Online Holocaust Denial Report Card: An Investigation of Online Platforms' Policies and Enforcement

INTRODUCTION

The Holocaust, the systematic murder of approximately six million Jews and several million others by the Nazis, is one of history's most painstakingly examined and well-documented genocides.
Nevertheless, in the decades since World War II, a small group of antisemites has repeatedly attempted to cast doubt on the facts of the Holocaust concerning Jewish victims. They claim that Jews fabricated evidence of their genocide to gain sympathy, extract reparations from Germany and facilitate the alleged illegal acquisition of Palestinian land to create Israel.

This phenomenon, known as Holocaust denial, is founded on stereotypes of Jewish greed, scheming, and the belief that Jews can somehow force massive institutions — governments, Hollywood, the media, academia — to promote an epic lie. In the United States, until the early 2000s, Holocaust denial was dominated by the extremist right, including white supremacists, who had a vested interest in absolving Hitler from having committed one of the most monstrous crimes the world has ever known. Today, Holocaust denial in the U.S. has moved far beyond its original fringe circles on the extremist right to become a phenomenon across the ideological spectrum.

On October 12, 2020, after nearly a decade of advocacy by ADL and others, Facebook announced that it changed how its platforms categorized Holocaust denial content. Holocaust denial would no longer be classified as misinformation. Facebook would now address Holocaust denial as hate speech. This change resulted from a campaign by ADL and others in the Jewish community demanding Facebook treat Holocaust denial as one of its most serious content violations.

Holocaust denial is not just a problem on Facebook, however. A September 2020 survey found that 49 percent of American adults under 40 years old were exposed to Holocaust denial or distortion across social media. ADL's 2020 survey of hate, harassment and positive social experiences in online games found that one in ten (10%) of American adult gamers have encountered Holocaust denial in online games. While the actions of Facebook, the largest communication platform in human history, are important, it is also important to keep track of how other
platforms with tens, and sometimes hundreds, of millions of users are addressing pernicious Holocaust denial on their platforms.

In January 2021, ADL investigated how a variety of platforms address Holocaust denial content. Our investigators both reviewed the policies regarding Holocaust denial on various digital social platforms and looked for examples of Holocaust denial across different tech platforms from social media to online gaming sites. We then reported the Holocaust denial content under the platform's anti-hate policies from non-official accounts to see how platforms would enforce their policies for ordinary users. We created a report card on how tech companies are currently managing Holocaust denial content on their digital social platforms based on these results. The results present a snapshot in time, but give rise to significant concern.

Investigation Report Card

Limitations and Caveats

- The investigation our researchers undertook was non-exhaustive, and cannot be said to represent the overall prevalence of Holocaust denial content across any of the stated platforms. What we are evaluating in this investigation are the stated policies of the platforms in question and their responsiveness to the average user in reporting egregious violations of those policies.

- We did not include every digital social platform in this investigation. Our investigators focused on a cross-section of mainstream tech platforms and services that have publicly expressed their commitment to addressing hate on their platform and whose platforms would be easily searchable by an average user. For example, we did not include in this investigation platforms such as Gab, 4chan or Telegram that have few if any policies regarding hate, extremism, or Holocaust denial. We also did not include hosting providers,
payment processors, or online games whose services are not broadly searchable.

- It is important to note that each of these platforms are serving billions of users. We recognize the challenge of enforcing platform policies consistently and at scale. If a platform of a billion users who each post once a day enforced its policies accurately 99 percent of the time, that would still leave 10 million errors on the platform every day. Even so, each of those 10 million errors would not just be a statistic. Each error represents the lived experience of at least one human being per comment, a person exposed to harassment, a threat of violence, misinformation or some other harmful phenomena resulting from platforms operating at these massive scales. Whether enforcement can ever be perfect, we still believe there is significant room for improvement across all platforms regarding how they design their products and prioritize resources toward the moderation of harmful content. This investigation intends to capture how platforms function around Holocaust denial from the perspective of an affected individual user and provide constructive criticism, while also acknowledging that it cannot represent every user of every platform at all times.

