



# Stand Up For Digital Rights

## Key Issues: Moderation and Removal of Content

### Policy Measures by Intermediaries

Among the major factors behind the success of the Internet has been the open, honest and freewheeling nature of online discourse. Internet users who are connecting from the comfort of their home, and through the (perceived) anonymity of being behind a computer or mobile screen, feel comfortable sharing opinions and accessing information that they otherwise might not, due to official censorship or fear of legal or social reprisals. There is a brutal, no-holds-barred honesty to online speech that can be liberating and refreshing. However, this sense of anonymity, and the fact that online communications generally feel more remote than face-to-face communication, can also encourage people's darker impulses. The Internet provides a seemingly bottomless well of humour, storytelling and political commentary, but it is also a prime vehicle for vitriol and threats, as well as for the distribution of illegal material such as child sexual abuse imagery.

This dichotomy puts private sector intermediaries in a difficult position. On the one hand, for many the free flow of information is their bread and butter. Internet users, predictably, dislike having their thoughts and ideas controlled and have grown used to the freedom of being able to say whatever they like. Private sector intermediaries, as a consequence, have been keen to burnish their image as open and unfiltered platforms. Dick Costolo, a former CEO of Twitter, once described the company as being "the free speech wing of the free speech party."<sup>1</sup> In a post to the site's users, Reddit's then-CEO Yishan Wong said:

We uphold the ideal of free speech on reddit as much as possible not because we are legally bound to, but because we believe that you – the user – has the right to choose between right and wrong, good and evil, and that it is *your responsibility* to do so. [emphasis in original]<sup>2</sup>

---

<sup>1</sup> Emma Barnett, "Twitter chief: We will protect our users from Government", The Telegraph, 18 October 2011. Available at: [www.telegraph.co.uk/technology/twitter/8833526/Twitter-chief-We-will-protect-our-users-from-Government.html](http://www.telegraph.co.uk/technology/twitter/8833526/Twitter-chief-We-will-protect-our-users-from-Government.html).

<sup>2</sup> "Every Man Is Responsible For His Own Soul", Reddit, 6 September 2014. Available at: [www.redditblog.com/2014/09/every-man-is-responsible-for-his-own.html](http://www.redditblog.com/2014/09/every-man-is-responsible-for-his-own.html).

At the same time, the growing influence of private sector intermediaries has placed them under increasing pressure to mitigate the less desirable aspects of online speech. This can include pressure from their own users, who may prefer an online experience which is free from abusive or offensive material. It is, in particular, no secret that the Internet can be an especially hostile place for women. On 24 September 2015, two prominent online figures, Anita Sarkeesian and Zoe Quinn, spoke at the United Nations about the threats and harassment they faced as part of 'GamerGate', a controversy over ethics in journalism related to video games that spiralled into a campaign of anger against prominent women in the industry.<sup>3</sup> Both women were subjected to thousands of explicit rape and death threats and their personal contact information was widely disseminated. There were also attempts to steal or manipulate their online identities.<sup>4</sup>

While the experience of Anita Sarkeesian and Zoe Quinn was extreme, due to the fact that they were the public faces of a major conversation about sexism, harassment is a routine part of life for many women online. Caroline Criado-Perez, an activist who successfully lobbied to have Jane Austen replace Charles Darwin on the face of a British banknote, was similarly targeted with threats of death and rape.<sup>5</sup> In October 2015, Mia Matsumiya, a musician and blogger, started an Instagram account profiling the over one thousand abusive or sexually explicit messages she had received online over the period of a decade.<sup>6</sup> It is worth noting that Ms. Matsumiya is not a particularly prominent online figure and there is no reason to believe her experience was particularly exceptional. Writers at Jezebel, a feminist blog, have complained about visitors repeatedly and systematically posting images of violent pornography in the comment sections which follow their articles, which their staff must then sort through manually.<sup>7</sup> Although it is arguably the most pervasive "civility" issue on the Internet, gender-based harassment is part of a broader problem. Reddit, for example, contains dozens of forums dedicated to racial abuse, holocaust denial, pictures of dead children and many other forms of highly offensive content.

---

<sup>3</sup> A good summary of how this happened can be found in Jay Hathaway, "What Is Gamergate, and Why? An Explainer for Non-Geeks", Gawker, 10 October 2014. Available at: [gawker.com/what-is-gamergate-and-why-an-explainer-for-non-geeks-1642909080](http://gawker.com/what-is-gamergate-and-why-an-explainer-for-non-geeks-1642909080).

