

CALIFORNIA REGIONS

CENTRAL PACIFIC (N CA, HI, UT)

720 Market Street, Suite 800
San Francisco, CA 94102
(415) 981-3500
Central-Pacific@adl.org

SETH BRYSK
Regional Director
LINDA HORODAS
Chair, Regional Board
GILBERT R. SEROTA
Chair, Civil Rights

LOS ANGELES

10495 Santa Monica Boulevard
Los Angeles, CA 90025
(310) 446-8000

LA@adl.org
JEFFREY I. ABRAMS
Regional Director
SHARYN NICHOLS
Chair, Regional Board
BRENT SOKOL
ERIC KINGSLEY
Co-Chairs, Civil Rights

ORANGE COUNTY/LONG BEACH

1201 Dove Street, Suite 390
Newport Beach, CA 92660
(949) 679-3737

orange-county@adl.org
PETER LEVI
Regional Director
MICHAEL ROSEN
Chair, Regional Board
DAN TARMAN
Chair, Civil Rights

SAN DIEGO

4950 Murphy Canyon Road, Suite 250
San Diego, CA 92123
(858) 565-6896

san-diego@adl.org
TAMMY GILLIES
Regional Director
ADAM ROSENTHAL
Chair, Regional Board
ALAIN AVIGDOR
Chair, Civil Rights

SANTA BARBARA

1528 Chapala Street, Suite 101
Santa Barbara, CA 93101
(805) 564-6670

santa-barbara@adl.org
DAN MEISEL
Regional Director
MARK A. GOLDSTEIN
Chair, Regional Board
JASON SALTOUN-EBIN
Chair, Civil Rights

WESTERN DIVISION

ROBERT TRESTAN
Vice President
LAURA FENNEL
Director of Strategic Communications
KENDALL T. KOSAI
Director of Policy
ANN M. ORTEGA-LONG
Director of Education Programming

NATIONAL STAFF

STEVEN M. FREEMAN
Vice President, Civil Rights
KAREN LEVIT
Civil Rights National Counsel
LAUREN KRAPP
Technology Policy & Advocacy
Counsel
BETH HOLTZMAN
Western Civil Rights Area Counsel

OFFICERS OF THE

NATIONAL COMMISSION

ESTA GORDON EPSTEIN
National Chair
JONATHAN GREENBLATT
CEO and National Director
KEN JACOBSON
Deputy National Director



Lauren T. Krapf, ADL Technology Policy & Advocacy Counsel

Testimony to California's Senate Judiciary Committee

November 9, 2021

Good afternoon.

Thank you Chair Umberg, Vice Chair Borgeas, Members of the Committee. My name is Lauren Krapf and I serve as ADL's Technology Policy & Advocacy Counsel. It is my honor to speak to the committee today about how social media platforms can take action and address the hate and disinformation running rampant on their platforms.

The way news and current events reverberate online matters. What was once on the fringes of our society is now front and center in our platform newsfeeds. False allegations of voter fraud, hosted and amplified on social media, led to the violent insurrection at our nation's Capitol. The most predictable terror incident in American history was literally planned and promoted, out in the open, on social media.

It is important to note: it's not just card-carrying members of extremist groups that commit acts of violence. According to ADL research, over 70 percent of the people who stormed the Capitol on January 6th were not self-identified extremists. These individuals were motivated through a steady diet of lies and misinformation that social media algorithms pumped out because this content is highly engaging.

ADL's Center for Technology and Society has also engaged in reporting on content management for platforms' Holocaust Denial policies and response to Antisemitism. We put out report cards to see how consistent platforms enforced their policies. When it came to Antisemitism, YouTube and Twitter got the highest grades with a B-. Tik Tok and Facebook got a C-. When it came to Holocaust Denial, YouTube Twitter and Tik Tok got C's; Facebook got a D. This is alarming and deeply troubling.

In addition to misinformation and extremism, ADL has been tracking the growing threat of online harassment on social media. As my colleague David Sifry discussed earlier, according to ADL's latest data, 1 in 3 Americans who were harassed online attribute at least some harassment to a protected characteristic. This hate isn't only taking place on dark web message boards. 75% of those harassed said at least some harassment happened on Facebook—24% on Twitter, 21% on YouTube, 15% on Snapchat: well-resourced, highly profitable platforms.

Major social media companies will tell you that it's too hard to address these issues, that the legal framework will prevent us from regulating their platforms. It's simply ridiculous. Automobiles, food, prescription drugs, and securities. Time and time again, lawmakers have crafted good policy to protect consumers and industry alike. **There is a lot we can and must do to push hate and extremism back to the fringes of the digital world and repairing our internet starts with transparency.**

Transparency is about shining the light on social media companies' content management policies and practices. It is about getting access to consistent and regular data to ensure that big social media platforms are taking the necessary steps to protect consumers.