THE ONLINE HOLOCAUST DENIAL REPORT CARD EXPLAINED
Our report card methodology involves both evaluating the publicly stated policies of the platforms and their enforcement responses as part of our investigation. As policies are only as strong as how they are enforced, we weighted each tech company's enforcement more heavily in evaluating its efforts to address Holocaust denial on its platform.

In response to this scorecard, Reddit has described its platform as having an explicit Holocaust denial policy, pointing to this post about Reddit's use of "quarantines", which used holocaust denial as an example of content that could be "quarantined". However, ADL views "quarantining" not as a policy but as an enforcement action. Quarantining does not constitute a policy applicable to a type of hate content, but an action to be taken against content that violates some other platform policy. We therefore have not allocated a point to Reddit for having an explicit Holocaust Denial policy. We note that even if we did allocate a point to Reddit for quarantining, our ultimate grade assessment of the platform's response to Holocaust denial content would not change.

### Explicit Holocaust Denial Policies
For this category, we looked across the existing policies of digital social platforms to see if they had a policy explicitly designed to address Holocaust denial. Both Twitch's and Facebook's policies ban Holocaust denial under a subcategory of their hate policies.

From Facebook's hate speech policy, January 2021

From Twitch's hateful conduct policy, January 2021

Facebook has one of the most explicit Holocaust denial policies, but it is also the only platform in our investigation that either failed to respond to our reports or claimed the content we reported did not violate its Holocaust denial policy. On the other hand, Twitch has an explicit Holocaust denial policy and implemented it against the content we reported as part of our investigation.

On other platforms, such as YouTube and TikTok, Holocaust denial policies take a slightly more broad form that focuses on content that attempts to "deny that a well-documented, violent event took place." This would include the denial of the
Holocaust, but also the denial of other well-documented genocides, mass shootings and atrocities such as the Sandy Hook elementary school shooting or the genocide of the Rohingya Muslims in Myanmar. Other platforms go further still. Roblox, a platform primarily for young people under 14, has a policy that prohibits all content regarding "atrocities, massacres, and other shocking real (or pseudo-real) world events." We gave credit to platforms with an explicit Holocaust denial policy under one of these categories in our report card.

**General Hate Policies**

For this category, we looked at tech company's existing policies to see if they had a broad policy around hateful content on their platforms. Here, we looked for cross-platform rules regarding content or behavior that target individuals or groups who belong to a vulnerable or marginalized community. In a U.S. context, this would include people or groups targeted because of their actual or perceived race, ethnicity, religion, gender, gender identity, sexual orientation or disability. Given how online platforms have contributed to the spread and amplification of hate in the past several years, this type of policy should be foundational to any attempts at content moderation of a platform. There is no excuse for platforms not to have robust and broad policies around hate. For this reason, this category was weighted less heavily than the explicit Holocaust denial policy category.

**Effective Product Level Efforts to Address Holocaust Denial**

Many of the platforms included in this investigation have done extensive work removing Holocaust denial content, redirecting users to credible information about the Holocaust, or putting safeguards in place to prevent the easy discoverability and amplification of Holocaust denial content. For example:

- Searches for "Holocaust denial" or "Holohoax" on YouTube directs the user to credible videos that debunk Holocaust denial.
- TikTok and Roblox do not allow users to search for "Holohoax" when trying to find content.
- Facebook has removed some of its larger Holocaust denial groups after its policy change in October 2020.

This report card category gives a platform rating regarding the degree to which we found these efforts effectively mitigated our investigators' efforts in finding Holocaust denial content. We gave credit to those platforms for whom a search for common terms or phrases associated with Holocaust denial, such as "Holohoax," resulted in few or no search results, redirected users to credible information, or blocked a search around such terms and phrases.

**Reporting Effectiveness and Transparency**

Users have to trust that their reports are being addressed to be effective as part of content moderation on digital social platforms. Two categories in our report card speak to this. The first focuses on response time: Did the platform investigate the report and promptly respond to the user? This category does not include auto-generated messages to the user affirming receipt of a report. Instead, it counts any follow-up messages indicating that the platform conducted an investigation and decided whether to enforce its policies. We awarded credit to platforms that responded with the results of their investigation and enforcement decisions within 24 hours.

