

Copyright's Chilling Tale of Content Moderation: Understanding the Impact of Takedowns on Non-Infringing Content Creators

by Casey Fiesler and Corian Zacher

1 INTRODUCTION AND BACKGROUND

User-generated content platforms engage in a kind of privatized governance over content, expression, and behavior on those platforms, through both policy and design affordances. In addition to decisions about what is allowed and not allowed via their policies, decisions about what actually violates those policies is a question of content moderation. However, as platforms grow in scale, how to do content moderation at that scale is an issue that many platforms are currently struggling with. One alternative to armies of paid human moderators (a paradigm that comes with a number of serious problems) is moderation by algorithm.

Whether conducted by human or algorithm, moderation decisions on platforms are often opaque, and users often (intentionally) have no access to policy processes. Researchers have examined issues such as lack of explanation and ambiguity/vagueness in policies, and in general, appeals processes for moderation decisions are often simplistic or confusing. Though platforms are making governance decisions, there is little sense of due process as there would be in a government context—and when the decision is made by an algorithm rather than a person, this difference is even more stark.

Though we are only just now beginning to see examples of algorithmic moderation in contexts like harassment, hate speech, and adult content, one example that has been ongoing for over a decade is YouTube's ContentID system that automatically flags and removes content posted on YouTube that might infringe copyright. One concern about algorithmic moderation is this concern about due process, and what impact that might have (particularly with respect to false positives and appeals) on the platform's users. Additionally, content moderation often intersects with broader policy—for example, restrictions on child pornography, which predicated the disastrous algorithmic removals of adult content on Tumblr—and copyright is a particularly strong example since Section 512 of the Digital Millennium Copyright Act *requires* platforms to remove content accused of infringing. Therefore, algorithm-based content removals for copyright infringement are a longstanding example of the intersection of policy, online platform design, and algorithms.

This paper examines the impact of content takedowns on a community of content creators who primarily deal in fanworks, which are a type of *transformative works*—that is, new content that

make use of copyrighted material that are transformative of the original. Examples of fanworks include fanfiction (written stories about a character from popular culture, e.g., the continuing adventures of Captain Kirk) or fanvids (remix videos that use media footage edited to music to share an interpretation of the televisual text, e.g., a series of meaningful looks between Captain America and Iron Man to the backdrop of Elvis' "Can't Help Falling in Love"). Prior work has examined the complexity of copyright-related decisions in this particular online creative community, particularly around what constitutes *fair use* of copyrighted content, which is a notoriously gray area. This prior work revealed that misunderstandings of law or platform policy, lack of confidence of their own interpretation, and/or fear of "getting in trouble" resulted in chilling effects where creators decided not to create or share certain kinds of content or use certain platforms (notably, YouTube). This study builds off of that work by specifically examining the impact of takedowns—how they are experienced and responded to—through a large-scale survey of fan creators.

1.1 DMCA 512 and YouTube's ContentID

[Because I haven't written this section yet, I will point you to this recent EFF white paper that does a nice job of laying out how the ContentID system works and how it relates to DMCA 512: <https://www.eff.org/wp/unfiltered-how-youtubes-content-id-discourages-fair-use-and-dictates-what-we-see-online>]

Because of the connection between ContentID removals and formal DMCA takedowns, though a DMCA takedown still has to be triggered manually by a copyright holder, they typically *begin* as an algorithmic decision via ContentID. It is important to note, therefore, that though our survey questions as described below refer to the DMCA and "takedowns," it is likely that many respondents are actually referring to content removals and/or appeals via ContentID *before* a formal DMCA takedown. Because our interest in this paper is largely on the impact of algorithmically initiated content takedowns generally, this distinction is not necessarily important for our findings.

2 METHODS

2.1 Data Collection

In 2017, in collaboration with the Organization for Transformative Works, we conducted an online survey about fan experience with and knowledge of copyright. OTW disseminated the call for participation via social media channels (e.g., Facebook, Twitter, Tumblr) and mailing lists; the recruitment was targeted at adults who self-identified as part of fandom. The survey

was disseminated widely, and there were a total of 2376 participants who answered some part of the survey; all questions were optional.

