Introduction

The Coalition for Creativity (C4C) – more details below – welcomes the consultation by the U.S. Copyright Office on technical measures to identify or protect copyrighted works online (hereafter 'the consultation'). For further questions on our submission, please do not hesitate to contact Caroline De Cock, C4C's Coordinator, at secretariat@coalition4creativity.org.

About C4C

C4C is a broad-based coalition that seeks an informed debate on how copyright can more effectively promote innovation, access, and creativity. C4C brings together libraries, scientific and research institutions, digital rights groups, technology businesses, and educational and cultural heritage institutions that share a common view on copyright. See our members here: https://coalition4creativity.org/about-us/.

Preamble

As an organisation that has its core of activities and members based in the European Union, we have been heavily involved in the debate on the EU Directive on Copyright in the Digital Single Market. One of our main concerns is and has been the worrying push for (mandatory) filtering measures. By providing various examples of filter fails we want to depict why the adoption and broad implementation of such measures have negative effects (see questions 3 and 6).

Our recommendations

In December 2020, the Electronic Frontier Foundation (EFF) issued a report[1] depicting three case studies of long-time video creators' experiences with Content ID, filters, and YouTube as a platform. Herein, Lindsay Ellis, a New York Times best-selling author, film critic, and video essayist with over one million YouTube subscribers, and her channel moderator, Elisa Hansen, described their experience with YouTube as being a place where "the only thing that matters is are you smarter than a robot."

"By shifting the neutral presumption accompanied with fair use against the uploader, Content ID effectively overrides judicial precedent," scholar Taylor B. Bartholomew observed in 2015.

Cory Doctorow, author, journalist, and activist, summarised the harsh reality of online content moderation, especially in the copyright realm: removed in 10 seconds, reinstated after 53 months.

Our overall recommendation is simple: the uptake of technical measures to identify or protect copyrighted works online should not be mandated, nor should filtering become the de facto modus operandi online through 'voluntary' agreements.

The evidence outlined below shows the negative effects faced by users, citizens and creators. These detrimental effects are reinforced when users are confronted with algorithmic failure combined with 3-strikes policies on platforms.

Moreover, growing calls from the copyright industry for stay-down ignore the fact that a piece of content can be used in multiple ways and that the copyright status of content has an expiry date, all contextual elements which algorithms cannot comprehend.

[1] 'Unfiltered: How YouTube's Content ID discourages fair use and dictates what we see online'

Instead, we believe that close attention should be paid to how technical measures are deployed, to ensure that a framework is developed for the use of these measures that guarantees that users' fundamental rights as well as their right to due process are respected, while ensuring that:

1. Rightholders are actively involved in redress procedures and are the ones taking responsibility. They must be obliged to respond to counter notices in a detailed manner and in due time. It has been noted that the first 24h to 48 hours after a video has been uploaded are the most crucial, especially for content creators who monetise their efforts or for organisations who use these tools to campaign on certain political or societal issues.
2. Rightholders are actually penalised for bad faith or negligent inaccurate copyright claims. These penalties should especially target repeat offenders, who abuse their privileges to stifle speech and/or to unfairly appropriate themselves the monetisation revenues for content.
3. Safeguards are put in place against active copyright trolling practises, whereby malicious users game the system, in order to extort benefits, be it financial or anti-competitive, through false copyright claims.

Our argumentation

Main issues and experiences

Content moderation practises, both with human intervention and through automated systems, are known to generate false positives (content wrongfully removed) and false negatives (illegal content not identified as such).

Rightholders are in arms about the latter, while the former, which represent a threat to fundamental rights of users, seem to bother those same rightholders much less.

Various journalistic investigations have exposed the shortcomings of human intervention in the content moderation process.[2]

Looking at automated systems, most filter mechanisms admit a 2 to 5% rate of false positives, implying that hundreds of thousands of uploads have been blocked incorrectly on some platforms.

Organisations such as the Lumen Database have been documenting the origin of false positives.

The reasons behind erroneous removals can be categorised as follows:

1. Glitches of the system: Content moderation tools are known to have blocked white noise, a 2012 video of the uploader gardening in the wilderness, with no accompanying music or inserted audio, a sound check of a microphone on a private YouTube livestream, a purring cat, and most recently numbers on Google Drive.

   The glitches can also appear on the side of the rightholder, as illustrated by a January 2020 report that Studio 71, part of German media company ProSiebenSat.1 Media, wrongfully issued copyright claims on YouTube videos due to a 'technical problem.'

