

**UNDER EMBARGO
UNTIL TUESDAY,
MAY 13, 2025**



SOCIAL MEDIA SAFETY INDEX 2025

TABLE OF CONTENTS

Meeting This Moment: A Letter from GLAAD President and CEO Sarah Kate Ellis	03
Executive Summary, Findings and Recommendations, Methodology	04
2025 Social Media Safety Index Platform Scorecard <ul style="list-style-type: none">• TikTok• Facebook• Instagram• Threads• YouTube• X	10
2025 Platform Scorecard Indicators	15
Social Media Safety Index Advisory Committee	16
Acknowledgements	17

2025 SMSI Articles and Reports Appendix

2025 Appendix of Policies, Guidelines and Reporting Links

MEETING THIS MOMENT:

Letter from GLAAD President and CEO Sarah Kate Ellis

Now in its fifth year, GLAAD's Social Media Safety Index (SMSI) continues to be the leading analysis of anti-LGBTQ online hate, harassment, and disinformation, with a year-round program dedicated to protecting safety, privacy, and expression for LGBTQ people online.

Recent years undeniably illustrate how online hate speech and misinformation negatively influence public opinion, legislation, and the real-world safety and health of lesbian, gay, bisexual, transgender, and queer (LGBTQ) people. The landscape of social media platform accountability work has shifted dramatically since GLAAD's first SMSI report in 2021, with new and dangerous challenges in 2025.

The recent major ideological shifts from Meta have been particularly extreme. Earlier this year, the company announced it would retreat from established norms of trust and safety in favor of welcoming hate speech, and further place the onus on users to block blatantly harmful content that would otherwise violate its policies. The **Platform Scorecard** and the **Key Findings and Recommendations** of this edition of the SMSI reflect the early fallout of the current situation and point to why collective advocacy to hold social media companies accountable is so urgent.

Given the current dehumanizing attacks on LGBTQ people and new challenges to social media safety, this year's SMSI includes refreshed strategies to advance our mission. GLAAD continues monitoring, researching, and reporting on anti-LGBTQ hate, harassment, and disinformation on major social media platforms, while leaning in to new approaches including: expanding our reach by providing stakeholder guidance to additional tech and AI companies; presenting and advocating SMSI recommendations to other social media platforms;

providing tools, information, and resources directly to LGBTQ creators, activists, organizations, and everyone. When companies put our community — and other marginalized communities — at risk, GLAAD leads with facts, education, organizing, and humanity.

Forty years ago, GLAAD's founders launched this non-profit advocacy organization because they recognized the immense power and opportunity that all forms of media have for advancing acceptance and understanding of LGBTQ people. As new forms of media have proliferated over the years, GLAAD's work to hold these companies accountable has grown in scope and importance as well. We stand ready and strong to meet this moment.



GLAAD PRESIDENT AND CEO
SARAH KATE ELLIS

EXECUTIVE SUMMARY, FINDINGS AND RECOMMENDATIONS, METHODOLOGY

// Companies must do better to protect their users and the public interest.

— Ranking Digital Rights

In addition to the annual **Platform Scorecard** ratings below, this year's Social Media Safety Index (SMSI) report provides a set of **Key Findings** and **Recommendations** as guidance for companies to improve LGBTQ social media safety, privacy, and expression across their platforms. The **Scorecard** employs 14 LGBTQ-specific indicators to evaluate policies and product features of six major platforms (TikTok, X, YouTube, and Meta's Instagram, Facebook, and Threads), adapting the standard methodology of noted tech and human rights research organization [Ranking Digital Rights](#) (RDR). GLAAD urges platforms to review RDR's extensive research and recommendations,¹ and exhorts all platforms and tech companies to prioritize the **Key Findings** and **Recommendations** of the SMSI.

An Important Contextual Note About the 2025 Platform Scorecard and Methodology Changes

The 2025 scores are not directly comparable to the 2024 scores due to extensive revising of the 2025 Platform Scorecard methodology. In some cases, existing indicators and elements were revised to clarify our evaluation standards (e.g., for Q12 — language was revised to clarify that content moderator trainings should take place annually). We also added a new indicator and several elements addressing emerging threats to LGBTQ safety, privacy, and expression that have gained traction since we initially developed the Scorecard. These revisions to the methodology resulted in universal score declines across platforms. Therefore, year-to-year comparisons of the 2025 scores to previous scores will not be an accurate reflection of relative platform progress. Relevant policy changes for each of the platforms are discussed below.

EXECUTIVE SUMMARY OF PLATFORM SCORECARD

The 2025 SMSI **Platform Scorecard** scores reflect the fact that some platforms have made product and policy improvements in the past year, while others have retreated from best practices in LGBTQ safety, privacy, and expression (some, drastically so). Overall the scores remain extremely low. The most significant changes this past year are the draconian rollbacks from Meta² and YouTube³, particularly their retractions of policy protections for transgender and nonbinary people. In a positive change from last year's evaluation, YouTube updated its advertising policies which now prohibit advertisers from excluding users from seeing ads based on their sexual orientation and gender identity.