Much like the importance of tracking and documenting hate crimes, tracking disinformation, extremism, hate, harassment and foreign interference on social media platforms through transparency reports will provide critical data that can be used to inform policy and protect targeted groups.

Transparency reform would motivate platforms to be more explicit about their policies on hate, harassment, and misinformation, and apply their rules consistently. It would act as a deterrent from making changes, exceptions, or other decisions that end up amplifying hate. It would create an environment where social media companies can compete on how well they are protecting users, not on how they can optimize the most corrosive content to keep us scrolling for as long as possible to sell as many ads as possible.

Today, most major social media companies only publish limited information about their content policies and enforcement. Efforts by these companies have been opaque, arbitrary, biased, and inequitable. Why is this the case? Currently, social media companies have little to no legal or financial incentives to give consumers comprehensive information. Their reports end up serving as a deflection away from the truth about what content proliferates on platforms.

Recent revelations from leaked documents show that this isn't just a theory--organizations like Facebook had an unannounced program called XCheck that allowed over 5 million celebrities, politicians, and influencers to effectively skirt all of the published rules and policies. That meant that none of the posts by millions of public figures went through automated systems that normally flag violating content. The most influential accounts got a free pass for posting almost anything they wanted whenever they wanted. **We only know about this because of a whistleblower. None of this was disclosed in a single platform report.** What other bad behaviors would they stop if they knew they had to report on their policies and enforcement consistently?

There are clear issues with the way platforms currently share the limited information they do publish.

- **First, the information platforms do share is missing crucial context.** For example, Facebook reported that last quarter it actioned 31.5 million pieces of hate content. But documents uncovered by whistleblower Frances Haugen showed that the company took action on as little as 3 to 5 percent of all hateful content on the platform. That means Facebook may have missed over a billion pieces of hate content. How can we reconcile these numbers? Is Facebook being intentionally deceptive in their reporting? What we're seeing is numerators without denominators.

- **Second, platforms are cherry picking the information they share**, resulting in an inability to meaningfully compare how platforms are responding to hate, extremism, and misinformation.
- **Third, information is hard to find.** Many big social media platforms have established microsites, sometimes known as “transparency centers;” however, transparency centers are inaccessible and not user friendly. A few examples: while people may be able to locate the most up-to-date reports, it is nearly impossible to find and compare past reports. If someone downloads Twitter’s report, the 8 page PDF only shows half the text of the microsite. Consumers should not have to expend so much time or energy searching for information that should be readily accessible.

Today we are in the dark. These transparency reports are exercises in obfuscation.

Consistent and comprehensive transparency will finally clarify platform policies and reveal data regarding enforcement of those policies. We’ll understand rules, changes to rules, exceptions to rules, and categories of violations. We will also get high-quality data: not just publishing numbers, but explaining the significance of those numbers.

Transparency is a foundational step to creating a safer and more equitable internet. To be clear, transparency is not forcing companies to reveal trade secrets. It’s not about requiring companies to have specific guidelines. It is not about restricting platforms’ editorial decisions or suppressing freedom of speech. It’s not about reforming Section 230. Transparency is about companies being honest with consumers about the rules of the road and how they enforce those rules.

Additionally, transparency reports are not overly burdensome or expensive. Especially for social media companies that make hundreds of millions-billions-or trillions of dollars in gross revenue. Currently, most social media platforms already track much of the information we are seeking. If we reform transparency, the relatively small quarterly cost of compliance would be offset by the enormous benefit to public safety.

Finally, inflated claims that bad actors and extremists would somehow take advantage of social media platforms if they were more transparent is a straw argument. Bad actors and extremists are already exploiting social media platform capabilities--anonymized, aggregated data about content moderation won’t make this worse, rather it will give tools to researchers, civil society and policymakers to counter the threat.

Transparency is nothing new.

We require it for companies that choose to sell their shares in the public market. We require it for companies that sell us food and cosmetics. We should demand the same for social media platforms that have the power to shape the way we digest information.

Unchecked, hateful and extremist content pushes women, people of color, LGBTQ+ people, and other marginalized groups out of online spaces. It has fractured our democracy and spread health misinformation. We know these harms exist but we don't know the extent of them. What we do know is that the problem is too severe and the consequences are too grave. Transparency regulation and reform is a necessary and important start to repairing our internet.

Thank you and I look forward to answering your questions.