up the next time she logs onto YouTube.[21] Associated penalties are clearly described on YouTube's Policy Enforcement page and include a clear timeline—the first violation leads to a warning, the second violation within a six-month period results in the disabling of the user's ability to post new content to YouTube for two weeks, and the third violation within a six-month period results in account deactivation.[22]

In both examples, these intermediate steps allow for good faith misunderstandings of ToU while giving users opportunities to align their behavior with the rules of the platform and thereby avoid account deactivation.

Provide clear channels through which users can contact the platform with complaints, questions, and issues

Terms of use should include a mechanism by which users can contact the company with questions or complaints about its policies. For example, Flickr offers an easy-to-find web form through which users can email their questions about the platform. The body of Flicker's form poses directed questions that prompt the user to provide the most helpful types of information about their question or issue.[23]

Given the volume and diversity of potential complaints and the geographic and linguistic diversity of end-users on some platforms, managing such a contact mechanism can present practical challenges. Companies should nonetheless seek to implement and communicate to users a reasonable policy for responding to questions and complaints in a timely manner. Such policies can further user education, while also helping the service provider better understand how their platform is being used by a range of users.

C. Provide Opportunities for Recourse: Appeals, Due Process, and Data Export

Content hosts use a variety of tools and community-driven models to discover and police harmful content and violations of the ToU. Even in cases where platforms set out clear guidance for when content may be removed or under what circumstances user accounts are disabled, the means through which they monitor and enforce those standards have a marked impact on the number of erroneous content removal or account deactivation decisions. For example, a number of platforms deploy versions of a community-based model, relying on other users to monitor and flag content that potentially violates community guidelines or ToU standards. YouTube, for example, offers multiple tools to enable community members to alert staff to allegedly offending

---

[21] YouTube, "General Policy Enforcement," http://www.google.com/support/youtube/bin/answer.py?answer=92486.

[22] *Ibid.*

[23] Flickr, "Help/Other Issues," http://www.flickr.com/help/with/other.

or infringing content, starting with a flagging tool embedded in the video player.[24] Staff members then review those flagged videos "24 hours a day, seven days a week" to determine whether they violate Community Guidelines.[25]

Given the large quantity of material uploaded each minute to the most popular platforms for user-generated content,[26] combining user-driven flagging with some automated processes can be invaluable to maintaining community norms. However, a system in which users are enlisted to monitor others can lead to abuse;[27] users may flag content that they disagree with, that offends the government or company that they represent, that is posted by individuals against whom they have a personal grudge, or that offends their sensibilities, but that does not violate the ToU.

Moreover, even the most sophisticated automated processes and human review teams can and will make mistakes in identifying violations of a platform's ToU.[28] By allowing users to appeal content removal or account deactivation decisions, and by instituting processes for handling those appeals, service providers can help ensure that a mistake made during the content evaluation process does not permanently silence a user or a user's message or cause other undue harm.

Mechanisms that allow users to flag content should be coupled with mechanisms that allow the creators of content to defend the material that they post. Companies have legitimate concerns regarding their ability to devote resources to reviewing and responding to groundless appeals, especially given the sheer volume of UGC on some platforms. However, a system in which it is easy for users to contact companies to report alleged violations but difficult or impossible for users to respond to those allegations will naturally see abuse.

---

[24] YouTube, "YouTube Help Center page – How do we implement YouTube's content policies?," http://www.google.com/support/youtube/bin/answer.py?answer=92486.

[25] *Ibid.*

[26] As of November 2010, the amount of video uploaded to YouTube grew to 35 hours per minute. *See* You Tube, "Great Scott! Over 35 Hours of Video Uploaded Every Minute to YouTube," *The YouTube Blog*, http://youtube-global.blogspot.com/2010/11/great-scott-over-35-hours-of-video.html. As of January 2011, Facebook has more than 600 million users worldwide. *See* Nicholas Carlson, "Facebook Has More Than 600 Million Users, Goldman Tells Clients," *Business Insider*, January 5, 2011, http://www.businessinsider.com/facebook-has-more-than-600-million-users-goldman-tells-clients-2011-1.