Although the current state of LGBTQ social media safety may seem worse than ever, GLAAD continues to work with all platforms to improve their products, speaking out as a constant voice of advocacy urging all companies to protect LGBTQ people, and especially trans people, online.

The indicators of the Platform Scorecard are collectively measuring against the following LGBTQ safety, privacy, and expression best practices.

Every platform should have public-facing policies that: protect LGBTQ people from hate, harassment, and violence on the platform; prohibit targeted misgendering⁴ and deadnaming⁵ on the basis of gender identity; prohibit content promoting so-called “conversion therapy;”⁶ prohibit advertising content that promotes hate, harassment, and violence against LGBTQ individuals on the basis of protected characteristics; explain the proactive steps it takes to stop demonetizing and/or wrongfully removing legitimate content and accounts related to LGBTQ topics and issues; and explain its internal structures to best ensure the fulfillment of its commitments to overall LGBTQ safety, privacy, and expression on the platform.

Companies should also provide users with a dedicated field to add and change gender pronouns on their user profiles; and explain what options users have to control or limit the company’s collection, inference, and use of data and information related to their sexual orientation and their gender identity.

Companies should state that: they do not recommend content to users based on their disclosed or inferred sexual orientation or gender identity, unless a user has proactively opted in; and that they do not allow third-party advertisers to target users with, or exclude them from, seeing content or advertising based on their disclosed or inferred sexual orientation or gender identity, unless the user has proactively opted in.

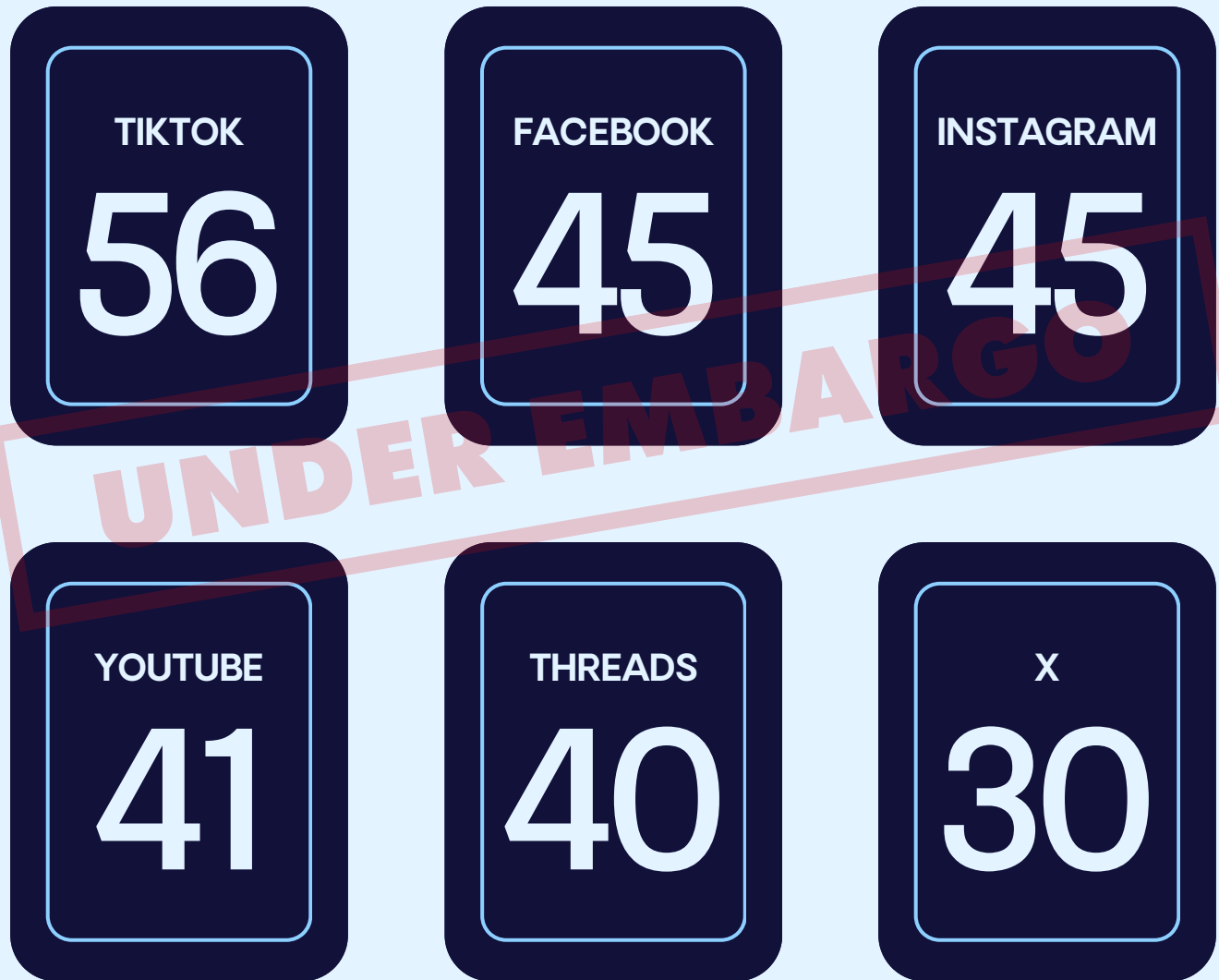
In the realm of transparency, every platform should regularly publish data about the actions it takes to restrict content and accounts that violate policies protecting LGBTQ people; and about the actions it takes to stop demonetizing and/or wrongfully removing legitimate content and accounts related to LGBTQ topics and issues.

