Before the
FEDERAL COMMUNICATIONS COMMISSION
Washington, DC 20554

In the Matter of	)	
	)	
Section 230 of the	)	RM-11862
Communications Act of 1934	)	

COMMENTS OF REDDIT, INC.

I.   INTRODUCTION

Reddit, Inc. hereby files these comments in opposition to the Petition for Rulemaking to

Clarify Provisions of Section 230 of the Communications Act of 1934—as reflected in the

Federal Communications Commission's August 3, 2020 Public Notice. The American ideal of

free expression requires that users, as well as platforms, can count on not being subject to

liability for undertaking good-faith moderation efforts in community-moderated online spaces.

As a small, privately-held company with limited legal resources, we will leave it to other

Commenters to argue whether the Petition is retaliatory, unconstitutional, or fatally flawed in its

gross misunderstanding of Section 230 and Congress's intent in passing it. Instead, Reddit's most

valuable contribution in this matter comes from our status as among the most popular

forum-based websites in America, and the primarily user-led way in which content moderation

on Reddit happens. It is our view that the debate on Section 230 too often focuses solely on very

large, centrally moderated platforms—and individual grievances with them—to the exclusion of

smaller, differently organized websites that take an alternative approach. Indeed, Section 230's

ingenuity is in protecting and providing for a diversity of business models, and encouraging

innovation, experimentation, and adaptability. Whereas Section 230 has been improperly portrayed as a gift to "big tech," we argue that it is crucial to those smaller companies, like ours, that seek to compete with and offer alternatives to the largest corporate entities. In this way, the law makes an important contribution to consumer choice and marketplace diversity.

However, the most important point that we offer, as we hope to make clear in this filing, is that with regard to Reddit and other community-moderated websites, Section 230 protects our individual users just as much as it does us. Their continued protection is crucial to the viability of community-based moderation online.

II.   REDDIT'S UNIQUE STRUCTURE AND APPROACH TO MODERATION

The way Reddit handles content moderation today is unique in the industry. We use a governance model akin to our own democracy—where everyone follows a set of rules, has the ability to vote and self-organize, and ultimately shares some responsibility for how the platform works.

A. Community Moderators

While more than 430 million users visit Reddit in any given month, it would be improper to think of them or reddit.com in monolithic terms, as our website is divided into more than 130,000 active communities, known as "subreddits," based on shared interests regarding everything from history and science to advice on relationships, parenting, and pet ownership. Each of these communities is created and moderated not by Reddit employees, but by the users themselves. These self-appointed community moderators set specific, often highly-detailed rules tailored to the unique needs of their members and interests. For example, the r/science

community's first rule sets out that the only external links permitted in the subreddit are those that lead to peer-reviewed scientific journals.[1] This rule ensures that the conversation remains evidence-based and rooted in the scientific method. The r/movies community, on the other hand, requires users to mark their post with a specific warning if it contains plot spoilers.[2]

These volunteer community moderators are empowered to remove any post that does not follow their rules, without any direction from Reddit, Inc. As a result of this system, 99.7% of all non-spam content removals on Reddit are undertaken by users themselves, rather than by employees of Reddit, Inc.[3] This self-moderation our users do every day is not only the most scalable solution we've seen to the challenges of moderating content online, but also inherently accounts for differences in cultural norms and contextual nuance among a diverse global userbase.

B. User Upvotes and Downvotes

Each individual user plays a crucial role as well, regardless of whether they serve as a moderator or not, by voting up or down on every post and comment. Through this system of voting, users can accept or reject any piece of content. While most platforms have some version of the upvote function, an action to convey approval or agreement, we see the additional downvote as equally important. The downvote is where community culture is made, through rejecting transgressive behavior or low-quality content. If any community member, not just a moderator, sees poor quality content, they may downvote it. As people do so, the content becomes less visible and, in the case of a comment, may disappear from the default view of the community. Thus, Reddit's voting system essentially turns every user into a content moderator.

---

[1] https://www.reddit.com/r/science/wiki/rules#wiki_submission_rules
[2] https://www.reddit.com/r/movies/wiki/index
[3] https://www.redditinc.com/policies/transparency-report-2019#text-content3

Additionally, accrued upvotes and downvotes feed into the posting user's reputation score, which we call "karma," which is publicly visible to all other users. It's an indicator of the constructiveness of a user's participation on Reddit, and it's possible for karma to be negative. Quantifying a user's reputation in this way incentivizes good behavior.