[27] For a discussion of how Tunisian activists were targeted on Facebook by users who conducted content flagging campaigns, see Rebecca MacKinnon, "More problems in Facebookistan," *RConversation*, May 29, 2010, http://rconversation.blogs.com/rconversation/2010/05/more-problems-in-facebookistan.html. For examples of how takedown processes have been abused to silence political campaigners in the U.S., see Center for Democracy & Technology, "Campaign Takedown Troubles: How Meritless Copyright Claims Threaten Online Political Speech," September 2010, http://www.cdt.org/files/pdfs/copyright_takedowns.pdf.

[28] *See, e.g.,* Jillian York, "Would Anonymity Help Activists on Facebook? A Response to Luke Allnutt," *Jilliancyork.com*, February 24, 2011, http://jilliancyork.com/2011/02/24/would-anonymity-help-activists-on-facebook-a-response-to-luke-allnutt; Miguel Helft, "Art School Runs Afoul of Facebook's Nudity Police," *New York Times Bits Blog*, February 18, 2011, http://bits.blogs.nytimes.com/2011/02/18/art-school-runs-afoul-of-facebooks-nudity-police.

Develop an appeals process to mitigate the impact of mistakes or abuse by third parties
Because review processes are imperfect, companies should avoid absolute processes—a "decision is final" approach—that discourage users from legitimately appealing an account deactivation or a decision to remove content. Processes for appealing allegations of violations should be clear, transparent, and easily accessible. Platforms could, for example, enable users to re-post certain content the first time it is flagged if the user believes the adverse action is in error.

YouTube, for example, allows a user to appeal community flags on her video (and resulting "strikes" on her account). If YouTube agrees that the video did not violate the Community Guidelines, YouTube reinstates the video. If YouTube denies the appeal, the user is not allowed to appeal another flagged video for sixty days, which discourages meritless appeals.[29] Users should be empowered to appeal not only content removal and account deactivations, but also warnings issued by the service provider and restrictions on profile functionality that are part of a service provider's strategy of graduated response or escalation.

Respond to appeals promptly and communicate to users why the action was taken and what they should expect from the appeals process
Prompt responses to appeals should explain why the content was removed or the account was deactivated (or the warning issued) and, if the appeal was denied, the reason for which it was denied. As part of an appeals process, users may be asked to submit additional relevant information. In these instances, service providers should comply with the principle of data minimization: they should not ask for more personal information than is necessary to determine whether the content or account should be reinstated, and they should not store submitted information longer than is necessary for resolving the appeal. Companies will handle different volumes of complaints and therefore will necessarily respond to appeals with different degrees of promptness. All companies should, however, communicate to users the reasonable, approximate, expected duration of their appeals process.

Provide users with options to preserve and export data upon the deactivation of accounts
In cases of account deactivation, service providers should preserve data and content the user herself has uploaded to her account—and lists of associated "friends" or "followers" on the account—for the duration of the appeals process in case the user's appeal is successful and the account is restored.[30] Such harm mitigation strategies can be particularly important for advocates and activists who rely on social media for outreach and campaigns. However, even in cases in which an appeal is unsuccessful, service providers should also consider allowing the impacted user to export the content associated with the deactivated account. In some cases, service providers may wish to enable users to leave a "forwarding address" or final communication with their followers so they do not lose the network they have built on the platform.

---

[29] YouTube, "YouTube Help – Appealing your video strike,"
http://www.google.com/support/youtube/bin/answer.py?answer=185111.

[30] This situation (where a service provider is disabling an account for ToU violations) differs from the situation where the user herself wishes to delete her profile or account. In the latter context, users may have some expectation that the service provider will not unreasonably retain profile information or her UGC.

D. Embed Human Rights Considerations into Company Practice and Platform Design

While the risk of abuse related to ToU enforcement mechanisms is present whenever users express controversial views on UGC platforms, the impact of such abuse can be especially pernicious for human rights defenders, activists, and those engaged in political speech or democratic reform efforts. The popular press has amply documented the increasing use of social media by human rights advocates and activists of all creeds. The human rights community has also analyzed efforts to manipulate ToU enforcement mechanisms—sometimes by governments—to silence or shut down these uses.[31] In other cases, the content in question may violate the letter of the ToU or community guidelines but may serve an important, legitimate public good or human rights end.