Lastly, to create products that better serve all of its users, the company should make a public commitment to continuously diversify its workforce, and ensure accountability by annually publishing voluntarily self-disclosed data on the number of LGBTQ employees across all levels of the company.



Gender Spectrum Collection

THE 2025 PLATFORM SCORECARD SCORES:



EXECUTIVE SUMMARY OF KEY FINDINGS AND RECOMMENDATIONS

All platforms evaluated in the **Scorecard** have (some⁷) policies prohibiting hate and harassment on the basis of sexual orientation, gender identity and/or expression, and other protected characteristics. Given the difficulty of assessing policy enforcement methodologically — which is further complicated by a lack of transparency from the companies — these failures are not quantified in the **Scorecard** scores. However, in GLAAD’s day-to-day research and monitoring, and in reports from other organizations, researchers, and journalists, failures are seen repeatedly in both the development of policies and in their enforcement across major platforms.⁸

The **Key Findings and Recommendations** bullet points below are drawn from GLAAD’s year-round work and research, to accompany the **Platform Scorecard**. The most notable highlight of the 2025 research is the pair of findings that: in addition to inadequate moderation of harmful anti-LGBTQ material (for example, see GLAAD’s 2024 report, *Unsafe: Meta Fails to Moderate Extreme Anti-trans Hate Across Facebook, Instagram, and Threads*), platforms also frequently over-moderate legitimate LGBTQ expression. This includes wrongful takedowns of LGBTQ accounts and creators,⁹ mislabeling of LGBTQ content as “adult”¹⁰ or “explicit,” unwarranted demonetization of LGBTQ material,¹¹ shadowbanning,¹² and other kinds of suppression¹³ of LGBTQ content.¹⁴ (Such unwarranted restrictions occur with non-LGBTQ content as well.¹⁵)

Additional LGBTQ safety, privacy, and expression issues include: lack of algorithmic transparency and harmful algorithms; inadequate transparency and user controls around data privacy; lack of transparency with regard to content moderation protocols including information about moderator trainings; apparent over-reliance on AI moderation without human review; failures to effectively moderate anti-LGBTQ content in many non-English languages; reductions in transparency tools and

access for independent researchers; among other issues — all of which disproportionately impact LGBTQ users and other marginalized communities who are uniquely vulnerable to the harms of online hate, harassment, and discrimination.¹⁶ These areas of concern are exacerbated for those who are members of multiple communities, including people of color, women, immigrants, people with disabilities, religious minorities, and more.¹⁷ Social media platforms should be safe for everyone, in all of who we are.

As a US-based non-profit organization GLAAD’s focus is primarily domestic; however, there are enormous global implications of this work, and GLAAD calls upon platforms to take responsibility for the safety of their products worldwide.¹⁸ Social media platforms are vitally important for LGBTQ people, as spaces where we connect, learn, and find community.¹⁹ While there are many positive initiatives these companies have implemented to support and protect their LGBTQ users,²⁰ they simply must do more. Lastly, as GLAAD has long noted, proposed legislative social media safety solutions must be mindful of not censoring LGBTQ resources or causing unintended harm to LGBTQ users, especially LGBTQ youth.²¹



Gender Spectrum Collection

KEY FINDINGS

- Recent hate speech policy rollbacks from Meta and YouTube present grave threats to safety and are harmful to LGBTQ people on these platforms.²²
- Platforms are largely failing to mitigate harmful anti-LGBTQ hate and disinformation that violates their own policies.²³
- Platforms disproportionately suppress LGBTQ content, via removal, demonetization, and forms of shadowbanning.²⁴
- Anti-LGBTQ rhetoric and disinformation on social media has been shown to lead to offline harms.²⁵
- Social media companies continue to withhold meaningful transparency about content moderation, algorithms, data protection, and data privacy practices.²⁶