<sup>4</sup> Jessica Valenti, "Anita Sarkeesian interview: 'The word "troll" feels too childish. This is abuse'", The Guardian, 29 August 2015. Available at: [www.theguardian.com/technology/2015/aug/29/anita-sarkeesian-gamergate-interview-jessica-valenti](http://www.theguardian.com/technology/2015/aug/29/anita-sarkeesian-gamergate-interview-jessica-valenti).

<sup>5</sup> See: Katie Roiphe, "The Bank of England wanted to put Jane Austen on a 10-pound note. Then all hell broke loose.", Slate, 6 August 2013, available at: [www.slate.com/articles/double\\_x/roiphe/2013/08/the\\_anger\\_over\\_jane\\_austen\\_on\\_a\\_10\\_pound\\_note\\_proves\\_people\\_can\\_rage\\_over.html](http://www.slate.com/articles/double_x/roiphe/2013/08/the_anger_over_jane_austen_on_a_10_pound_note_proves_people_can_rage_over.html); and "Two jailed for Twitter abuse of feminist campaigner", The Guardian, 24 January 2014, available at: [www.theguardian.com/uk-news/2014/jan/24/two-jailed-twitter-abuse-feminist-campaigner](http://www.theguardian.com/uk-news/2014/jan/24/two-jailed-twitter-abuse-feminist-campaigner).

<sup>6</sup> Her account is available at: [instagram.com/perv\\_magnet/](https://www.instagram.com/perv_magnet/).

<sup>7</sup> "We Have a Rape Gif Problem and Gawker Media Won't Do Anything About It", Jezebel, 11 August 2014. Available at: [jezebel.com/we-have-a-rape-gif-problem-and-gawker-media-wont-do-any-1619384265](http://jezebel.com/we-have-a-rape-gif-problem-and-gawker-media-wont-do-any-1619384265).

In response to these problems, there has in recent years been a trend towards more active content management by some major private sector intermediaries. However, this gives rise to tricky debates about when and how companies should intervene. It is conceptually easy to defend a laissez-faire approach, where companies only intervene when they are legally required to do so, on freedom of expression grounds. Once companies choose to go beyond that, the debate becomes far more tangled.

A good example of these challenges came in the aftermath of the murder of journalist James Foley in August 2014. Foley was killed by the Islamic State, which then attempted to disseminate propaganda footage of the murder online. Twitter and YouTube, the two main platforms being used to share the material, moved swiftly to try and remove it from their networks and block users who uploaded, shared or linked to it. This muscular reaction resulted in at least some collateral damage against users who merely discussed or commented on the video. For example, Zaid Benjamin, a journalist who posted analysis and still images from the video, but not the moment of Foley's death or links to the video itself, had his account temporarily blocked. He reported that he lost 30,000 followers as a result.<sup>8</sup>

Although no sensible observer would fault Twitter or YouTube for attempting to remove graphic footage of a murder being disseminated as propaganda for a violent extremist group, some expressed unease at platforms with such a high level of power and influence exercising what is effectively editorial control over content being shared by their users. As James Ball, a writer for *The Guardian*, put it:

Twitter, Facebook and Google have an astonishing, alarming degree of control over what information we can see or share, whether we're a media outlet or a regular user. We have handed them a huge degree of trust, which must be earned and re-earned on a regular basis.

If Twitter has decided to make editorial decisions, even on a limited basis, it is vital that its criteria are clearly and openly stated in advance, and that they are consistently and evenly applied.<sup>9</sup>

Journalist Glenn Greenwald echoed these sentiments:

[A]s a prudential matter, the private/public dichotomy is not as clean when it comes to tech giants that now control previously unthinkable amounts of global communications... These are far more than just ordinary private companies from whose services you can easily abstain if you dislike their policies. Their sheer vastness makes it extremely difficult, if not impossible, to avoid them... It's an

---

<sup>8</sup> Shane Harris, "Social Media Companies Scramble to Block Terrorist Video of Journalist's Murder", *Foreign Policy*, 19 August 2014. Available at: [foreignpolicy.com/2014/08/20/social-media-companies-scramble-to-block-terrorist-video-of-journalists-murder/](http://foreignpolicy.com/2014/08/20/social-media-companies-scramble-to-block-terrorist-video-of-journalists-murder/).