When an activist's account is shut down or content is removed, it can cripple an ongoing campaign, remove a vital channel for disseminating information, or wipe out online networks. It is also important to note that there are many cases in which the line between being a 'typical user' and an activist becomes blurred or evolves in response to changing political, social, or other circumstances. Many users—who have long used social media to share content, connect with friends, develop a public profile, etc.—may not identify as activists but may be unknowingly engaging in sensitive expressive activities that require the adoption of new privacy, security, and content-related practices.

This section is therefore focused on strategies that companies can employ to embed human rights concerns into their processes and practices, with a particular focus on mechanisms for engaging and protecting activists and other communities that may face heightened risks and costs in the wake of ToU enforcement, account deactivation, and content removal practices.

Consider temporal, political, geographic, and other contexts when evaluating content
Platforms should evaluate posted content in the context of current events or emerging activities that may require special attention. In an interview with BeetTV, for example, Olivia Ma, Manager of News at YouTube, discussed the challenges of curating video coming out of recent protests in Egypt, Libya, Syria, Yemen, some of which were violent in nature.[32] Recognizing the potential value and importance of such media, YouTube worked closely with knowledgeable community members to ensure that content posted by activists in these regions was categorized as "educational, documentary or scientific in nature" and therefore not removed.[33]

In seeking solutions to the particular challenges faced by activists, some commentators have suggested encouraging widely used social media sites—such as Facebook, Flickr, and YouTube—to create a dedicated area of their platform for human rights-oriented work. Such proposals, however, may generate additional risks for activists by isolating them or making them easier for governments to identify and censor. Such visibility may also make it more dangerous

---

[31] *See* Section IV. Additional Resources.

[32] Andy Plesser, "YouTube is Managing Graphic, Violent Videos from the Middle East with Community Help," *BEET.TV*, May 5, 2011, http://www.beet.tv/2011/05/youtubeviolence.html.

[33] YouTube, "YouTube Community Guidelines," http://www.youtube.com/t/community_guidelines.

for other users to link to or 'friend' activists, which can undermine the efficacy of their work. Activists often deliberately post their writing, photos, and videos to prominent popular sites in order to easily reach people who may not be engaged with or are less inclined to be actively seeking out advocacy materials; these sites are most likely to be visited by the audiences that activists are trying to reach.[34] Similar concerns arise with regard to proposals to create dedicated sites for human rights-oriented content. As Global Voices co-founder Ethan Zuckerman notes, access to such a space could be blocked by governments hostile to free speech more quickly and less noticeably than 'general purpose' platforms; these spaces would also prove a clearer target for other forms of attack.[35]

Provide relevant human rights education for staff, including review teams and platform managers

Key staff members involved in reviewing complaints should be educated about the potential for abuse of the complaint system and about how to respond to complaints related to politically sensitive topics. It may be appropriate for the focus of staff education to vary by language. For example, education may differ for staff that review complaints in Arabic and staff that review complaints in Chinese. Some companies may designate a small number of employees as point persons for complaints that raise human rights-related concerns.[36] These staff members could cultivate experience in pertinent cultures, languages, and political issues, while establishing channels for regular communication with relevant human rights communities. Platforms should endeavor to hire local staff from regions that receive the most (or most complex) politically sensitive complaints, as they may be able to bring important linguistic, cultural, and other contextual information to the process of evaluating content and behaviors on the platform.

Meetings, summits, and workshops convened by online activists and civil society groups also provide important opportunities for company staff members to better educate themselves about relevant issues. Active engagement by companies in international and multi-stakeholder dialogues, from the GNI to the Internet Rights & Principles Coalition,[37] also provide channels for reaching out to and consulting with the digital activist community and other groups of high-risk users.