KEY RECOMMENDATIONS

- **Strengthen and enforce (or restore) existing policies and mitigations** that protect LGBTQ people and others from hate, harassment, and misinformation;²⁷ while also reducing *suppression* of legitimate LGBTQ expression.²⁸
- **Improve moderation** by *providing mandatory training for all content moderators* (including those employed by contractors) focused on LGBTQ safety, privacy, and expression; and moderate across all languages, cultural contexts, and regions.²⁹ AI systems should be used to flag for human review, not for automated removals.³⁰
- **Work with independent researchers to provide meaningful transparency** about content moderation, community guidelines, development and use of AI and algorithms, and enforcement reports.³¹
- **Respect data privacy.** Platforms should reduce the amount of data they collect, infer, and retain,³² and cease the practice of targeted surveillance advertising,³³ including the use of algorithmic content recommender systems,³⁴ and other incursions on user privacy.³⁵
- **Promote and incentivize civil discourse** including working with creators and proactively messaging expectations for user behavior, such as respecting platform hate and harassment policies.³⁶

METHODOLOGY

For the **Key Findings and Recommendations** of the SMSI, GLAAD's Social Media Safety (SMS) team reviewed research, journalism, and reports across the field of social media safety and platform accountability. The SMS team also consulted with the SMSI advisory committee and other organizations and leaders in technology and human rights. The past year's developments in the field of LGBTQ social media safety have been tracked in the [2025 SMSI Articles and Reports Appendix](#). Please also refer to the [2024](#), [2023](#), [2022](#), and [2021](#) SMSI reports.

The 2025 **Platform Scorecard** methodology and research guidance from research analyst Andrea Hackl can be found [here](#). The full detailed scoring sheets are available [here](#).

SIGNIFICANT 2024-2025 REPORTS ON LGBTQ SOCIAL MEDIA SAFETY

In 2021, the inaugural GLAAD Social Media Safety Index report [offered](#) the first-of-its-kind dedicated analysis on LGBTQ safety and social media platforms. [There are](#) now many powerful reports and studies devoted to these issues, and some of the most significant of the past year are listed in our [2025 Appendix of Articles and Reports](#). We urge [everyone](#), especially platform leadership and executives, to read the full reports.

On the Firewall Between Financial Sponsorship and GLAAD's Advocacy Work

Several of the companies that own products and platforms listed in this report are current financial sponsors of GLAAD, a 501(c)3 non-profit. A firewall exists between GLAAD's advocacy work and GLAAD's sponsorships and fundraising. As part of our media advocacy and media watchdog work, GLAAD publicly calls attention to issues that are barriers to LGBTQ safety, as well as barriers to fair and accurate LGBTQ content and coverage — including issues originating from companies that are current financial sponsors.

2025 SOCIAL MEDIA SAFETY INDEX PLATFORM SCORECARD

The 2025 SMSI **Platform Scorecard** consists of 14 indicators that draw on best practices from the [Ranking Digital Rights \(RDR\) Big Tech Scorecard](#), the highly-respected evaluation of the world's most powerful digital platforms on their policies and practices affecting people's rights to freedom of expression and privacy. All evaluations are conducted looking at the public-facing policies of each platform. While results have global implications for LGBTQ safety, privacy, and expression online, our analysis is based on the English-language versions of public-facing policies and documents covering the U.S. market. Additional methodology information is available in the [Research Guidance](#). The full list of indicators and elements used in the evaluation process is available [here](#). **Please also see the important contextual note above about the 2025 Platform Scorecard and Methodology Changes.**

The Platform Scorecard evaluates six major social media platforms: Facebook, Instagram, and Threads (whose parent company is Meta); TikTok (parent company: ByteDance); YouTube (parent company: Alphabet/Google), and X.

TIKTOK

SMSI
SCORE: **56**

In the 2025 SMSI Platform Scorecard, TikTok receives a score of 56. Out of all of the platforms evaluated in this year's Platform Scorecard, TikTok's hate and harassment policies provide the most comprehensive protections for LGBTQ people. TikTok also clearly states that public figures are protected under its Community Guidelines. Of the scored platforms, TikTok continues to be one of only two with a policy expressly prohibiting both targeted misgendering and deadnaming. According to its "[Inclusion and Belonging Guide](#)," TikTok also prohibits the promotion of dangerous "conversion therapy" content, and has several public-facing policies that explain how it enforces this policy.

However, the platform continues to fall short of sufficiently protecting LGBTQ privacy, safety, and expression in other key areas. TikTok should show greater transparency around the wrongful removal and demonetization of LGBTQ-related content and accounts, and clearly explain the steps it takes to address such actions. As recommended to other platforms, the company should also publish

corresponding data that sheds light on these processes and protocols.