<sup>9</sup> James Ball, "Twitter: from free speech champion to selective censor?" *The Guardian*, 21 August 2014. Available at: [www.theguardian.com/technology/2014/aug/21/twitter-free-speech-champion-selective-censor?CMP=twg](http://www.theguardian.com/technology/2014/aug/21/twitter-free-speech-champion-selective-censor?CMP=twg).

imperfect analogy, but, given this extraordinary control over the means of global communication, Silicon Valley giants at this point are more akin to public utilities such as telephone companies than they are ordinary private companies when it comes to the dangers of suppressing ideas, groups and opinions. It's not hard to understand the dangers of allowing, say, AT&T or Verizon to decree that its phone lines may not be used by certain groups or to transmit certain ideas, and the dangers of allowing tech companies to do so are similar.<sup>10</sup>

Facebook, it is worth noting, has long taken a far more active approach than Twitter towards regulating content, in line with its "Community Standards".<sup>11</sup> Reddit has struggled with this issue for years. In 2012, a series of articles on the website Gawker drew attention to large forums (or "subreddits") devoted to sexualising underage girls. These were initially defended by the website on freedom of expression grounds, but later banned as attention snowballed into the mainstream media. In 2015, Reddit introduced a policy whereby particularly offensive subreddits would be quarantined, so that they would only be visible to users who explicitly opted in.<sup>12</sup> This represents a sort of half-way house where content is not entirely blocked but its dissemination is limited.

It is easy to see why this issue has become such a minefield for private sector intermediaries. Supporters of Ms. Criado-Perez contrasted Twitter's swift and energetic response to distribution of the Foley video with its refusal to take action against users who harassed and abused her.<sup>13</sup> Reddit's users compared the decision to prohibit sexualised images of minors with the website's continued hosting of a subreddit devoted to pictures of dead children.<sup>14</sup> Inevitably, when a list of quarantined subreddits was published, users found a vast volume of highly offensive content which had escaped the restrictions.<sup>15</sup> Even Apple, primarily a hardware maker, faced criticism over policies on what content it allows to be sold through its App Store. The company banned an app which tracked the number of deaths caused by drone strikes in Pakistan, Yemen and Somalia in real-time, claiming that it contained "excessively crude or objectionable content".<sup>16</sup>

---

<sup>10</sup> Glenn Greenwald, "Should Twitter, Facebook and Google Executives be the Arbiters of What We See and Read?", Intercept, 21 August 2014. Available at:

[firstlook.org/theintercept/2014/08/21/twitter-facebook-executives-arbiters-see-read](http://firstlook.org/theintercept/2014/08/21/twitter-facebook-executives-arbiters-see-read).

<sup>11</sup> Available at: [www.facebook.com/communitystandards](http://www.facebook.com/communitystandards).

<sup>12</sup> "Content Policy Update", Reddit, 5 August 2015. Available at:

[www.reddit.com/r/announcements/comments/3fx2au/content\\_policy\\_update/?limit=500](http://www.reddit.com/r/announcements/comments/3fx2au/content_policy_update/?limit=500).

<sup>13</sup> James Ball, "Twitter: from free speech champion to selective censor?" The Guardian, 21 August 2014. Available at: [www.theguardian.com/technology/2014/aug/21/twitter-free-speech-champion-selective-censor?CMP=tw\\_t\\_gu](http://www.theguardian.com/technology/2014/aug/21/twitter-free-speech-champion-selective-censor?CMP=tw_t_gu).

<sup>14</sup> "Why is it that r/jailbait was shut down, but not r/picsofdeadkids?", Reddit, 7 September 2012.

Available at:

[www.reddit.com/r/AskReddit/comments/zhd5d/why\\_is\\_it\\_that\\_rjailbait\\_was\\_shut\\_down\\_but\\_not/](http://www.reddit.com/r/AskReddit/comments/zhd5d/why_is_it_that_rjailbait_was_shut_down_but_not/).

<sup>15</sup> "Content Policy Update", Reddit, 5 August 2015. Available at:

[www.reddit.com/r/announcements/comments/3fx2au/content\\_policy\\_update/cttd2li](http://www.reddit.com/r/announcements/comments/3fx2au/content_policy_update/cttd2li).

<sup>16</sup> Stuart Dredge, "Apple removed drone-strike apps from App Store due to 'objectionable content'", The Guardian, 30 September 2015. Available at:

[www.theguardian.com/technology/2015/sep/30/apple-removing-drone-strikes-app](http://www.theguardian.com/technology/2015/sep/30/apple-removing-drone-strikes-app).