[2] Understaffed, underpaid and/or insufficiently trained people (often outsourced) need to make crucial decisions affecting freedom of speech and other fundamental freedoms within a very short timeframe, decisions that take expert courts months to years of analysis to reach. Moreover, they are often guided in their work by very summary and often puzzling 'training manuals' setting out content moderation guidelines that do not necessarily meet the standards of the rule of law in all countries.

2. Non-recognition of a legitimate use of copyrighted work/incorrect information provided by rightholders to content recognition tools (CRTs)

   Non-recognition of a legitimate use happens more often with automated systems, as they are incapable of assessing context and hence of identifying if a user uploaded content benefitting from fair use. It is also incapable of identifying if a use is incidental or inadvertent, as is the case for example with background music during public manifestations.

   These limitations are exacerbated by the practises of some big music and audiovisual rightholders which feed reference files to CRTs. These CRTs then automatically block new uploads of the content contained in those reference files, and also remove existing uploads. As a result, legitimate uses of these audio or video files end up being blocked by platforms.

   Examples include:

   - Educational material: Works inserted for educational purposes in online courses get blocked.

     A March 2020 case actually demonstrates with a certain irony as the blocking relates to a course by the NYU School of Law explaining the inner workings of copyright law. The explanation "over how and when one song can infringe the copyright of another and how exactly one proves in a courtroom if the accused song is 'substantially similar' enough to be deemed illegal" was blocked on YouTube and not reinstated even after a counter-notice. The same thing happened in 2016 to a Harvard Law presentation including snippets of Jimi Hendrix.

     In 2018, a German music professor received a Content ID claim for a short video he uploaded about a project to digitise materials that are unequivocally in the public domain. In this video, he explained his project, while examples of the public domain music (e.g. Beethoven) played in the background.

   - Public domain: Videos of the Mars Curiosity Rover, the recent SpaceX launch, a video celebrating a 1923 song finally entering the public domain, and musicians playing Mozart or Bach have all been taken down because of mistakes made by algorithmic copyright enforcement.

     In a similar vein, a speech by French President Macron broadcasted on national television and on the French Government's official channel 'Palais de l'Élysée' was blocked on YouTube as a result of a claim by French broadcaster Canal+ in April 2020.

     In the context of classical music being taken down, The Washington Post spoke, in May 2020, with Duncan Hammons, Senior Manager of video and new media at Naxos – representing 2.5 mln tracks, who are relying on the content recognition tools of YouTube and Facebook to protect their labels' content. Hammons noted that:

     > "Though the technology works most of the time in terms of correctly identifying instances of our clients' content on-platform, it still generates a not-insignificant amount of mismatches that require human review to differentiate[.] The chances of conflicts with this amount of content are considerable. For these reasons there is always a volume of potentially erroneous auto-generated claims that unless contested, I may never be made explicitly aware. (…) We would love to work with these platforms to improve their technologies so that they are better adapted for classical music," Hammons says, "but as the situation stands, our input on the issue is limited."

- Background music or video: Everyone knows about the Dancing Baby case. More recently, in July 2020, a video of an interview with Black Lives Matter protestors posted by the alternative media organisation Unicorn Riot was removed from various social media platforms due to the presence of copyrighted audio in the background of the video.

  The same happened to BBC's Gary Lineker who had his account locked by Twitter over him posting a video of Alan Shearer, his 'Match of the Day' co-host, celebrating the Newcastle win vs Chelsea, and notes that: "Presumably the issue was that footage of the game could be seen in the background of the brief clip."

  The 2019 Time 100 Summit livestream was blocked due to claims from Warner Music and Universal Music Group (ab)using the YouTube Content ID system.

  Games developer Epic had to patch their game Fortnite to add a feature to enable players to mute the musical content of Fortnite emotes (i.e. 'dances'). The motivation for this functionality: YouTubers were receiving copyright strikes for uploading such emotes based on Rick Astley's song 'Never Gonna Give You Up', despite the song being licensed by Epic. Techdirt noted that the label considered that "every YouTuber […] has to license the song as well".

  A worrisome trend is that of law enforcement abusing copyright by 'weaponising' music. Police officers started abusing music to prevent protesters from uploading videos of their intervention, as one officer was caught saying: "you can record all you want, I just know it can't be posted to YouTube."

  Vice has also reported on multiple attempts (see here and here) by US Beverly Hills cops to trigger copyright filters on footage being recorded by citizens of their intervention by playing music with their own mobile phone.