The survey contained questions about fandom participation, perceived level of copyright knowledge, copyright-related opinions (e.g., "Are fanworks copyright infringement?"), knowledge of certain copyright concepts (e.g., "Do you know what fair use refers to in U.S. law?"), and experiences with copyright (e.g., "Has concern about copyright law ever impacted your decisions or behavior when it comes to fanworks?"). The survey also asked for demographic information--age, gender, location, occupation (all open answer), and level of education (multiple choice)--as well as length of fandom participation and whether they consider themselves as a fan creator.

2.2 Data Analysis

Members of the OTW legal committee conducted initial analysis that was published in a report and covered some basic aspects of the data (e.g., level of fandom involvement, perceived knowledge of copyright law) as well as a close look at answers regarding fair use. For the current study, we conducted an analysis of only the responses that came from the two questions related to the DMCA and takedowns and a third question about general copyright experiences where the answers often touched on these topics. These questions included:

Q34. Are you familiar with the Digital Millennium Copyright Act (DMCA)?

- No, I have never heard of it
- I have heard of it, but I do not know what it is
- Yes, and I think the DMCA is: _____

Q35. Do you know what to do if your fanwork is the subject of a takedown notice on YouTube?

- I do not know what a takedown notice is
- I have received a takedown notice and here is what I did do: _____
- I am familiar with takedown notices but would not know what to do
- Here is what I would do if I received a takedown notice _____

Q44. Have you ever experienced an issue with copyright related to your fanwork? (e.g., your work being taken down, or finding out that someone else copied your work)

- No
- Yes (explain what happened) _____

Q34 was intended to get at a baseline of copyright knowledge, and Q25 to get at actual past experiences as well as hypotheticals for those who have not had those experiences. We analyzed

Q44 as well, because though some respondents reproduced their answers about YouTube experiences, it also provided a number of non-YouTube DMCA experiences--including those in which the DMCA helped participants to *protect* their work in other contexts.

In addition to discerning descriptive statistics for these responses, we also conducted a qualitative analysis of open responses to parts of the above questions. To conduct this analysis, we began with two researchers conducting inductive, open coding on a subset of the data and then coming to a consensus on high-level categories to be applied deductively to the remainder of the data. One researcher conducted the majority of the deductive coding, and met weekly with the second researcher to discuss, iterate, and make collective decisions on edge cases. We then wrote theme memos based on cross-cutting themes across questions and categories, which became the primary themes in our findings. Both researchers are law school graduates.

2.3 Participants

The demographics of survey participants track to other large-scale fandom surveys, showing that participants are majority female (N%), between the ages of N and N (mean N). The majority (N%) reside in the United States, and N% have an education level of at least college. The vast majority are fan creators rather than just consumers, and the population represents heavy fandom participants (including writers and artists) who use a range of platforms but dominantly Archive of Our Own and Tumblr. Previous studies of copyright and fan creation have been drawn from similar populations.

3 FINDINGS

Our findings largely highlight the challenges and problems associated with DMCA and takedowns for content creators with good faith beliefs of fair use. In contrast to previous work on copyright challenges for remixers and fan creators, rather than focusing on the ambiguity of the law itself (e.g., “Is this actually fair use?”), our findings that deep dives into content removals specifically reveal *process*-related issues, including complications involving the use of algorithms to make legal judgments and lack of access to due process. First, we see a large amount of evidence of chilling effects, which we categorize into three major types: policy-based, process-based, and fatalistic. We then cover perceptions of DMCA misuse beyond big copyright owners, and end with examples in our data of when the takedown process works as intended.

Of the 1844 total survey respondents who answered Q34 (“Are you familiar with the DMCA?”), only about a quarter responded that they did, though a smaller number of those could correctly describe it. As described in our background section, the DMCA includes two major pieces: 1201, which covers copyright anti-circumvention (and is relevant to fandom due to ripping DVDs to create remix videos), and 512, which covers notice-and-takedown. Unsurprisingly, only 7

respondents mentioned *both* of these parts. Another 97 correctly described just DMCA 512, 56 respondents had a completely incorrect answer, and the rest had a partially wrong, incomplete, or ambiguous answer.