## Illegal Content

Although private sector intermediaries have considerable flexibility in terms of the material they classify as offensive or against the standards of their services, they have little control over what material is prohibited by law. However, there are significant differences in how private sector intermediaries decide to deal with content which is illegal or of questionable legality. Among the most important factors in determining this is whether, and under what circumstances, intermediaries are protected against liability for the content in relation to which they provide services. Many legal systems grant intermediaries some degree of immunity, although this can come with various conditions. For example, in the United States, private sector intermediaries are protected by section 230 of the Communications Decency Act<sup>17</sup> and section 512 of the Digital Millennium Copyright Act (DMCA).<sup>18</sup> However, the DMCA protections against liability for copyright infringement depend on private sector intermediaries' compliance with "notice and takedown" procedures designed to promote the expedited removal of infringing material.

Although legal rules on immunity from liability are a significant factor in guiding their behaviour, many intermediaries commit to or take actions which go significantly beyond the minimum requirements. This is particularly true in relation to combating the spread of child sexual abuse imagery, which is of course a particularly heinous social ill.

For example, the GNI Implementation Guidelines,

Acknowledge and recognize the importance of initiatives that seek to identify, prevent and limit access to illegal online activity such as child exploitation. The Principles and Implementation Guidelines do not seek to alter participants' involvement in such initiatives.<sup>19</sup>

Although the Guidelines broadly support measures to combat illegal activity, the specific reference to child exploitation should be seen in light of the fact that many intermediaries have demonstrated a willingness to take more intrusive action in this area. This is likely due to the fact that child sexual abuse is vastly more harmful than, say, copyright infringement, and because contextual considerations like fair use or fair dealing are far less relevant, making it easier to identify illegal content definitively.

Several major tech firms maintain databases of identifying markers (hashes) which automatically identify child sexual abuse imagery. This includes Microsoft's

---

<sup>17</sup> 47 U.S.C. § 230. Available at: [www.law.cornell.edu/uscode/text/47/230](http://www.law.cornell.edu/uscode/text/47/230).

<sup>18</sup> 17 U.S. Code § 512. Available at: [www.law.cornell.edu/uscode/text/17/512](http://www.law.cornell.edu/uscode/text/17/512).

<sup>19</sup> Available at: [globalnetworkinitiative.org/implementationguidelines/index.php](http://globalnetworkinitiative.org/implementationguidelines/index.php).

PhotoDNA technology, which has been in use since 2009.<sup>20</sup> The same system has been used by Facebook since 2011.<sup>21</sup> In 2014, a similar programme run by Google came to light after a tip off from the company to the authorities led to a conviction for child pornography in the United States.<sup>22</sup> Although this particular activity by Google attracted little controversy, some commentators expressed unease at the possibility that a similar approach might be used in other areas of law enforcement, leading to searches for broader incriminating phrases, such as “assassinate the president”.<sup>23</sup>

Some intermediaries also go beyond minimum legal requirements to combat hate speech. In particular, intermediaries often face significant pressure from governments to take a more proactive stance in situations where there is a risk of hate-sponsored violence. In Germany, in the wake of xenophobic attacks on refugee camps, the Justice Minister called on Facebook to do more to reign in abusive posts. In response, the company promised to work with the government to create a task force aimed at flagging and removing hateful content more quickly and to help finance organisations which track online speech.<sup>24</sup>

## Copyright

By far the most pervasive illegal content issue online is the use of the Internet to violate copyright rules. By making it vastly easier to copy, manipulate and share information, the digital age has led to an explosion in copyright infringement. Some have argued that the mass violation of copyright laws suggests that those laws are poorly adapted to the digital age, and badly in need of reform.<sup>25</sup> But the reaction of many States has been to expand copyright rules rather than to revise them to take digital realities into account.

---

<sup>20</sup> Anthony Salcito, "Microsoft donates PhotoDNA technology to make the Internet safer for kids", Microsoft Developer Blog, 17 December 2009. Available at: [blogs.msdn.microsoft.com/microsoftuseducation/2009/12/17/microsoft-donates-photodna-technology-to-make-the-internet-safer-for-kids/](https://blogs.msdn.microsoft.com/microsoftuseducation/2009/12/17/microsoft-donates-photodna-technology-to-make-the-internet-safer-for-kids/).

<sup>21</sup> Catharine Smith, "Facebook Adopts Microsoft PhotoDNA To Remove Child Pornography", Huffington Post, 20 July 2011. Available at: [www.huffingtonpost.com/2011/05/20/facebook-photodna-microsoft-child-pornography\\_n\\_864695.html](http://www.huffingtonpost.com/2011/05/20/facebook-photodna-microsoft-child-pornography_n_864695.html).