---

[34] The experience of WITNESS, an international human rights organization that provides training and support for activists to use video in their advocacy, is instructive in this regard. WITNESS had originally hosted its own platform on which activists could post videos. However, the myriad challenges, technical and otherwise, of hosting video ultimately proved too much and in spring 2010 the organization stopped hosting content. *See* Yvette, "Update on the Hub and WITNESS' New Online Strategy," *Video for Change*, August 18, 2010, http://blog.witness.org/2010/08/update-on-the-hub-and-witness-new-online-strategy.

[35] *See, e.g.,* Ethan Zuckerman, Hal Roberts, Ryan McGrady, Jillian York & John Palfrey, "Distributed Denial of Service Attacks Against Independent Media and Human Rights Sites," Berkman Center for Internet & Society at Harvard University (December 2010), http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/2010_DDoS_Attacks_Human_Rights_and_Media.pdf. For additional information regarding the innovative use of common Web 2.0 tools by activists, *see* Ethan Zuckerman, "The Cute Cat Theory Talk at ETech," *...My Heart is in Accra*, March 3, 2008, http://www.ethanzuckerman.com/blog/2008/03/08/the-cute-cat-theory-talk-at-etech.

[36] Yahoo!'s Business & Human Rights Program provides one good model. *See* Yahoo!, "Business & Human Rights Program," http://www.yhumanrightsblog.com/blog/our-initiatives.

[37] Internet Governance Forum Dynamic Coalition on Internet Rights & Principles, http://internetrightsandprinciples.org/node/14.

Engage academic institutions, human rights groups, and local NGOs regarding emerging issues, local contexts, and other factors that may influence user activity and content and inform company decision-making

While the costs of developing a staff that has enough knowledge of political issues, movements, and disputes in all corners of the world (in all languages) may be prohibitive and unrealistic,[38] companies can join or cultivate networks via established communities of bloggers, online and offline human rights activists, university legal clinics and academic institutions, and like-minded companies to help them to keep abreast of political and other developments that may require extra consideration. Such a step may be especially useful for smaller platforms, as these networks and partners can also assist in employee education and provide real-time expertise and strategic advice in response to emerging events. They can also proactively identify for companies periods when risk of abuse of community flagging systems may increase—for example, if there is an upcoming sensitive political event or campaign launch. Multi-stakeholder coalitions like the GNI offer one model for such an approach.[39]

Companies might also consider constituting advisory groups composed of individuals with strong connections to the human rights community. Such groups can help mediate and resolve issues between activists and companies regarding account deactivation and content removal. They might also assist with problems of authentication and verification of user claims and act as trusted sources of information. While there are compelling arguments both for and against such a system—around security, practical application, and degree of formality—the idea may support the overarching objective of strengthening lines of communication and engagement between the human rights community and companies.

Companies should also find ways to engage with at-risk user communities in advance of developing and deploying new settings, features, and products in order to ensure that there are not unintended consequences for users. Product designers should develop use cases that anticipate the needs of such at-risk users.

Incorporate human rights considerations into both the policy development process and ongoing evaluations of existing policies

At an early phase of development, service providers should evaluate the human rights impact of different features and policies related to their products and services. For example, platforms should carefully consider whether real-name policies are necessary; such policies can be especially problematic for activists and a range of other users who engage in sensitive, expressive activities.[40] Transparency can also go a long way toward making a platform more

---

[38] Rebecca MacKinnon, "More problems in Facebookistan," *RConversation*, May 29, 2010, http://rconversation.blogs.com/rconversation/2010/05/more-problems-in-facebookistan.html.

[39] In June 2011, the GNI announced the launch of its Global Advisory Council, which is designed to deepen understanding of emerging issues in different contexts and jurisdictions by reaching out to and engaging partners and other stakeholders around the world. *See* GNI, "Global Network Initiative Launches Advisory Council," June 10, 2011, http://www.globalnetworkinitiative.org/newsandevents/Global_Network_Initiative_Launches_Advisory_Council.php.

[40] *See, e.g.,* "Who is Harmed by a 'Real Names' Policy?," *Geek Feminism Wiki*, http://geekfeminism.wikia.com/wiki/Who_is_harmed_by_a_%22Real_Names%22_policy%3F (last visited July 28, 2011).

user-friendly for activists. For example, when Google+ launched, users who activated the chat feature were notified that their email address might be visible through chat. This notification gave users who might be sensitive about sharing their email address the information they needed to decide whether or not to use the feature.