Additionally, Q35 gives us an idea of how many participants have actually had experiences with content removals on YouTube, either through formal DMCA takedowns or ContentID removals. N participants said that they had received a takedown notice on YouTube; a number of the quotes below come from these participants (where the quote comes from Q25(2)). Additionally, analysis of Q44 revealed some experiences with the DMCA outside of YouTube; a total of 148 mentioned takedown experiences, and N were unrelated to YouTube. This may seem like a small overall number out of N total survey respondents, but recall that less than 300 participants indicated that they create fanvids, which means that that makes up most of the participants who would be uploading fanworks to YouTube.

3.1 Chilling Effects

Regardless of whether they understand the basic practicalities of the DMCA or ContentID, most participants saw these mostly as a hindrance to their creativity and legal rights. Our data suggests that this attitude is in part due to the “black box” nature of both the ContentID algorithm and the takedown process itself. Without being prompted to express an opinion, a number of respondents were negative or hostile in explaining the policy--however, it is important to note that this negativity does not relate to a desire to infringe copyright, but is rather rooted in feelings of confusion or injustice. They described experiences of fear and confusion about what the state of the law is, cynicism about the law’s technical and political legitimacy, and concern that it reinforces the power imbalance between copyright holders and internet users.

The dominant theme across all of the responses that we analyzed was that of chilling effects. In the context of the law, a chilling effect describes a situation where speech is suppressed because of a fear of penalty. Prior research about fanworks and copyright has confirmed that this attitude generally exists when operating in the legal gray area of fair use, though our findings point to specific patterns of causes for chilled content on YouTube specifically: policy-based chilling effects, process-based chilling effect, and fatalistic chilling effects.

3.1.1 Policy-based Chilling Effects

As shown in prior work, fair use is ambiguous enough that many fan creators are concerned about not knowing whether their content is infringing or not. However, this concern is

amplified on YouTube where their copyright policy makes the process to fight the algorithmic decision seem arduous and even actively discourages doing so.

A number of participants specifically mentioned the policy around takedowns on YouTube in which videos are taken down automatically via an algorithmic flag, and those flagged users are then required to appeal the decision. Some see this policy as favoring the copyright owner or providing no real recourse to the content creator.¹

Usually [takedowns] happen because of music as youtube has quite strict views on audio copyrights. There's nothing an author can do about that, the site's policy is that this content is taken down with or without the creator's consent.

I knew there was a chance I could contest it, but I thought even if my vids were legal YouTube might not know that and might just tell me they were illegal, so it wasn't worth bothering.

Others mentioned that a reason they did not engage with the counter-notice process was that the policy is “scary”:

I sent back a 'fair use' claim? something like that? the wording is very scary though because they make it sound like they're going to take you down.

Filing a DMCA counter-notice is done under penalty of perjury, which could understandably be scary on its own, particularly given that it requires a “good faith belief” that content is not infringing, and fan creators are often unsure about fair use. However, YouTube’s 2011 “Copyright School” video, which is required watching for anyone who receives a copyright strike on their account. The video (starring a cartoon rabbit) explains the concept of remix, noting that it might require permission from the copyright owner--unless the new work is a fair use! A voiceover then proceeds to rattle off text from the fair use statute as it appears on the screen and literally crushes the poor rabbit (see Figure 1)—ending with the suggestion that he “consult a qualified copyright attorney.” Then in explaining the counter-notice process, it highlights that submitting false information will result in account termination, and that “if you misuse the process, you could end up in court.” A gavel hits the cartoon protagonist in the head as a voiceover says: “You would get in a lot of trouble. That’s how the law works.”

¹ In this paper, we use the term “copyright owner” to refer to the party who makes a copyright claim, and “content creator” (typically our participants) to refer to the person who received the claim for a piece of (typically remixed) content.



Figure 1: a screenshot from YouTube’s “Copyright School” video

From making fair use sound like an impenetrable concept that will never be understood, to suggesting that amateur content creators hire lawyers to emphasizing legal consequences—all of these things appear to be designed to discourage remix. In our data, we see that even in cases that could have a very strong fair use argument, creators were afraid to go through the counter-notice process. And YouTube’s “copyright school” is not only seen as potentially insulting, but it further reinforces that “you could get in a lot of trouble” for going through the counter-notification process.