<sup>22</sup> James Vincent, "Google scans Gmail accounts for child abuse - and has already helped convict a man in the US", The Independent, 4 August 2014. Available at: [www.independent.co.uk/life-style/gadgets-and-tech/google-tips-off-us-police-to-man-storing-images-of-child-abuse-on-his-gmail-account-9647551.html](http://www.independent.co.uk/life-style/gadgets-and-tech/google-tips-off-us-police-to-man-storing-images-of-child-abuse-on-his-gmail-account-9647551.html).

<sup>23</sup> Jonathan Zittrain, "A Few Keystrokes Could Solve the Crime. Would You Press Enter?", Just Security, 12 January 2016. Available at: [www.justsecurity.org/28752/keystrokes-solve-crime-press-enter/](http://www.justsecurity.org/28752/keystrokes-solve-crime-press-enter/).

<sup>24</sup> Amar Toor, "Facebook will work with Germany to combat anti-refugee hate speech", The Verge, 15 September 2015. Available at: [www.theverge.com/2015/9/15/9329119/facebook-germany-hate-speech-xenophobia-migrant-refugee](http://www.theverge.com/2015/9/15/9329119/facebook-germany-hate-speech-xenophobia-migrant-refugee).

<sup>25</sup> Centre for Law and Democracy, "Reconceptualising Copyright: Adapting the Rules to Respect Freedom of Expression in the Digital Age", (Halifax: Centre for Law and Democracy, 2013). Available at: [www.law-democracy.org/live/wp-content/uploads/2013/07/Final-Copyright-Paper.pdf](http://www.law-democracy.org/live/wp-content/uploads/2013/07/Final-Copyright-Paper.pdf).

The pervasiveness of copyright infringement has led to the establishment of robust systems for identifying and removing infringing content. Despite this, there is little evidence that these systems have made a dent in the illegal spread of copyrighted material and infringement remains as ubiquitous as ever. At the same time, the systems put in place to address copyright have proven susceptible to abuse.

When Ashley Madison, a website that facilitates adultery, was hacked in 2015, resulting in the publication of sensitive user information, the company responded by sending out a barrage of copyright notifications under the DMCA to try to remove the material.<sup>26</sup> Although the Ashley Madison hack represented a serious invasion of the privacy of millions of individuals, this is unrelated to the purpose of the DMCA and the takedown requests were frivolous and clearly abusive. For example, targets which were successfully taken down included a website which allowed individuals to check whether their private information had been compromised, a critically important service in the aftermath of a major data breach.

### **Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE)**

Across Latin America, there are many examples of abusive uses of the DMCA system, particularly for political purposes. In Ecuador, President Rafael Correa has become notorious for this behaviour:

- On 9 October 2013, Ecuadorian filmmaker Pocho Alvarez discovered that one of his documentaries had been removed from his YouTube page due to alleged copyright infringement. The documentary in question, Assault on Intag, is a short exposition on the harassment suffered by the indigenous community for its resistance to mining activities in the region. It included less than 20 seconds of images of Ecuador's President Correa, including a short clip of his voice. The removal was based on a claim that Alvarez had violated copyright rules by using footage of President Correa taken from his weekly national broadcast. It is interesting to note that Correa filed the claim through a Spanish agency in the United States, rather than in his own country. Another documentary, by filmmaker James Villa, which criticised the Correa administration was also removed due to having used images from his weekly public address. These clearly fall into the scope of exceptions to copyright protection.
- In September 2014, a video depicting the violent repression of a student demonstration, which included apparent police abuses as well as depictions of President Correa praising the police's actions, was removed from Facebook and YouTube after a copyright complaint.
- A Twitter account belonging to Diana Amores was subject to several removal requests after she posted images of politicians with humorous taglines. The

---

<sup>26</sup> Adam Clark Estes, "Ashley Madison Is Sending Out Bogus DMCA Takedown Notices", Gizmodo, 20 August 2015. Available at: [gizmodo.com/ashley-madison-is-sending-bogus-dmca-takedown-notices-1725372969](http://gizmodo.com/ashley-madison-is-sending-bogus-dmca-takedown-notices-1725372969).

volume of complaints led to her account being suspended on multiple occasions. The complaints originated from EcuadorTV, the State-run TV station, and Movimiento Alianza País, the country's governing party.