C. Moderation actions taken by Reddit, Inc.

While this user-led system generally works well, we recognize that we as a company still have responsibilities, and we proactively work to ensure communities stay within our rules, known as our Content Policy, which is obligatory for all users and communities on Reddit.[4] Any user may report Content Policy violations directly to the specialized group of employees at Reddit who form our Safety Team. Their role is to enforce the rules against malicious users, or take down particularly egregious content violations that may have urgent legal or safety implications. We can take action against individual users (for example, through account suspensions), or against entire communities. We try to be as transparent as possible when we take such actions, and we publish our enforcement actions annually in our Transparency Report.[5] Our decisions are also appealable, and we likewise publish the appeal intake and acceptance rate in the Report. Owing to these practices, we are proud that we were the only company to earn a perfect six-star rating from the Electronic Frontier Foundation in their most recent "Who Has Your Back?" report on tech company transparency.[6]

---

[4] https://www.redditinc.com/policies/content-policy
[5] https://www.redditinc.com/policies/transparency-report-2019
[6] Gebhart, Jennie. "Who Has Your Back? Censorship Edition 2019." Electronic Frontier Foundation. 12 June 2019. https://www.eff.org/wp/who-has-your-back-2019.

## III.	REDDIT VOLUNTEER MODERATORS: IN THEIR OWN WORDS

Section 230 protects the moderation decisions of our users as much as it protects Reddit. The changes to Section 230 interpretation and enforcement proposed by NTIA would undermine our community-centered moderation model and place undue burdens on everyday users who make everyday decisions to curate their community. Even with Section 230, our user-moderators have faced lawsuits directed at them by other users angry at moderation decisions. In one recent example, a volunteer moderator of the r/screenwriting community—a largely hobbyist subreddit dedicated to the screenwriting craft—was frivolously sued in connection to good faith actions taken to protect users from what the moderation team perceived to be predatory pay-to-play screenplay competitions.[7]

But rather than take our word for it, we thought that it would be best to let some of our moderators tell you in their own words the thoughtful and often intricate steps they take to care for their communities, and why they choose to volunteer their time to do so. The following section (in *italics)* contains the direct contribution of the moderators of our r/LGBT community.[8]

### A.	What is r/lgbt?

r/lgbt is a community, or "subreddit," on Reddit. We're one of a number of communities on the site dedicated to content by and for the LGBT+ community. Because of our name, we're often deemed "The Home of LGBT+ Content on Reddit." At the time of writing, we have 539,520 redditors who have subscribed to our community. These are our members. We often have between 1.5k and 2k members online in the community at any given time.

---

[7] Compl., *Niebich, J. v. Reddit, Inc., et al.*, Case No. 20STCV10291 (Cal. Sup. Ct., Los Angeles).
[8] https://www.reddit.com/r/lgbt/. In keeping with Reddit's strong privacy values, these moderators prefer to offer this contribution using their Reddit usernames. The r/lgbt community moderator masthead can be found here: https://www.reddit.com/r/lgbt/about/moderators.

B. r/LGBT Community Moderators

Moderators of Communities (often called "mods"), like us with r/lgbt, are not employed by Reddit. We are normal redditors helping to run our communities and keep them safe and welcoming. We volunteer our time because we care for our community and know our community well.

The structure of Reddit gives us wide range to do so. We as users were the originators of our community, and we also set the specific rules for it. Reddit the company gives us and other moderators full power to do so. Accordingly, each subreddit has its own set of community-generated rules specific to the topic of that subreddit. For example, the rules that we have set in r/LGBT contain a specific protocol on the use of trigger warnings on posts, so as to protect those members of our community who may have experienced trauma and are seeking a safe space.

Reddit's paid employees are known as Admins, and they focus on the larger side of site safety, such as dealing with communities which break the Reddit Content Policy. Additionally they work on tools to make the work of community moderators a lot easier. Having this structure of community moderators helps keep content on Reddit organic, as it is we, the community members, who know what is most likely to cause harm to our community.

C. How does r/lgbt prevent abusive content from reaching users?

On r/lgbt we use a number of tools—some automated, some manual, some from Reddit, and some custom designed—to keep our users safe from harmful and abusive content.