  European police forces started copying their American counterparts: French police officers notably attempted to censor yellow vest protestors with Dirty Dancing's soundtrack.

  EFF's Katharine Trendacosta highlights the fact that cops have shown they understand what too many policymakers do not, namely that "copyright can offer an easy way to shut down lawful expression." Law enforcement officers on both sides of the Atlantic have been caught trying to stifle citizens' attempt to film them by starting to play copyrighted music in the background so that the videos or live streams would get caught up in the platforms' copyright filters. "Since YouTube is so dominant when it comes to video sharing, knowing how to game Content ID can be very effective in silencing others," she writes.

- Use of pre-existing material: Automated systems do not have chronological capabilities: they cannot see which piece of content predates another.

  As a result, the insertion of pre-existing material can lead to that original material being blocked. In 2016, the makers of the TV-cartoon show Family Guy 'borrowed' a gameplay clip first uploaded to YouTube in 2009 for one of their episodes. Once this Family Guy episode was aired on Fox, the reference file for this episode was automatically submitted to YouTube's Content ID system. As a result the original clip, which was uploaded 7 years prior to the show's episode airing on TV, was suddenly blocked on YouTube, for allegedly infringing Fox's copyright. The clip was only reinstated on YouTube after public outcry and the intervention of Family Guy creator Seth MacFarlane.

A similar situation occurred after a royalty-free track used by many YouTubers was hit by takedown requests from notably Sony for having sampled a track from 1956 owned by them.

Gaming footage from a virtual reality game was blocked after the game was featured on The Tonight Show with Jimmy Fallon, even though the game developer was perfectly fine with the uploads.

A video from German feminist organisation Pinkstinks was blocked in 2018 by YouTube's Content ID for an alleged copyright infringement of material from broadcaster RTL, while in reality it was RTL that had used the organisation's content in a broadcast.

- Similar material: In June 2020, a claim by the Guinness book of records against the video footage of a record-setting Super Mario brothers speedrun, uploaded by the speedrunners themselves.

  This led not only to the concerned footage being taken down but also videos of lots of other speedruns that looked similar. More oddly, a picture of a freshly baked bread was blocked by Twitter as 'adult content'. It required human intervention for the 'banned buns' to be reinstated.

- Odd unchecked claims of ownership: Techdirt has reported how one YouTuber recently saw two videos demonetised for copyright claims over the numbers '36' and'50'.

  Another user had his scheduled livestream taken down before it was even made available to the public, likely on the basis of a manual notice by Warner Bros out of fear that the live stream would infringe CNN content. On top of that, due to this one copyright strike this user had their live-streaming functionality disabled for three months.

  A Dutch YouTuber got a ContentID claim for 'infringing' his own content, as he explained to the BBC: "Someone took my track, added vocals and guitar to make their own track, and uploaded it to YouTube, but I got the copyright infringement notice!"

  65daysofstatic, an experimental band from Sheffield, England, got a video to promote their upcoming album removed by automated content filters on behalf of Sony, who distributes the band's label.

  In August 2021, Universal Music Group (UMG) issued a copyright claim for a Facebook post by British filmmaker Philip Bloom of a video, with no audio, showing the moon as seen from Greece. The automated claim by Facebook's algorithm alleged that Bloom's video matched 30 seconds of video owned by UMG. It took five days for UMG to release the claim, after which Bloom could safely re-upload his video. He noted that an unknown number of other moon videos, some of them from over 3 years ago, also got caught up in UMG's claim over the moon.

  A programmer live-streaming his coding faced a copyright claim on YouTube over the sound of him typing on his keyboard.

- Anglo-Saxon automated systems' difficulties with other languages: Users have seen their uploaded content or even their account blocked for using words or expressions that are perfectly harmless in their native language but not so in English.

  This includes such words as 'die Boomer' by a German (blocked by Twitter) or 'hoes' (cough) by a South African (also blocked by Twitter).

- Claim by a non-owner of the copyright: YouTuber 'TheFatRat' saw the copyright for one of his songs wrongfully being transferred and attributed to an unknown company by YouTube.
- Lack of common-sense or nuance: Algorithms fail in assessing situations where context matters. Take for example, Tokyo 2020 olympic gold medalist Elaine Thompson-Herah of Jamaica, who got blocked for copyright violations on Instagram after posting snippets of her victories. "After lots of people expressed reasonable outrage at this clearly ridiculous abuse of copyright law, Facebook claimed that it was all a mistake and her account was reinstated," Techdirt explains, adding that "the International Olympic Committee seems to be insisting that Facebook did exactly the right thing." See also Techdirt's reflection on another Olympics-related case involving Sports Illustrated writer Chris Chavez.