Political abuse of the DMCA system is not limited to Ecuador:

- The Ministerial Church of Jesus Christ International, associated with the Colombian political party MIRA, has repeatedly sought the removal of YouTube videos that feature, for example, declarations made by the church's founder. One of the videos that YouTube blocked upon the church's request informed the viewer explicitly – in its title – that the video was a parody.
- In Brazil, the DMCA was used to remove critical videos of 2014 presidential candidate and former governor, Aécio Neves. Although the requester's identity has not been confirmed, many speculated that Neves himself was responsible for the takedowns.

The public interest is affected each time legitimate content is removed from the Internet. The public interest is engaged if the content removed can be legally sent or received according to intellectual property laws (such as content in the public domain, "fair use" or other copyright exceptions). In many cases, content is removed based on an incorrect balancing between copyright and freedom of expression. This is a serious imbalance because freedom of expression is a fundamental human right, while copyright is not.

In 2015, a hacker leaked an enormous trove of internal information from Hacking Team, a spyware and surveillance company, onto the Internet.<sup>27</sup> The leak included evidence that the company had sold their equipment to Sudan, potentially in breach of UN sanctions, as well as to intelligence agencies in Egypt, Ethiopia, Kazakhstan, Russia and Saudi Arabia, all States which are known to persecute journalists and opposition figures. The company's immediate response was to send out frivolous DMCA notifications in an attempt to stop the spread of the leaks.

The DMCA system was even used by the United States' National Association for Stock Car Racing (NASCAR) to try and remove footage of a major car crash at one of their events.<sup>28</sup> NASCAR defended its actions as a matter of respecting the privacy of those injured, again not the problem the DMCA was designed to address. From a human rights perspective, measures which can easily be expanded beyond their intended purpose, like the DMCA, are troubling since they are by definition overbroad, running counter to the cardinal principle, as spelled out in the

---

<sup>27</sup> Cory Doctorow, "Hacking Team leak: bogus copyright takedowns and mass DEA surveillance in Colombia", BoingBoing, 7 July 2015. Available at: [boingboing.net/2015/07/07/hacking-team-leak-bogus-copyp.html](http://boingboing.net/2015/07/07/hacking-team-leak-bogus-copyp.html).

<sup>28</sup> Mike Masnick, "NASCAR Abuses DMCA To Try To Delete Fan Videos Of Daytona Crash", Techdirt, 25 February 2013. Available at: [www.techdirt.com/articles/20130224/22411222089/nascar-abuses-dmca-to-try-to-delete-fan-videos-daytona-crash.shtml](http://www.techdirt.com/articles/20130224/22411222089/nascar-abuses-dmca-to-try-to-delete-fan-videos-daytona-crash.shtml).

*International Covenant on Civil and Political Rights (ICCPR)*,<sup>29</sup> that laws which restrict expression should be carefully and narrowly construed.

Furthermore, many private intermediaries go beyond what is legally required when dealing with potentially infringing content. The starkest example of this is in South Korea, where legal ambiguities and an eagerness to avoid liability have led to intermediaries complying with virtually every request they receive, resulting in a rate of removal that far exceeds that of other comparable countries.

### **Open Net Korea**

The Korea Communications Standards Commission (KCSC), the administrative body responsible for monitoring and restricting Internet content in Korea, generally attempts to remove information through the use of “non-binding” requests rather than formal takedown decisions. This avoids having to provide subjects with notice and a hearing or any other procedural safeguards. Although private sector intermediaries can refuse to comply with these requests, the compliance rate is effectively 100 percent, partly because South Korea has extremely weak protections against intermediary liability, incentivising intermediaries to comply with requests without questioning them.

No intermediary has ever challenged a KCSC decision in court. Although users can file objections, they rarely do since the intermediary, rather than the user, is notified of the takedown request. This is particularly problematic in light of the fact that in some cases the users, properly notified, would likely volunteer to remove just the offending material. Instead, takedowns are often vastly overbroad. For instance, an entire blog maintained by a 60-year-old man was shut down following a KCSC request because about one-third of 132 entries included content deemed to be supportive of North Korea, which is illegal under the National Security Act (which is a highly problematic document on its own). About half of the entries were photos of his grandchildren, pictures of his own paintings, music and singing files of his own composition, and cooking recipes, accumulated over 3-4 years late in the man’s life. Had he been notified, it is likely that he would have deleted the pro-North Korean statements in order to protect his other, legal content, or at least have backed-up the other content to prevent it from being lost.

Overall, South Korea’s system of content removal is extremely pervasive. In 2013, the KCSC ordered the blocking or deletion of 104,400 websites. By comparison, their counterpart in Australia, the Australian Communication and Media Authority (ACMA), only blocked about 500 websites in 2013.