Useful resources exist for helping platforms design programs and policies that are consonant with human rights norms and practices. Service providers should consult the Global Network Initiative's Principles on Free Expression and Privacy;[41] the Protect, Respect, and Remedy framework created by John Ruggie, the United Nations Special Representative of the Secretary General on transnational corporations and other business enterprises;[42] and the OECD's "Guidelines for Multinational Enterprises—Recommendations for Responsible Business Conduct in a Global Context."[43]

Develop specific guidance for human rights activists regarding good practice and relevant rules

Many commentators have suggested developing guides specifically geared towards human rights activists and other users whose activities and affiliations may put them at risk for account suspension and content takedown.[44] Such topical guides could be based on how popular platforms are presently used by self-identified advocates but could also be expanded to reach, and be useful to, those users who may not identify as activists but are nonetheless using the platform to share information and content related to specific, potentially controversial, activities, such as reporting police abuses. Many platforms already offer such guides, but they tend not to include specific advice about how these classes of users can navigate and avoid violating the ToU.

E.   Recommendations for Users

Users, like companies, have a clear role to play in protecting their content from removal and their accounts from being deactivated. They should proactively educate themselves regarding the UGC platforms they choose to engage with and should familiarize themselves with the ToU and other guidelines.

---

[41] The Global Network Initiative is a multi-stakeholder group of companies, civil society organizations (including human rights and press freedom groups), investors, and academics that has created a collaborative approach to protect and advance freedom of expression and privacy in the ICT sector. *See* GNI, "Principles on Freedom of Expression and Privacy," http://www.globalnetworkinitiative.org/principles/index.php.

[42] *See* John Ruggie, "Protect, Respect and Remedy: a Framework for Business and Human Rights," Human Rights Council, A/HRC/8/5, April 7, 2008, http://www.reports-and-materials.org/Ruggie-report-7-Apr-2008.pdf.

[43] *See* OECD, "OECD's Guidelines for Multinational Enterprises – Recommendations for Responsible Business Conduct in a Global Context," May 25, 2011, http://www.oecd.org/dataoecd/43/29/48004323.pdf.

[44] *See, e.g.,* Rebecca MacKinnon, "More problems in Facebookistan," *RConversation*, May 29, 2010. http://rconversation.blogs.com/rconversation/2010/05/more-problems-in-facebookistan.html.

Educate yourself about the platforms that you engage with and use these services in responsible manner

Before deciding to post content to a particular social networking or other UGC platform, users should seek to determine whether it is the right service for their intended purpose(s). They should browse the site to get a good understanding of its functionality, the amount of control that users seem to have over their content, and the intended audiences for the site.[45] They should also familiarize themselves with the basic contours of the ToU and with associated processes and guidelines for user behavior. For example, users should carefully review the platform's name policy—are they required to register under their real names? They should also take care not to engage in activity that might be considered "spamming" behavior, such as messaging or connecting with excessively large numbers of users at once. Users should pay close and proactive attention to the system the platform utilizes to identify ToU violations, enforce platform rules, or address abuse of community flagging. Where possible, they should educate themselves about how moderation, takedown, and abuse-prevention mechanisms work.

Finally, just as UGC platforms have a responsibility to take seriously their roles as mediators of expressive activities, users have a responsibility to act as good citizens in their uses of these platforms. Where a user is in clear violation of a service provider's ToU, she should not appeal a content removal or account deactivation decision, even if she may find certain provisions in the company's ToU objectionable. Frivolous appeals divert resources from legitimate appeals.

Use labels, tags, and other cues to provide contextual and other relevant information regarding your content

Users should provide contextual cues—in the titles, descriptions, labels or other information surrounding potentially controversial content—that indicate that their material has been posted for educational (or awareness raising) purposes. For example, a user posting a video of police brutality should label it as such so that content reviewers do not mistakenly assume it was posted to be entertaining, gratuitous violence. By providing context for their postings, users can alert human reviewers that content that may initially appear to be prohibited (gratuitous violence, for example), may in fact be of a nature that does not violate the platform's ToU or that serves a purpose that requires additional consideration or attention. If their work uses a less common language, where possible, users should seek to provide such cues and information in English or another major language.