The goal of r/lgbt is to provide a safe space online for the LGBT+ community. For this reason we take a more active approach to moderation than many similar online communities. The vast majority of the automated actions get flagged to our moderators for manual review within Reddit's Mod Queue, a list of posts and comments awaiting moderator action. These actions can include removing the post, banning the user from the community, or, if the content is found to be within the rules, approving it.

a. Reddit-developed automated safety tools

Two of our most used tools are developed by Reddit for use by all of their communities. These are AutoModerator and Crowd Control. Each operates differently. AutoModerator focuses on content, following highly customizable rules to target certain instances of harmful and/or abusive content, as defined by community moderators. Crowd Control focuses on users, flagging content from users who have had little to no previous participation in the Community. The point of this is to protect communities from sudden interference, either innocent (for example, if the subreddit is featured in a news article), or organized interference attempts known online as "brigading." AutoModerator is a very powerful tool, but it can be daunting to configure, as some advanced features require a level of familiarity with computer programming concepts. Crowd Control, by contrast, requires almost no effort to implement, but has only a few possible settings.

i. Crowd Control

Crowd Control security ranges from from no filtering to strict filtering. On r/lgbt we have the Crowd Control tool set up using two different security levels. First, the majority of content on our subreddit goes through the "strict" filtering setting, which means that comments from users

who haven't joined our community, new users, and users with negative karma in our community are automatically collapsed. A collapsed comment is neither removed nor fully hidden—Reddit displays the username associated with a collapsed comment, but the text of that comment only becomes visible when a user clicks to expand it.

Second, comments on live broadcasts through the Reddit Public Access Network, go through the slightly eased "Moderate" filtering. This means that only comments from new users and users with negative karma in our community are automatically collapsed.

b. Custom-developed Safety Tools

u/A_Gay_Tornado (AGT) is a custom developed safety bot for r/lgbt developed by our community members. We utilize AGT to set posts to "Community Mode," a type of enhanced Crowd Control used on a post-by-post basis. Community Mode requires commenters to have positive karma in r/lgbt for their comment to be publicly visible. Users who do not meet that requirement have their comments automatically removed by the bot.

u/A_Gay_Tornado is not on by default across the community; it must be enabled by a moderator for individual posts where we need a higher-threshold for participation. This usually occurs when a thread becomes extremely popular or for some reason attracts many spammers or harassers. AGT is a useful tool in these situations to ensure moderators do not have to manually read through the sometimes hundreds of comments on a single post.

c. Active human moderation

Along with our automated tools, which help to prevent abusive content from reaching the community in the first place, our Moderators also do active moderation to tackle abusive content

that slips through. We get reports from Redditors of content that breaks either our Community

Rules, or the Reddit-wide rules, and we review and take action on them as needed.

i.   Community bans

One of the actions that we as community moderators can take is issuing a community

ban. When we as mods ban a user from our community, it does not stop that Redditor from seeing

the posts in our community, but it does stop them from commenting and posting themselves. It

also does not ban them from across the entire Reddit website, a function that is restricted to the

Reddit Administrators.

We typically ban for 3 days, 7 days, 14 days, 31 days, or a Permanent Ban. The ban

length given depends on the context of the ban. We may utilize the different lengths as steps, so a

user would get a 3 day ban, be problematic again, get a 7 day ban, etc. Or if the issue is

significant, we can also just go straight to a permanent ban.

On the 21st of May 2020, we compiled data around the increase of bans from our

community:[9]

Year	Users banned
2010	1
2011	5
2012	1156
2013	755
2014	427
2015	435
2016	1621

---

[9] https://www.reddit.com/r/lgbt/comments/go4fnn/bans_from_rlgbt_by_year_through_may_21_2020_with/

2017	*4784*
2018	*12211*
2019	*11424*
2020 [to 21 May]	*5971*

We're expecting 2020 to finish with a figure somewhere between 13500-14000. This is due to a number of factors including political events, the COVID situation, and other spikes in our forecast for this year.

D. The Impact Online Abuse Has on Our Community

We're seen as the home of LGBT+ content on Reddit not only because of our name r/lgbt, but also because we're tough on online abuse. Typically, abusive content does not become visible in our community. However, we recognize that nothing is perfect, and there are times when abusive content breaks through, or a targeted online attack happens against our community.

For example in April, we had a period of around 3-4 days where some hostile Redditors attempted constantly to post abusive content. We call situations like this a brigade. We had a stream of Redditors coming from other communities, both on-Reddit and off-Reddit, who posted deliberate homophobic and transphobic posts, memes, and comments. They also reported genuine content that should be on the subreddit (we call this "report abuse") and sent abusive and threatening messages to our moderator team. Every few months we will have an issue with brigades, but this one in April was certainly the biggest and most obvious one that we have had in quite a while, since at least 2016. We've definitely seen an increase in abusive content since certain areas began COVID lockdowns and stay home orders, we suspect because of the growth

of people having time to waste doing these sorts of brigades. We also notice that the spikes do tend to occur when certain public figures are homophobic/transphobic, etc.