TikTok is more forthcoming than its peers about its engagement with LGBTQ organizations and other key stakeholders, but does not sufficiently explain whether it receives input on key issues such as content promoting "conversion therapy," and the wrongful removal and demonetization of LGBTQ-related content and creators. The company's policies should provide more details on the specific issue areas it receives guidance on, and explain whether meetings with key stakeholders take place at least once per year.

Although it states a public commitment to diversifying its workforce, TikTok is the only platform besides X that does not publish any data on its LGBTQ employees. TikTok should track and publish annual data showing its progress toward reaching diversity and inclusion goals, including publishing voluntarily disclosed data on the number of LGBTQ employees across different US-based teams.

META PLATFORMS

FACEBOOK
SMSI SCORE:

45

INSTAGRAM
SMSI SCORE:

45

THREADS
SMSI SCORE:

40

In the 2025 Platform Scorecard, Facebook and Instagram each receive a score of 45, while Threads receives a score of 40. In a major set of policy changes that have been widely denounced by human rights advocates and tech policy experts,³⁷ Meta revised its “Hateful Conduct” policy this year to expressly allow and encourage hate, harassment, and discrimination against LGBTQ people.³⁸ As one example, according to the updated policy, Meta “allow[s] allegations of mental illness or abnormality when based on gender or sexual orientation, given political and religious discourse about transgenderism [sic] and homosexuality [sic]...”³⁹ These harmful policy changes coincide and align with the increasing political attacks on LGBTQ people and their rights, and contribute to a larger societal anti-LGBTQ animus that leaves LGBTQ people vulnerable to increasing levels of hate and harassment online and off.

Meta should remove these harmful exceptions from its “Hateful Conduct” policy and provide LGBTQ people with strong protections against hate, harassment, and violence on its platforms.

Facebook’s policies also fail to adequately protect LGBTQ people in several other key areas. According to its “Gender Identity Policy and User Tools” policy, Meta prohibits targeted misgendering. However, this policy document is not easy to locate and doesn’t reside in Meta’s Transparency Center. Tier 3 of Meta’s “Bullying and Harassment” policy further states that “all private minors, private adults (who must self-report), and minor involuntary public figures are protected from: ... Claims about romantic involvement, sexual orientation or gender identity.” However, unlike the “Gender Identity Policy and User Tools” policy, the “Bullying and Harassment” policy does not contain any explicit examples that make it

clear whether targeted misgendering falls under this policy. In addition, it is not clear whether the company will continue to enforce its targeted misgendering policy in light of Meta’s significant changes to its “Hateful Conduct” policy. While the updated policy does not specifically address targeted misgendering, several provisions explicitly permit different forms of hate speech and harassment against LGBTQ people (and other historically marginalized groups). GLAAD will continue to closely monitor how Meta’s policy changes will impact LGBTQ people on its platforms. We will also closely monitor how Meta’s enforcement of its newly revised “Hateful Conduct” policy will impact its apparent commitment to protecting transgender, nonbinary, and gender-nonconforming people from targeted misgendering.

In addition to other critical changes to its “Hateful Conduct” policy noted above, we expect the company to include an explicit commitment in its Community Standards that protects LGBTQ people from targeted deadnaming on the platforms. The company should also explain in detail how this policy is enforced. Moreover, this policy should also cover public figures and should not require self-reporting by the targeted individual.

To better protect LGBTQ users’ privacy and give them full control over their own data, Meta’s public-facing policies should provide comprehensive information on the options that users have to control the collection, inference, and use of data related to their sexual orientation and gender identity. Meta currently provides insufficient transparency on this topic.

Facebook and other Meta platforms do receive credit on the Scorecard’s new “conversion therapy” content indicator. In its “Hateful Conduct” policy,

Meta prohibits content promoting so-called “conversion therapy.” However, the policy also contains a notable caveat as Meta states that it “require[s] additional information and/or context to enforce” this prohibition. It is not clear from the policy how users can provide this additional information or context. According to this policy, users should not post “content explicitly providing or offering to provide products or services that aim to change people’s sexual orientation or gender identity.” Since advertisers must comply with Meta’s Community Standards, harmful “conversion therapy” content is therefore also prohibited in Meta ads.

In its “[2024 Responsible Business Practices Report](#),” Meta made a commitment to “building a workforce that’s as diverse as the communities we serve.” In the same report, the company also published partial data on its LGBTQ workforce. However, the company announced more recently that it would abandon its diversity and inclusion efforts.⁴⁰ It is important to note that this significant policy reversal in Meta’s commitment to DEI is not reflected in the 2025 SMSI Platform Scorecard data. As our research for the SMSI Platform Scorecard is based on public-facing company policies published in the previous year, we expect that this policy change will be reflected in our 2026 Scorecard.