The second category around the effectiveness of reporting involves explainability. Does the user understand why a platform has made a certain decision based on its stated policies? None of the reports we submitted as part of this investigation that received a response was given a specific reason why an action was or was not taken. This is problematic because it further obfuscates how these platforms make decisions and erodes trust in whether or not the work that users put in to report hateful content is being reviewed or addressed in any meaningful way.
For example, Facebook determined multiple videos from the BitChute channel called "Holocaust Lies Exposed" did not violate its specific community guidelines around Holocaust denial. These videos included captions such as "Holohoaxers talking shit" and "Holohoax tales," and comments such as, "How was it ever possible that people with even half a brain believed this garbage?" In communicating this decision to the user, Facebook stated, "The video was reviewed, and though it doesn't go against one of our specific Community Standards, we understand that it may be offensive to you and others." It offered no additional context and took no further action.

At the same time, Twitter decided to remove content and lock a user's account for a single tweet calling the Holocaust a "#sixmillion #HoloHoax #lie," despite Twitter not having an explicit public-facing policy regarding Holocaust denial. Twitter has said that it has enforced against Holocaust denial under its hate policy in the past based on internal guidance, but has not updated its public-facing policy to reflect this statement. While the above tweet was reported under...
Twitter's hate speech policy, the follow-up on the report stated, "After our review, we've locked the account for breaking our rules." We can assume this was enforced under the company's hate policy, but no specific rule or category of rule was cited. Even when platforms make strong and swift decisions, users should know why both for their edification and build trust in the reporting system.

Transparency at the platform and the user level is imperative for technology companies operating digital social platforms. Unfortunately, as this investigation found, no company was transparent about the policy rationale for its decisions.

**Action Taken on Holocaust Denial Content**
The category we weighted most heavily in this investigation and our report card was whether the platforms to which we reported Holocaust denial content took any action to enforce their stated policies based on our user reports. Holocaust denial is hate speech, rooted in antisemitism and antisemitic conspiracy theories and should have no place on private digital social platforms that purport to be inclusive spaces for all people.

There can certainly be space on platforms for discussion of hateful phenomena, but ADL’s experts judged each piece of content that was reported to the platform as part of this investigation to contain blatant and uncritical Holocaust denial.

The following are a few examples of unaddressed content that remain active as of this writing, despite being reported to the platforms in question under their anti-hate policies.

Only Twitter and Twitch decided to take action against the Holocaust denial content reported. Other platforms stated the content reported did not violate their policies or provided no response at all.

The following are a few examples of unaddressed content that remain active as of this writing, despite being reported to the platforms in question under their anti-hate policies.

Facebook
In this video, noted antisemite "Brother Nathaniel" Kapner makes a variety of Holocaust denial statements, such as "there were no gas chambers."

Instagram

Internment Camps: IRC Records disclose more than just death figures of 271,301 from Typhus during WWII - Soviet sabotage and German patronage

6 million. Ya right. Another deception. Holocaust story is a lie. #holocaustisalie

Many of them in the uk fema camps

In Europe and Canada it is against the law to deny the Holocaust. The truth fears no investigation. Kinda like now with Covid-19. They will strip your medical licence from you.
This post is tagged with the hashtag #holocaustisalie and questions the validity of six million Jews' deaths in the Holocaust.

YouTube

This video features participants expressing Holocaust denial. One person said:

- "The point is the Holocaust has no standing as a historical event. It has standing as a religious concoction because the cult of the Holocaust and Israel are tied in with quasi-religious mindset that is impervious to criticism, which is why people are persecuted even for uttering a rational defense of information that contradicts the official version. You can criticize any aspect of history but if you try to criticize the Holocaust you're persecuted and that calls into question well what is the Holocaust."

TikTok
This account is one of several with the username “hol0caustwasahoax”, replacing the second “o” in Holocaust with a 0 to avoid filters put in place by TikTok.
Holocaust denial is prevalent and easy to find on Reddit. As of this writing, none of the reported posts was addressed.

**Twitter**

Twitter took action against many of the tweets reported as part of this investigation, despite not having a specific policy against Holocaust denial. However, of note was that Twitter only took action on 12 percent of non-English language tweets reported versus 85 percent of English language tweets reported.

**Discord**
This post on Discord shows a user in a channel posting various Holocaust denial videos from BitChute, the extremist-friendly video sharing site.