When this content slips through, it can have a profound effect on our users: suddenly a space where they were not being confronted with abusive content has become poisoned. Members of the LGBTQ+ community come to r/lgbt not just to celebrate the good times, but also to reach out for help, sometimes when they're at their most vulnerable. Imagine being on the brink of self-harm or worse, and being faced with a meme saying that you deserve to burn in hell, or that it would be funny if you killed yourself. The consequences could be devastating.

E. *The impact online abuse has on our moderators*

Being a Community Moderator for r/lgbt is very fulfilling: you're helping your community, helping keep a safe space for hundreds of thousands of LGBTQ+ Redditors, and there are some amazing posts from our Redditors. At the same time, however, it can be very difficult, because whilst the abusive content is not reaching the community, we are still having to sift through it as moderators. You're looking at a lot of comments, reports, and posts telling you that gay people shouldn't exist, that suicide's the better option, that you're a tranny freak or a faggot, or that all dykes need is a bit of dick. You can be targeted through private direct messages with similar types of content, or even death threats. Your crime? Being openly LGBT+ on the internet, and working to keep a small corner of it safe.

Reddit recognizes the challenges that we face in this area and does provide a number of different means of support. They're not perfect and more can be done, but they form a starting point. For one, there is a team of specialized Administrators, known as the Community Team, whose job it is to work directly with moderators, understand their problems, and answer their

questions. This team provides an important link to the engineering teams at Reddit that build moderator tools such as Crowd Control, which was developed and piloted in a collaborative effort this past year. In addition to blocking abusive users from our subreddit or from contacting us, we can also report them to Reddit Admins for potential banning from the entire site. Reddit has also launched some new programs aimed at providing mental wellness support to Moderators. These include dedicated support availability through Crisis Text Line, as well as offering free memberships to mindfulness services for Moderators.

F. How can we make things better?

Let us be clear, unfortunately there is no magic solution to online abuse. It's not something that can be solved overnight, but action does need to be taken. As you can hopefully see from our contribution, we do use a wide-array of technological solutions to prevent abusive content reaching our users. By combining these with human moderators from our community, we are able to specifically target things that have a greater impact on our redditors. This would not be as easily achievable if using general, centralised moderators across the entirety of Reddit, as successfully moderating a community like ours requires specific knowledge of the issues facing LGBT+ people.

Whilst we agree that it is important to consider ways to ensure that online platforms work and are fair for everyone, we are also concerned about steps that would limit us as mods from taking the measures we think are necessary to protect our community, or even worse, put us at legal risk for doing so. If these proposed changes go through, they would profoundly affect our ability as moderators to keep our community safe for our marginalized users, including at-risk categories such as those who are under 18, those who have mental health difficulties, and others.

It would mean that we could no longer guarantee that we could keep those users safe in our community, and would lead to an end of our community as our users currently enjoy.

IV.   CONCLUSION

In the debate on Section 230, it is crucial to account for the diverse array of platforms and services that depend on the flexibility the law offers. Too many of the current arguments regarding Section 230 have focused on the centralized content moderation decisions of large platforms, whose dominance is rightly subject to serious question and criticism. But the debate often ignores that ***Section 230 protects the moderation decisions of individual users*** in the context of Reddit and other community-moderated fora.

Accordingly, the Argument laid out in the Petition would unquestionably harm not only platforms like Reddit, but also the community of volunteer Reddit users that curates them, by subjecting them to liability. This is especially the case for communities serving marginalized groups, such as Reddit's r/lgbt, which rely on the flexibility and protection afforded by Section 230 to guard against abusers. In doing so, they create a shared space with real connection, value, and meaning, across boundaries and borders, in a way that can only be done online.

User-moderated forums like Reddit hearken back to the very charm and community-mindedness of the early internet that Section 230 was seeking to protect. And in doing so, Congress in Section 230 crafted a law that protects not only Reddit, but the volunteer moderators of our many different communities. It protects the moderators of r/science, who prevent unpublished pseudoscience from hijacking evidence-based discussions. It protects the moderators of the r/screenwriting subreddit, including those who were dragged into a frivolous

lawsuit over their good faith efforts to protect their community from scams. It protects the moderators of the r/lgbt subreddit as they guard their community against homophobic abuse. Imagine a universe where trolls could use Section 230 loopholes to sue based on the decisions of these individual moderators. That is the universe that the Petition is leading the internet towards. To allow this to happen would change the very trajectory of the internet. The health of the internet, and users' right to create their own online spaces, hangs in the balance, and for these reasons, the Commission should not undertake a rulemaking proceeding based upon NTIA's petition.

Dated: September 1, 2020