Gender Spectrum Collection

In the 2025 SMSI Platform Scorecard, YouTube receives a score of 41. In a deeply concerning update to YouTube's "[Hate Speech](#)" policy, the company removed "gender identity and expression" from its list of protected characteristic groups, which suggests that the platform is no longer protecting transgender, nonbinary, and gender-nonconforming people from hate and discrimination according to its Community Guidelines. The company [has claimed](#) that the policy has not changed, however it is an objective fact that the gender identity protection is no longer expressly present in its public-facing policy. YouTube should reverse this dangerous policy change and update its "Hate Speech" policy to expressly include gender identity and expression as a protected characteristic.

As the only platform evaluated in the SMSI Platform Scorecard that does not have a policy prohibiting targeted misgendering and deadnaming, YouTube should adopt this best practice policy commitment to protect transgender, nonbinary, and gender-nonconforming people from targeted misgendering and deadnaming. As previously mentioned, this policy should apply to public figures and should not require self-reporting. Further, YouTube should also publicly explain that it employs various processes and technologies — including human and automated content moderation — to detect content and behaviors violating these policies, and explain its enforcement process in detail.

YouTube has a public-facing policy that clearly prohibits advertising content promoting so-called "gay conversion therapy," but does not have a similar policy prohibition for user content. We call on YouTube to address this policy gap and implement a comprehensive policy that clearly bans user content that promotes the dangerous practice of so-called "conversion therapy" practices on the platform. The company should also clearly explain the processes and technologies it would use to identify content violating this policy, and how the policy would be enforced.

In a positive change from last year's evaluation, YouTube updated its advertising policies which now prohibit advertisers from excluding users from seeing ads based on their sexual orientation and gender identity. In its "[Community Guidelines enforcement](#)" report, the company publishes limited data on the actions it has taken to remove content and accounts that violate policies protecting LGBTQ people.

In its [2024 Diversity Annual Report](#), the company made a public commitment to taking proactive steps to diversify its workforce, and published voluntarily self-disclosed data on the number of LGBTQ employees. However, as has been reported more recently by mainstream news outlets,⁴¹ Google is rolling back its DEI commitments in response to the Trump administration's Executive Orders on DEI.⁴² Research for the 2025 SMSI Platform Scorecard is based on public-facing company policies published over the course of the last year, since the company has not added this major stated policy reversal in Google's commitment to diversity and inclusion to its published policies, this change is not reflected in this year's SMSI Platform Scorecard data and will be reflected in the 2026 research.



Gender Spectrum Collection

In this year's evaluation, X receives a score of 30. While X is one of only two platforms that prohibits both targeted misgendering and deadnaming, a recent policy revision indicates that X only provides this policy protection "where required by local laws." The policy also does not provide sufficient protections for public figures, and X says it "must always hear from the target to determine if a violation has occurred," effectively requiring targeted individuals to always self-report policy violations. X should reverse this policy change and adopt a comprehensive policy that prohibits targeted misgendering and deadnaming, regardless of local laws. This policy should also cover public figures and should not require self-reporting by the targeted individual.

The company also does not provide adequate policy protections in other key areas. Like YouTube, X should implement a comprehensive policy prohibiting the promotion of dangerous "conversion therapy" content in user content. The company should also clearly explain the processes and technologies it uses to identify content violating this policy, and explain in detail how it enforces the policy. In addition, the company should also clearly state that it prohibits the promotion of "conversion therapy" in advertising.

X provides more comprehensive policy protections in other key areas. For example, the company's "Hateful Conduct" policy clearly prohibits hate, harassment, and violence against LGBTQ people: "You may not directly attack other people on the basis of race, ethnicity, national origin, caste, sexual orientation, gender, gender identity, religious affiliation, age, disability, or serious disease ... We prohibit the dehumanization of a group of people based on their religion, caste, age, disability, serious disease, national origin, race, ethnicity, gender, gender identity, or sexual orientation ... We consider hateful imagery to be logos, symbols, or images whose purpose is to promote hostility and malice against others based on their race, religion, disability, sexual orientation, gender identity or ethnicity/national origin ..."

The company continues to provide no transparency on any proactive steps it takes to hire employees from diverse backgrounds. X should make a public commitment to diversifying its workforce, and should publish voluntarily disclosed data showing its progress towards reaching diversity and inclusion goals.