The KCSC’s takedowns often target frivolous sites, or sites that criticise politicians. Government officials often make private takedown requests for postings that

---

<sup>29</sup> UN General Assembly Resolution 2200A(XXI), adopted 16 December 1966, in force 23 March 1976.

criticise their policy decisions. Some examples of this include:

- A posting criticising a Seoul City mayor's ban on assemblies in Seoul Square;
- A posting criticising a legislator's drinking habits and publicising his social media account;
- Clips of a television news report on the Seoul Police Chief's brother who allegedly runs an illegal brothel;
- A posting criticising politicians' pejorative remarks about the recent deaths of squatters and police officers in a redevelopment dispute;
- A posting calling for immunity for labour strikers from criminal prosecutions and civil damage suits;
- A posting by an opposition party legislator questioning a conservative media executive's involvement in a sex exploitation scandal related to an actress and her suicide; and
- A Twitter account titled 2MB18NOMA was blocked because the phonetic name of the account resembles an epithet against the then-President Lee Myung-Bak.

Although the DMCA offers private online intermediaries greater protection from liability that they have under South Korean law, it nonetheless heavily incentivises over-compliance, since protection is predicated upon their promptly removing content upon receiving notice from the rights holder. Consequently, some intermediaries have been criticised for failing to stand up for their users in the face of frivolous DMCA takedown requests, or their failure to investigate whether a complaint is meritorious, or engage with users after a complaint has been filed.

YouTube's ContentID system, which is another voluntary mechanism, automates the process of flagging and removing allegedly infringing content.<sup>30</sup> This can lead to mistakes. For example, the system has repeatedly flagged footage posted by the National Aeronautics and Space Administration (NASA), despite the fact that, like all United States government agencies, its content is in the public domain.<sup>31</sup> There have also been reports of users having original material which they created flagged.<sup>32</sup> In addition to these mistakes, the automation of the system means that it is unable to take into account possible defences to copyright infringement, such as fair use.

## Centre for Internet and Society

<sup>30</sup> A brief explanation of how the system works is available at:

[www.youtube.com/watch?v=9g2U12SsRns#t=33](http://www.youtube.com/watch?v=9g2U12SsRns#t=33).

<sup>31</sup> Mike Masnick, "Curiosity's Mars Landing Video Disappears From YouTube Due To Bogus Copyright Claim", Techdirt, 6 August 2012. Available at:

[www.techdirt.com/articles/20120806/11053019945/curiositys-mars-landing-video-disappears-youtube-due-to-bogus-copyright-claim.shtml](http://www.techdirt.com/articles/20120806/11053019945/curiositys-mars-landing-video-disappears-youtube-due-to-bogus-copyright-claim.shtml).

<sup>32</sup> Erik Kain, "YouTube Responds To Content ID Crackdown, Plot Thickens", Forbes, 17 December 2013. Available at: [www.forbes.com/sites/erikkain/2013/12/17/youtube-responds-to-content-id-crackdown-plot-thickens/#339f50001086](http://www.forbes.com/sites/erikkain/2013/12/17/youtube-responds-to-content-id-crackdown-plot-thickens/#339f50001086).

ISPs in India often respond to takedown requests by removing far more material than is required. One solution to this is for courts to be more specific in their orders, but ISPs also need to take a stronger stand in favour of freedom of expression and interpret these orders as narrowly as possible.

Only one of the companies we examined, SingTel, provided their users with notice when material had been removed on copyright grounds. None of the companies we examined provided a specific redress mechanism to individuals whose material was wrongfully removed.

Internet access providers in the United States have also agreed to participate in voluntary schemes aimed at combating copyright infringement, most notably the Copyright Alert System (CAS), otherwise known as “Six Strikes”.<sup>33</sup> This system, which was launched in February 2013, allows for escalating responses to instances of copyright infringement beginning with “educational” alerts and escalating to more intrusive measures, including penalties. The specific enforcement measures vary among access providers, and there is a lack of consistency, or transparency, as to how users may be impacted. For example, Verizon has stated that, on the fifth alert, users’ Internet access speed will be throttled to 256kbps for a period of two days.<sup>34</sup> Optimum Online, another Internet access provider, states that upon receiving an alert it “may temporarily suspend your Internet access for a set period of time, or until you contact Optimum.”<sup>35</sup> It is worth noting that a 2011 Report by the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression states:

The Special Rapporteur considers cutting off users from Internet access, regardless of the justification provided, including on the grounds of violating intellectual property rights law, to be disproportionate and thus a violation of article 19, paragraph 3, of the International Covenant on Civil and Political Rights.