I think my video was fair use (a small handful of clips from a movie series rearranged to highlight the director's use of racism) but I was too scared to fight it

Currently, my YouTube account bans uploads and has had a banner on it for month about how I need to attend 'copyright school.' I find that just plain condescending and humiliating.

3.1.2 Process-based Chilling Effects

Beyond the policy itself, confusion around and/or the difficulty of the counter-notification process also results in chilling effects. Even when creators have less anxiety about or are more confident about their rights and the fair use status of their content, some creators are either unaware that there is a way to fight a takedown, are confused about the process, or simply find it too arduous--or just pointless.

Fiesler and Bruckman’s 2014 paper about fair use and fanworks included a quote from a fan creator who, after a single takedown, stopped using YouTube altogether and instead kept their fanvids on a password-protected website, because they were afraid of getting into trouble. We

saw many similar examples in our data, sometimes because they didn't *know* there was recourse:

Eventually I took the works down altogether. I was unaware that I might have recourse and so felt this was the only option.

Other participants who received takedown notices were aware of the appeal process--but decided that fighting it would be too much trouble or hassle:

I had a song on a video that was copyrighted and I simply took the video down. I'm sure there are ways to dispute that but it didn't seem worth it.

I complied - the prospect of further hassle is unwanted stress.

Another participant who was very confident that the video was fair use still succumbed to the chilling effect of a seemingly arduous process:

I felt too unmotivated and busy to do anything but it was definitely unfair because the one taken down was a clip I modified into a parody which is perfectly legal for fair use.

One specific type of arduousness of the process relates to privacy concerns. Fandom has very strong privacy norms, including pervasive use of pseudonyms over real names. Though there are a number of reasons these norms exist, including a large LGBTQ population with safety concerns and a history of stigma, one origin point is fear over the entire community being vulnerable to copyright law. However, the DMCA counter-notice process requires revealing your real name (on YouTube and elsewhere), and moreover, even the process of appealing an algorithmic ContentID flag on YouTube requires a signature with a real name. As one participant revealed, they were unwilling to go through the process for this reason:

I've had a couple of podfic removed by MediaFire, presumably because I used pieces of music. I thought about fighting it, but then didn't feel like exposing my fannish identity too much, so I just used another download service.

In these cases, even if the policy itself is fair, or even if it favored the content creator over the copyright owner, the appeals process was what instigated chilling effects. Sometimes this was in the form of allowing the work (even when confident it was not infringement) to stay down and/or refraining from posting other content on YouTube for fear the same thing would happen.

3.1.3 Fatalistic Chilling Effects

Our findings also revealed that even when creators might understand the process and even are willing on principle to engage with it, they choose not to because, essentially, they think there's no point. This may be because of the resources required or power imbalance (e.g., they have lawyers and I don't), because of the fact that they're fighting an algorithm, or because of a general sense of "learned helplessness," or knowing normatively that there's no point because the outcome will always be bad.

Considering that the "Copyright School" video actually suggested that a remixer should talk to a lawyer *before* posting their remix, it is unsurprising that some participants were anxious about resources that might be required to fight a takedown. For example, one participant, even confident that their work was not infringing, both lacked the necessary resources and thought it pointless to fight an automated decision:

I let it stand, because I don't have the time/resources to fight for its posting. This was a short film clip to be used by students in a film class I taught, where the link was private and only available to students, and its intended purpose as educational material was clearly stated; all of which is meaningless because the clip was automatically flagged.

The concept of learned helplessness applies to situations in which one feels as if a given situation or outcome is unavoidable and that they have no control over changing it; we have seen this reflected in people's attitudes about their data privacy. Similar, many of our participants expressed a "what's the point?" attitude towards the idea of appealing the decision."

As much as I wanted to fight it, I figured it'd be a pointless uphill battle and sadly accepted the notice for what it was.

I googled stuff but was scared my appeal would fail so I did nothing.

Some of these comments pointed to a feeling of simply being worn down, that after some number of times being flagged, they simply stopped trying:

Youtube blocked several videos I'd put together featuring copyrighted music and images. I simply stopped posting anything there. Couldn't find the point and had nothing completely original to post at the time. Also, if you post anything that

even resembles something else already on the site, your video will be blocked, even if it's completely original work.