Gender Spectrum Collection

2025 PLATFORM SCORECARD INDICATORS

- 1: The company should have public-facing policies that protect LGBTQ people from *hate, harassment, and violence* on the platform.
- 2: The company should have a public-facing policy that states it provides users with a dedicated field to *add and change gender pronouns* on their user profiles.
- 3a: The company should have a public-facing policy that prohibits *targeted misgendering*⁴³ on the basis of gender identity.
- 3b: The company should have a public-facing policy that prohibits *targeted deadnaming*⁴⁴ on the basis of gender identity.
- 4: The company should have a public-facing policy that prohibits content promoting so-called “conversion therapy.”⁴⁵
- 5a: The company should have a public-facing policy that explains what options users have to *control or limit* the company’s collection, inference, and use of data and information related to their *sexual orientation*.
- 5b: The company should have a public-facing policy that explains what options users have to *control or limit* the company’s collection, inference, and use of data and information related to their *gender identity*.
- 6: The company should have a public-facing policy that states that it does not *recommend content* to users based on their disclosed or inferred sexual orientation or gender identity, unless a user has *proactively opted in*.
- 7: The company’s public-facing policies should state that it does not allow third-party advertisers to *target* users with, or *exclude* them from, seeing content or advertising based on their disclosed or inferred sexual orientation or gender identity, unless the user has *proactively opted in*.
- 8: The company should have a public-facing policy that prohibits *advertising content* that promotes hate, harassment, and violence against LGBTQ individuals on the basis of protected characteristics.
- 9: The company should regularly publish data about the actions it has taken to *restrict content and accounts* that violate policies protecting LGBTQ people.
- 10: The company’s public-facing policies should explain the proactive steps it takes to *stop demonetizing and/or wrongfully removing legitimate* content and accounts related to LGBTQ topics and issues.
- 11: The company should regularly publish data about the actions it has taken to *stop demonetizing and/or wrongfully removing legitimate* content and accounts related to LGBTQ topics and issues.
- 12: The company should publicly commit to providing *mandatory training for content moderators*, including those employed by contractors, focused on LGBTQ safety, privacy, and expression on the platform.
- 13: The company should have a public-facing policy that explains its *internal structures* to best ensure the fulfillment of its commitments to overall LGBTQ safety, privacy, and expression on the platform.
- 14: To create products that better serve all of its users, the company should make a public commitment to *continuously diversify its workforce*, and ensure accountability by periodically publishing voluntarily self-disclosed data on the number of LGBTQ employees across all levels of the company.

Read more information about the [2025 SMSI Platform Scorecard Indicators and Elements](#).

SOCIAL MEDIA SAFETY INDEX ADVISORY COMMITTEE

The GLAAD SMSI advisory committee includes an array of respected leaders working at the intersections of tech accountability and LGBTQ rights.

ALOK (they/them), Writer and performer

Lucy Bernholz, Ph.D (she/her), Founding Co-Director, Stanford Digital Civil Society Lab

Alejandra Caraballo, Esq. (she/her), Clinical Instructor, Cyberlaw Clinic, Berkman Klein Center for Internet & Society, Harvard Law School

Joan Donovan, Ph.D (she/her), Founder, Critical Internet Studies Institute and Assistant Professor of Journalism and Emerging Media Studies, Boston University

Jelani Drew-Davi (they/them), Senior Communications Specialist, Kairos

Liz Fong-Jones (she/her), Field CTO, Honeycomb

Evan Greer (she/they), Director, Fight for the Future

Leigh Honeywell (she/her), Lead Security Strategist, 1Password

Maria Ressa (she/her), Journalist, Co-founder, and CEO, *Rappler*

Tom Rielly (he/him), Founder, TED Fellows Program and Founder, PlanetOut.com

Sarah T. Roberts, Ph.D (she/her), Faculty Director, UCLA Center for Critical Internet Inquiry and Associate Professor, Gender Studies at UCLA

Brennan Suen (he/him), Deputy Director of External Affairs, Media Matters for America

Kara Swisher (she/her), Renowned journalist and Editor-at-Large at *New York Magazine*, and host of “On with Kara Swisher” and “Pivot” podcasts

Marlena Wisniak (she/her), Senior Legal Manager, AI and Human Rights, European Center for Not-for-Profit Law

ACKNOWLEDGMENTS

GLAAD is grateful to the many organizations and individuals with whom we do this important work. We especially want to acknowledge our SMSI advisory committee: ALOK; Lucy Bernholz; Alejandra Caraballo, Esq.; Joan Donovan; Jelani Drew-Davi; Liz Fong-Jones; Evan Greer; Leigh Honeywell; Maria Ressa; Tom Rielly; Sarah T. Roberts; Brennan Suen; Kara Swisher; and Marlena Wisniak.

Very special thanks to: Research consultant, Andrea Hackl and the team at Ranking Digital Rights. Thanks to GLAAD staff including Social Media Safety Program Manager, Leanna Garfield; Senior Director of Social Media Safety, Jenni Olson; Communications Manager, Jose Useche; Communications Consultant, Allison Bloom; Brand Creative Director, Crystal Stratford; Creative Studio Manager, Elizabeth Fernandez; Graphic Designer, Roho Yniguez; Web Producer, Heidi Spillum; Senior Manager of News & Research, Sarah Moore; and Chief Communications Officer, Rich Ferraro. Thanks are due as well to the staff at all of the social media companies who work with GLAAD to improve LGBTQ safety on their platforms. Extra special thanks to anyone else we forgot to mention, and to the countless journalists, researchers, activists, and others who continue to drive change forward.