The Special Rapporteur calls upon all States to ensure that Internet access is maintained at all times, including during times of political unrest. In particular, the Special Rapporteur urges States to repeal or amend existing intellectual copyright laws which permit users to be disconnected from Internet access, and to refrain from adopting such laws.<sup>36</sup>

---

<sup>33</sup> Center for Copyright Information, “The Copyright Alert System”. Available at: [www.copyrightinformation.org/the-copyright-alert-system/](http://www.copyrightinformation.org/the-copyright-alert-system/).

<sup>34</sup> Verizon, “Copyrights and Verizon's Copyright Alert Program”. Available at: [www.verizon.com/support/consumer/account-and-billing/copyright-alert-program-faqs#04FAQ](http://www.verizon.com/support/consumer/account-and-billing/copyright-alert-program-faqs#04FAQ).

<sup>35</sup> Optimum, “Copyright Infringement Alerts”. Available at: [optimum.custhelp.com/app/answers/detail/a\\_id/3592](http://optimum.custhelp.com/app/answers/detail/a_id/3592).

<sup>36</sup> Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, 16 May 2011, A/HRC/17/27, paras. 78 and 79. Available at: [www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27\\_en.pdf](http://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf).

## **Centro de Estudios en Libertad de Expresión y Acceso a la Información (CELE)**

Widespread misuse of the system suggests that, before any claimant completes the form to report an alleged infringement, they should be presented with instructions explaining:

- a. The conditions under which a copyright claim will be legitimate.
- b. The difference between being a copyright holder and the right to one's image.
- c. What constitutes abuse of the DMCA, as well as the possible sanctions for this abuse. Private sector intermediaries should make it clear that users who repeatedly file abusive complaints may also be subject to penalties, such as the cancellation of their accounts.
- d. A list of exceptions to copyright, as explained according to local legal standards.



# Stand Up For Digital Rights

## **Recommendations for Moderation and Removal of Content:**

### *Clarity and Communication*

- **Intermediaries should post, in a prominent place, clear, thorough and easy to understand guides to their policies and practices for taking action in relation to content, including detailed information about how they are enforced. Where policies need to be complex due to the fact that they form the basis of a legal contract with users, they should be accompanied by clear, concise and easy to understand summaries or explanatory guides.**
- **Intermediaries' copyright reporting mechanisms should provide information to both complainants and users about limitations and exceptions to copyright and, where applicable, warn complainants about the potential consequences of filing false claims.**
- **Policies to address problematic content (such as deletion or moderation) which go beyond formal legal requirements should be based on clear, pre-determined policies which can be justified by reference to a standard which is based on objective criteria (such as a family friendly service) which are set out in the policy, and which is not based on ideological or political goals. Where possible, intermediaries should consult with their users when determining such policies.**

### *Process for Receiving and Adjudicating Complaints*

- **Third parties who file a complaint about inappropriate or illegal content should be required to indicate what legal or policy rule the content allegedly violates.**
- **Intermediaries should be consistent in applying any content moderation policies or legal rules and should scrutinise claims under such policies or rules carefully before applying any measures. They should have in place processes to track abuses of their content moderation systems and should apply more careful scrutiny to claims from users who repeatedly file frivolous or abusive claims.**

- **Intermediaries should, subject only to legal or technical constraints, notify users promptly when content which the latter created, uploaded or hosts is subject to a complaint or restriction. The notification should include a reference to the legal or policy rule in question, and an explanation of the procedure being applied, the opportunities available to the user to provide input before a decision is taken, and common defences to the application of the procedure.**
- **Where action is proposed to be taken in relation to content a user has created, uploaded or hosts, that user should normally be given an opportunity to contest that action. Where possible, subject to reasonable resource and technical constraints, users should be given a right to appeal against any decision to take action against the content at issue.**

### *Restricting Content*

- **Actions to remove or otherwise restrict third party content should be as targeted as possible and should only apply to the specific content which offends against the relevant legal or policy standard.**
- **Intermediaries should consider whether less intrusive measures are available which provide protection against harmful content without necessarily taking that content down, such as providing for opt-ins to access the content.**
- **Where action is taken against content, the intermediary should, subject to reasonable technical constraints, retain the means to reverse that action for as long as any appeal against the action, including any legal appeal, remains pending.**
- **Where a user's account is deleted or de-activated, users should be given an option to preserve and export the data from that account, unless the material is patently illegal (i.e. in the case of child sexual abuse imagery) or has been declared to be illegal by a clear and binding legal order.**