Our findings combined with prior work suggest that the origin of this attitude is a combination of factors--not just personal experiences, but hearing about other people's over a long period of time, as well as some inherent distrust of the intentions of big tech companies like YouTube or an assumption that they will always care more about corporate copyright owners than about their users.

These three types of chilling effects--policy-based, process-based, and fatalistic--are not mutually exclusive, and overall attitudes are likely a combination of these factors. However, the end result is the same—some amount of non-infringing content not shared, either from the start, or due to removals that are not contested.

3.2 DMCA Misuse

The imbalance of power between big copyright owners and content creators on platforms like YouTube was a theme across our data. This particularly comes out in discussions of the scary and/or arduous process of filing a counter-notice. In contrast, the process for a copyright owner to file a claim is fairly simple, particularly if there are lawyers involved and they do a large number of these (e.g., through ContentID). These takedowns may or may not be in good faith--and we certainly have examples of huge overreach (e.g., *Lenz v. Universal's dancing baby* YouTube video). However, there is another kind of misuse of the DMCA process.

Some participants pointed to the potential for the DMCA takedown process to be abused for the purposes of trolling or harassment. Multiple participants had personal experiences with their content being taken down after being flagged for copyright infringement, not by the content owner, but by a third party using the process as a weapon for harassment.

Petitioned for reinstatement, because the takedown was placed on a ten second long parody video which mixed audio and video from three different sources. Clearly transformative. (It had been reported as a violation by someone who had a personal grudge against me. Ah, fandom wank.

People actually once reported me on behalf of a person who had copied my work and wanted my stuff taken down? ???? Drama drama drama.

[more here – unfinished section]

3.3 DMCA Proper Use

3.3.1 Successful Appeals

Though it was a small proportion of our data, we also see examples of when the DMCA or ContentID counter-notices work as intended. Typically, these involved the respondent having been very confident about fair use, otherwise having permission, or in one case, getting assistance from OTW. These findings beg the question: How do we get *more* of this and *less* chilling effects?

My video was fair use (a parody), and I filled out their protest form. Then they reinstated my video.

On cases where I was infringing (eg, [anime music videos] with copyrighted songs) I just accepted the strike, on one of the cases where I had permission from the band, I emailed YouTube with this and the strike was removed.

YouTube sent me a notice regarding music use in a vid. I followed the advice of the OTW in responding, citing fair use practices, and the vid was restored.

This last example, in which the creator sought out help from the OTW (most likely through an email exchange with the Legal Committee), points to the important of knowledge, education, and assistance in combatting chilling effects.

3.3.2 Protecting Fanworks

It is also important to remember that the DMCA is not just a tool for “big” copyright owners. Content creators also own rights in the original components of their work. Though the process is typically not automated as with ContentID (where copyright owners are notified when a video is algorithmically matched to their content), anyone can request a DMCA takedown if they believe there is content that infringes their copyright on a third-party platform.

Though this was outside the context of YouTube, there were also a number of examples in our data of DMCA takedowns working in *favor* of some respondents when they themselves are the copyright holder.

There have been a few scams in fandom where a for-profit group scrapes AO3 or another archive and puts up the works for its own profit, and sometimes my work was involved in that. I never

thought of that as a copyright issue though. As far as I know, those scammy websites were taken down. In one of the cases, I sent a notice myself, but I don't know how it works -- I was just following the instructions of the people who did.

My fics have been changed into ebooks and distributed where I didn't want them. I submitted DMCA and they were taken down.

In these cases, the power dynamics were flipped--as was the effort. In contrast to YouTube, where ContentID makes having work taken down a simple, automatic process, they had to put in work to allege infringement of their own content, often only through collective action. Cases like the ones described above are not uncommon; there have been a number of cases of unauthorized third parties reproducing content from Archive of Our Own without permission of the authors. The OTW Legal Committee helps fan creators deal with unauthorized copies by providing instructions for issuing DMCA takedown notices.