Where possible, keep copies of your materials or upload them to another service as a back-up

Those who use online platforms as the primary mode of communicating with their contacts, cultivating community, or maintaining important ties should create copies of critical materials. Even in cases where they are working within the guidelines put forth in the ToU of a particular platform, users should back up important information to a secure location. This will mitigate the effects of any account deactivation or removal of content by platforms.

---

[45] *See* Christopher J. Reese, "Best Practice Recommendations for Social Networking Sites," *American Society of Media Photographers*, http://asmp.org/articles/social-media-terms-service.html.

Participate in networks, dialogues, and other efforts aimed at deepening understanding of emerging issues that may impact user content, rights, and activities

Like companies, users can form and participate in efforts and networks designed to facilitate information-sharing and engagement between diverse stakeholders interested in mitigating risks related to account deactivation and content removal. Such communities can provide an opportunity to strengthen ties between users, companies, activists, academics, and other constituencies, and can also provide channels for dialogue when problems arise. Users have an important role to play in proactively communicating with platforms regarding emerging issues, events, or other considerations that may impact the risks that they face as they interact with the service in question. Participation in such networks can provide an important avenue for that exchange.

Conclusion

Activists and ordinary users are increasingly using a range of privately owned platforms to express themselves, highlight issues of public concern, and reach out to their communities. Tensions inevitably emerge where such uses do not fit within the parameters that have been put forth by privately owned companies and crafted to suit the company's legitimate business needs. As noted by Ebele Okobi-Harris, Director of Yahoo's Business and Human Rights program, there are no easy answers when "activists use tools and products that were not initially created for human rights aims, [as] activists are still subject to community rules."[46]

Such complex challenges require thoughtful and responsible behavior on the part of a variety of stakeholders, including companies, activists, human rights organizations, and the users themselves. Beyond some of the practical steps outlined in this report, many observers have emphasized the need for proactive engagement among companies, activists, and the human rights community in order to prevent the inadvertent suppression of speech on private platforms.[47] While not all issues (nor movements, uprisings, or other events) can be anticipated, such connections will ensure that there are mechanisms in place through which the public can reach someone within the company, give feedback, ask questions, and raise concerns in a timely manner. Proactive and transparent guidelines, coupled with pre-existing channels of communication, can help to mitigate unintended harm to human rights activists and other users.

Activists and the human rights community also have a role to play in drawing attention to their needs and issues and in ensuring that their community members are active in protecting themselves, making themselves aware of the rules, and taking conscious and informed steps to safeguard their security and safety on a given platform. As Global Voices co-founder Rebecca MacKinnon writes, "Anticipate problems and help solve them not only for yourself but for everybody else in the community. Act like a citizen. Not a passive 'user.'"[48]

---

[46] Ebele Okobi-Harris, "Thoughts on Flickr and human rights," *Yahoo! Business and Human Rights Program Blog*, March 15, 2011, http://www.yhumanrightsblog.com/blog/2011/03/15/thoughts-on-flickr-and-human-rights.

[47] Rebecca MacKinnon, "Human rights implications of content moderation and account suspension by companies," May 14, 2010, *RConversation*, http://rconversation.blogs.com/rconversation/2010/05/human-rights-implications.html.

[48] *Ibid.*

The complexity of these challenges should not be underestimated and it is inevitable that both users and companies will make mistakes. Creating practical, scalable, and globally applicable processes is no small task and requires innovative solutions and creative approaches that draw on technical, human, and community driven processes.

Additional Resources

News Articles

Sami Ben Gharbia, "Google has disabled the ability for Nawaat to upload new videos," *Samibengharbia.com*, originally posted on February 15, 2010, updated on February 17, 2010, http://samibengharbia.com/2010/02/15/google-has-disabled-the-ability-for-nawaat-to-upload-new-videos.