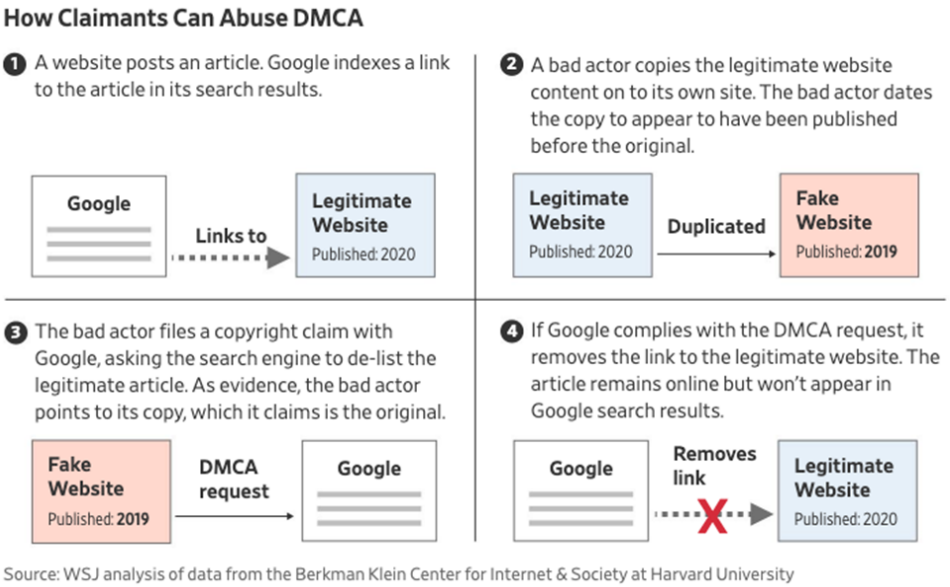
  A Czech gaming studio's game featured on The Tonight Show with Jimmy Fallon, back in 2019, with an unintended outcome: YouTube's Content ID started identifying 'Beat Saber' gaming footage as content from the 'The Tonight Show'. The result: Users who were uploading gaming footage, something that the game's developer encouraged, were seeing this footage being taken down.

  The live-stream of the Hugo Awards has been blocked for showing clips of the TV nominees that were provided by the studios for the purpose of the awards ceremony and for which the organisation had the explicit permission to broadcast them.

3. Censorship motivated use of notice systems: It is also worth mentioning that, next to issues arising from technical measures, it is a public secret that if one wants content to be removed quickly, a copyright notice is the most efficient route even if the end motive is different.

   - The 'backdating trick': Resourceful users have discovered that in order to block content they dislike on platforms, the swiftest method is to claim to be the copyright owner of such content. Claimants can simply (1) copy the original content and (2) backdate their publication prior to that of the original content, in order to get the original content delisted through a notice.

     This method tricks automated content removal mechanisms into thinking that the original website is infringing the claimants' copyright, whilst in reality it is actually the other way around. It was notably used by the owner of a café in Vietnam in March 2020 to block on Google a news story about two British clients with Covid-19, probably in order to avoid negative publicity. The trick was exposed by the Wall Street Journal and led Google to restore more than 52.000 links and identify more than 100 abusive submitters.



Source: *Illustration by the Wall Street Journal*

- The straightforward abusive claim approach: Some claimants simply lie when claiming they own the copyright on a file to get it taken down.

  Last year, an LA Court abused copyright to get an (unauthorised) recording of Britney Spears' court hearing taken down.

  In 2020, the Guardian and the Verge reported on the takedown of Michael Moore's film 'Planet of the Humans' based on what seems to be an abusive copyright claim from a British environmental photographer. In this example copyright is used to stifle free speech, which, regardless of one's opinion on the movie, is not what copyright is intended for.

  Techdirt has highlighted how a US gun rights advocate got a video taken down from Twitter with a bogus copyright claim. US celebrity Ellen DeGeneres took the same route to block a transformative video on Twitter that combined a monologue from her show on her friendship with former President Bush with a background of war images to highlight Bush' contribution to the Iraq war. A joint investigation of German media outlets netzpolitik.org and ARD FAKT uncovered how Acromax Media, a Hamburg-based company working for amongst others the Albanian governing party, used copyright claims to remove videos, mainly on Facebook, criticising the Albanian government in an attempt to suppress independent journalism.