Though there were so few examples of this in our data that we can only speculate, it is likely that based on privacy issues and concerns about filing counter-notices, these same barriers (e.g., linking a legal name to a fanwork posted under a pseudonym) would also prevent many fan creators from protecting their own copyright by filing counter-notices. Moreover, this is

4 DISCUSSION AND RECOMMENDATIONS

This section is also not written yet, but some initial ideas about strategies for helping with the power imbalance that might be causing chilling effects:

1. YouTube (and similar sites) policies that are not "scary," including:
 - a. clear, readable explanations
 - b. accurate representations of the process for both takedowns and appeal
 - c. scaffolding/help with the appeals process
 - d. active encouragement for users to appeal if they think their content might be non-infringing
 - e. removal of requirement to provide real names for appeals when not legally necessary
2. Education and outreach
 - a. more organizations like OTW to assist content creators beyond fan creators
 - b. advocacy on behalf of and recognition by the copyright office for non-infringing content creators as a legitimate stakeholder in DMCA reform

To expand upon that last point (taken from [a blog post](#) that we wrote on this topic):

On May 21st, the Copyright Office released a report documenting the results of a study five years in the making, its goal to evaluate the current effectiveness of DMCA 512. One obvious conclusion of the study was that “no potential solution(s) will please everybody” (pg 68). However, the problem is that “everybody” as defined by the Copyright Office study seems to include only two categories of stakeholders: online service providers (OSPs, like YouTube) and large-scale copyright owners like movie studios and record companies. Even in concluding that “the notice-and-takedown system as experienced by parties today is unbalanced” (pg 72) (which we agree is true), nowhere do we see the experiences of another relevant category: content creators who make fair uses of copyrighted content. The Copyright Office virtually ignored these “transformative creators,” along with other kinds of Internet content consumers.

However, these relevant stakeholders are also deeply impacted by the lack of balance in DMCA 512. The Copyright Office acknowledges that the goal of Section 512 is balancing the needs of all relevant parties, but at the same time, it vastly underestimates the extent that improper takedown notices affect users. The nod to “inaccurate notices” (when non-infringing content receives a takedown request) even focuses on the cost incurred to the OSP—with no acknowledgment of the cost to the non-infringing content creator (pg 148).

However, as revealed by our findings there is a very real cost to transformative content creators as well, and they should be a relevant stakeholder in policy discussions. Inaccurate takedown notices have a tremendous impact on all creators’ ability to share their work online, and those repeatedly accused of infringing may lose content and accounts as a result. These stakes are particularly heartfelt by small creators, who rely on the reach of platforms to share their content.

Overwhelmingly, the fan creators represented in our survey viewed the lack of balance in the takedown process as stifling their creativity and hindering their ability to share their work. The harm of failure to consider fair use was particularly powerful: Participants whose works were removed explained that they believed that their works constituted fair use, offered examples of uses that fit squarely within fair use by being noncommercial and being used for criticism and parody, and discussed how removal of those works harmed their individual creativity and ability to participate meaningfully in their community.

By assuming that all users are “infringers” or “pirates” and therefore an adversary to the 512 process, the Copyright Office does a disservice to the many content creators who reasonably

and legally rely on fair use. Instead, they should be considered a stakeholder along with OSPs and copyright holders, such that their benefits and harms are part of the balance as well.

Further Reading

Since this draft does not yet have references incorporated, some possible things of interest:

EFF white paper by Katherine Trendacosta: [Unfiltered: How YouTube's Content ID Discourages Fair Use and Dictates What We See Online](#)

[Copyright Office Report on 512](#)

[2017 preliminary results of OTW survey](#)

My prior work on copyright in fan creator communities:

Law review essay overview: [Everything I Needed to Know: Empirical Investigations of Copyright Norms in Fandom](#)

Fiesler, Casey, and Amy S. Bruckman. [Remixers' understandings of fair use online](#). In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pp. 1023-1032. ACM, 2014. **[interview study focusing on fair use]**

Fiesler, Casey, Jessica L. Feuston, and Amy S. Bruckman. [Understanding copyright law in online creative communities](#). In *Proceedings of the ACM CSCW Conference on Computer Supported Cooperative Work & Social Computing*, 2015. **[analysis of online forum posts about copyright]**

Fiesler, Casey and Amy Bruckman. [Creativity, Copyright, and Close-Knit Communities: A Case Study of Social Norm Formation and Enforcement Online](#). *Proceedings of the ACM Human-Computer Interaction*, GROUP. 2019. **[interview study focusing on social norms]**