Tim Carmondia, "Google+ Punts on Kafkaesque Name Policy," *Wired*, August 12, 2011, http://www.wired.com/epicenter/2011/08/google-punts-names/.

Jacqui Cheng, "Facebook Takedown Followup: What Happened and What Facebook Needs to Fix," *Ars Technica*, April 30, 2011, http://arstechnica.com/tech-policy/news/2011/04/facebook-takedown-followup-what-happened-and-what-facebook-needs-to-fix.ars.

Mike Giglio, "Middle East Uprising: Facebook's Secret Role in Egypt," *The Daily Beast*, February 24, 2011, http://news.yahoo.com/s/dailybeast/20110225/ts_dailybeast/12602_middleeastuprisingfacebooksbackchanneldiplomacy_1.

Miguel Helft, "Art School Runs Afoul of Facebook's Nudity Police," *New York Times Bits Blog,* February 18, 2011, http://bits.blogs.nytimes.com/2011/02/18/art-school-runs-afoul-of-facebooks-nudity-police.

Anick Jesdanun**, "Rights Like Free Speech Don't Always Extend Online,"** July 7, 2008*, Associated Press*, available at http://www.usatoday.com/tech/products/2008-07-07-1933136783_x.htm?loc=interstitialskip.

Amy Lee, "Facebook Apologizes for Censoring Gay Kiss Photo," *Huffington Post*, April 19, 2011, http://www.huffingtonpost.com/2011/04/19/facebook-gay-kiss_n_850941.html.

Rebecca MacKinnon, "Human rights implications of content moderation and account suspension by companies," May 14, 2010, *RConversation*, http://rconversation.blogs.com/rconversation/2010/05/human-rights-implications.html.

Rebecca MacKinnon, "More problems in Facebookistan," *RConversation*, May 29, 2010, http://rconversation.blogs.com/rconversation/2010/05/more-problems-in-facebookistan.html.

Ebele Okobi-Harris, "Thoughts on Flickr and human rights," *Yahoo! Business and Human Rights Program Blog*, March 15, 2011, http://www.yhumanrightsblog.com/blog/2011/03/15/thoughts-on-flickr-and-human-rights.

Jennifer Preston, "Ethical Quandary for Social Sites," *New York Times*, March 17, 2011, http://www.nytimes.com/2011/03/28/business/media/28social.html.

Phillip Weiss, "'YouTube Silences a Boycott Video," *Mondoweiss*, April 6, 2011, http://mondoweiss.net/2011/04/youtube-silences-a-boycott-video.html.

Jillian York, "Would Anonymity Help Activists on Facebook? A Response to Luke Allnutt," *Jilliancyork.com*, February 24, 2011, http://jilliancyork.com/2011/02/24/would-anonymity-help-activists-on-facebook-a-response-to-luke-allnutt.

Jillian York, "YouTube Bans Tunisian Site Nawaat from Uploading Videos," *Global Voices*, Advocacy, February 16, 2010, http://advocacy.globalvoicesonline.org/2010/02/16/youtube-bans-tunisian-site-nawaat-from-uploading-videos/.

Reports, Writings, and Recommendations

Jillian C. York, Policing Content in the Quasi-Public Sphere," *Open Net Initiative*, http://opennet.net/policing-content-quasi-public-sphere.

Ethan Zuckerman, "Public Spaces, Private Infrastructure—Open Video Conference," . . .*My Heart's in Accra*, October 1, 2010, http://www.ethanzuckerman.com/blog/2010/10/01/public-spaces-private-infrastructure-open-video-conference.

Rebecca MacKinnon, "Human rights implications of content moderation and account suspension by companies," *RConversation*, May 14, 2010, http://rconversation.blogs.com/rconversation/2010/05/human-rights-implications.html.

Global Network Initiative, "Account Deactivation and Content Removal: Takeaways from Oct 28 call," November 16, 2010, http://www.globalnetworkinitiative.org/newsandevents/Account_Deactivation_and_Content_Removal.php.

Access, "Access Recommendations to the GNI," October 2010, https://www.accessnow.org/page/-/docs/Access_recommendations_to_GNI_Oct_2010_1.doc.