The Innocence of YouTube

It's time for Internet giants to explain when censorship is and isn't OK.

BY SUSAN BENECH, REBECCA MACKINNON | OCTOBER 5, 2012

In 2006 Egyptian human rights activist Wael Abbas posted a video online of police sodomizing a bus driver with a stick, leading to the rare prosecution of two officers. Later, Abbas’s YouTube account was suddenly suspended because he had violated YouTube’s guidelines banning "graphic or gratuitous violence." YouTube restored the account after human rights groups informed its parent company Google that Abbas’s posts were a virtual archive of Egyptian police brutality and an essential tool for reform. After the Abbas case, Google concluded that some graphic content is too valuable to be suppressed, even where it is most likely to offend.

More recently, the Innocence of Muslims video led Google to bend its rules in the other direction, temporarily blocking the video in Egypt and Libya "given the very sensitive situations in these two countries," according to a statement given to reporters, even though those governments had not requested censorship and it was not violent, graphic, or directly hateful enough to violate YouTube’s guidelines banning gratuitously violent images and hate speech. (The video has since been quietly unblocked in both countries.) From the beginning, Google kept the video up in most of the world — and denied a request from the White House to remove it completely, but blocked it in countries including India and Indonesia where it has been ruled illegal, in keeping with Google policy to abide by its own rules as well as national laws.
In the crush of events, Google’s decision was the best it could have done under the circumstances. Yet little of the rationale behind Google’s decisions has been offered directly to YouTube users. Google has made a laudable public commitment to free expression and does a good job of disclosing how it responds to government demands around the world. Given the Internet giant’s power to shape global public discourse, it should be equally transparent about its private governance of global speech.

Sovereigns of cyberspace such as Google, Facebook, and Twitter have no legislatures or courts, yet they are carrying out private worldwide speech “regulation” — sometimes in response to government demands, sometimes to enforce their own terms of service and guidelines. As they attempt this unprecedented feat in a dizzying variety of social and political climates, they will continue to face complex dilemmas. In trying to develop "community guidelines" for the world, they will be both tempted and pressured to bend them to fit other new circumstances. With so many audiences ready to riot, so many provocateurs looking for excuses, and so many channels of communication between them, the companies cannot be expected to prevent — or take responsibility for — violence provoked by content posted on their services. What they can do, however, is to make a major contribution to both freedom and civility by explaining their de facto "jurisprudence" as it develops, to the public.

While Google does put up explanatory messages to users when a video has been restricted by government demand, it did not have an appropriate message on file for its decision in Libya and Egypt. For the first 24 hours, Egyptian users were shown a message claiming that the video had been blocked "due to a legal complaint by court order." After that, a message simply said "This video cannot be accessed from your country." While the Innocence of Muslims was an exceptional case of ad hoc censorship that must remain very rare if YouTube is to remain a platform for freedom of expression, it is inevitable that similarly urgent and difficult cases will arise in future. When such rare decisions are escalated to high-level executives who must sign off on deviations to standard procedures – as was the case in this instance – the company should consider posting more customized messages. This would have helped to promote the values that Google was attempting to balance.
The message might have read, for example, "This video clip has been temporarily blocked in your country due to a violent emergency, although no content posted by anybody on YouTube ever justifies violence." In the rest of the world where the video remains unblocked, a notice could read, "Some may prefer not to watch this video due to its offensive content, although it does not violate YouTube’s community guidelines against hate speech and graphic or gratuitous violence." (There is a precedent for this: Google posted an apology and disclaimer in 2009 when Google image searches for Michelle Obama were turning up a racist caricature as the top result.) Explanations would set a good example of transparency and would contribute to debate about the contours of freedom of expression, especially in countries that are working out their own rules for regulating speech as they make transitions toward democracy, often in volatile contexts.

Google has already developed a robust set of policies and practices to handle government censorship demands, and to inform users about those demands, in its Transparency Report. The report lists requests from governments to block content, including but not limited to YouTube videos, and records Google's responses. It is fascinating reading, by turns amusing and sobering. In 2011 for example, Passport Canada asked Google to remove "a YouTube video of a Canadian citizen urinating on his passport and flushing it down the toilet. We did not comply with this request." Also last year, Google said it had "received a request from the UK’s Association of Police Officers to remove five user accounts that allegedly promoted terrorism. We terminated these accounts because they violated YouTube’s Community Guidelines, and as a result approximately 640 videos were removed."

Google could expand its Transparency Report to include the decisions it makes of its own accord, whether to remove or block controversial content — and other Internet companies could follow suit. Columbia University Law School’s Internet law guru Tim Wu has suggested setting up a community of experts, or of YouTube users, to give advice on tough questions of regulating speech online: a fine idea but decisions to block content must be made very quickly, before they become moot. In practice, Internet companies will probably continue to make their own decisions; outsiders and experts will continue to critique them after the fact.
Companies could invite discussion about their decisions, though, in online spaces connected to their transparency reports — and to the very content that has sparked widespread controversy. Given that this entails a new kind of community platform, and editorial functions that may be incompatible with these companies’ roles, discussions could be hosted by a neutral third party.

Such a third-party role has a precedent. Both Google and Twitter now link to the independent non-profit website Chilling Effects, which hosts copies of the "cease and desist" notices and other court orders requiring the companies to remove content from their services. A third-party website run by an international community of netizens who share core values of free expression, non-violence, and empathy between different cultures could curate, and even translate, both mainstream and social media responses to controversial content. These cross-cultural stewards could then moderate a global conversation about what sorts of responses are reasonable and appropriate from different cultural perspectives.

Such a space could become a salutary forum for international debate: a gym for the civil exercise of freedom of speech in a world that needs the practice.