Steam
A user who posted on the community boards of the gaming platform Steam wrote at length on various Holocaust denialism texts and resources. This discussion occurred under the WWII themed game *War Mongrels*, due to be released in 2021.

Roblox
The gaming platform Roblox has numerous filters in place for discouraging users from searching for Holocaust denial content. This piece of content, named "Defaced Synagogue," circumvents filters for "Holohoax" or "Holocaust denial" by using the word "synagogue" in its title instead of more overt signals of Holocaust denial.

CONCLUSION AND RECOMMENDATIONS
The results of this investigation aim to encourage further research and investment by tech platforms into better addressing the spread of Holocaust denial. One noteworthy finding is that platforms with explicit Holocaust denial policies did not necessarily do better enforcing those policies against our reported content. This raises questions about what makes for an effective platform policy—does expanding platform rules to include a longer list of specific phenomena result in a meaningful change for targeted communities? Our investigation cannot answer this broad question, but it urgently requires a more definitive answer.

Another notable result is the opacity surrounding how platforms enforce their policies, despite years of advocacy from civil society and researchers toward greater transparency. Investigators who were part of this project hypothesized that Facebook was the only platform that determined Holocaust denial was a form of misinformation. As such, we posited that other platforms, whether or not they had Holocaust denial policies, would enforce against Holocaust denial content under their broader hate policies, not their misinformation policies. We cannot prove that hypothesis because no platform specified the policies that guided their enforcement decisions, even though we reported all content under their applicable hate policies. There may be credible safety concerns around being too specific in sharing the rationale related to enforcement decisions, but providing no information does not allow researchers to study how platforms enforce their policies. Such a lack of transparency keeps the public further in the dark on how and why platforms take action.

The September 2020 survey mentioned earlier also found that 63 percent of American adults under 40 did not know that six million Jews died in the Holocaust. Over half thought the death toll was less than two million. Forty-eight percent were unable to name a concentration camp or ghetto, and 56 could not identify the concentration camp Auschwitz. Roughly one in ten (11%) believed that Jews caused the Holocaust. These alarming findings underscore how important it is to combat Holocaust denial and distortion.
More hopefully, 80 percent of respondents in that survey believed that it was important to continue teaching about the Holocaust, so that it does not happen again. Research from ADL's partner organization Echoes and Reflections found that young people who are educated about the Holocaust are more open to various perspectives, have a greater willingness to challenge intolerant behavior in others and have a greater sense of social responsibility.

The combination of lack of knowledge about the Holocaust and the clear benefits of Holocaust education further illustrates the damage that is likely to be caused when Holocaust denial content is allowed to spread on mainstream digital social platforms. While some platforms did relatively well in our investigation, many did not. Below are a few recommendations to platforms to better address Holocaust denial on their platforms in the future.

- **Enforce policies on Holocaust denial consistently and at scale**
  Platform policies on their own are not sufficient; they require enforcement that is consistent across a digital social platform. We recommend that, especially for hateful content such as Holocaust denial, tech companies designate sufficient resources such as training for human moderators, greater numbers of human moderators, and expanding the development of automated technologies to commit as few mistakes as possible when enforcing their policies around this content.

- **Expert review of content moderation training materials**
  Tech companies should seek subject matter experts' advice when drafting and revising the guidelines they use to train their human content moderators. For Holocaust denial, ADL or other Jewish organizations could be an expert resource to review such materials to ensure that content moderators are adequately trained to recognize and address Holocaust denial.
- **Transparency at the user level**

  Platforms should provide users with more information on how they make their decisions regarding content moderation. They should at least provide the user with the policy that guided their decisions, and some semblance of information about why the content did or did not violate the policy in question.

- **Anti-hate by design product choices**

  Tech companies must make changes to their products to prioritize users' safety over engagement and reduce the amount of Holocaust denial and other hateful content on their platforms. Some non-exhaustive options: adjusting what kinds of content are discoverable in search and how content is amplified and to whom, changes to reporting flows to better escalate reported hateful content to the appropriate internal reviewers, closing the feedback loops to ensure that reports of policy violations are routed back to reporters, adding friction to the posting process, reviewing certain users' posts before they are amplified, and more.