This report would not have been possible without the support of: Craig Newmark Philanthropies and Logitech. Thank you to David Hornik, Kara Swisher, and the Stanford Center on Philanthropy and Civil Society for early support of this project.

If you'd like to support the work of the Social Media Safety program and this report, please donate to GLAAD at GLAAD.org/donate.

FOOTNOTES

¹ The RDR Index offers a robust evaluation of 14 of the world's most powerful digital platforms, looking at more than 300 aspects of company policies, including indicators on: Which companies commit to human rights? Who discloses the most about how they moderate content? Which have the safest data privacy policies and practices? And much more.

² Ina Fried, "Meta's new policies open gate to hate," Axios, January 9, 2025, <https://www.axios.com/2025/01/09/meta-moderation-transgender-women-hate>

³ Taylor Lorenz, "YouTube removes 'gender identity' from hate speech policy," User Mag, April 3, 2025, <https://www.usermag.co/p/youtube-removes-gender-identity-from>

⁴ Targeted misgendering is a form of hate speech that involves the intentional use of the wrong gender and/or gender pronouns when referring or speaking to a transgender, nonbinary, or gender non-conforming person. Source: <https://glaad.org/releases/glaad-responds-twitters-roll-back-long-standing-lgbtq-hate-speech-policy/>

⁵ Targeted deadnaming is a form of hate speech whereby a person intentionally "reveal[s] a transgender person's former name without their consent – often referred to as 'deadnaming' – [which] is an invasion of privacy that undermines the trans person's true authentic identity, and can put them at risk for discrimination, even violence."

Source: <https://glaad.org/releases/glaad-responds-twitters-roll-back-long-standing-lgbtq-hate-speech-policy/>

⁶ "Conversion therapy" is a widely condemned practice that involves any psychological or religious intervention aimed at changing an LGBTQ person's sexual orientation, gender identity, or gender expression. Complicating efforts to address the amplification of harmful "conversion therapy" content online, its purveyors also promote this dangerous practice under alternate labels such as "leaving homosexuality" and "unwanted same-sex attraction."

Sources: <https://glaad.org/reference/conversion-therapy>; <https://globalextremism.org/reports/conversion-therapy-online-the-ecosystem-in-2023/>

⁷ Meta's January 2025 changes to its Hateful Conduct policy eviscerated many protections for LGBTQ users. GLAAD's overview of these policy rollbacks can be found [here](#).

⁸ Clint Rainey, "The Hate Speech Landscape on Facebook Is Worse than You Thought," Fast Company, August 31, 2024, <https://www.fastcompany.com/91129012/meta-facebook-hate-speech-in-advertisements-content>.

⁹ Tricia Crimmins, "A Casting Call for Trans Actors Caused Instagram to Suspend Several Accounts for 'Human Exploitation,'" The Daily Dot, July 11, 2024, <https://www.dailydot.com/debug/san-francisco-trans-march-instagram/>.

¹⁰ Nick Wolny, "The Pink Shadow Ban: How LGBTQ+ Influencers Are Fighting Censorship," Out Magazine, July 1, 2024, <https://www.out.com/out-exclusives/tiktok-lgbtq-social-media-censorship>.

¹¹ Sara Kingsley et al., "'give Everybody [...] A Little Bit More Equity': Content Creator Perspectives and Responses to the Algorithmic Demonetization of Content Associated with Disadvantaged Groups," Proceedings of the ACM on Human-Computer Interaction 6, no. CSCW2 (November 7, 2022): 1–37, <https://doi.org/10.1145/3555149>.

¹² Tatum Hunter, "What Is Shadowbanning? Why Social Media May Be Hiding Your Posts. - The Washington Post," The Washington Post, October 16, 2024, <https://www.washingtonpost.com/technology/2024/10/16/shadowban-social-media-algorithms-twitter-tiktok/>.

- ¹³ Taylor Lorenz, “Instagram Blocked Teens from Searching LGBTQ-Related Content for Months,” User Mag, January 6, 2025, <https://www.usermag.co/p/instagram-blocked-teens-from-searching>.
- ¹⁴ Devin Coldewey, “Oversight Board Presses Meta to Revise ‘convoluted and Poorly Defined’ Nudity Policy,” TechCrunch, January 17, 2023, <https://techcrunch.com/2023/01/17/oversight-board-presses-meta-to-revise-convoluted-and-poorly-defined-nudity-policy/>.
- ¹⁵ Tomas Apodaca and Natasha Uzcátegui-Liggett, “Demoted, Deleted, and Denied: There’s More than Just Shadowbanning on Instagram – the Markup,” The Markup, February 25, 2024, <https://themarkup.org/automated-censorship/2024/02/25/demoted-deleted-and-denied-theres-more-than-just-shadowbanning-on-instagram>.
- ¹⁶ Rachel Keighley, “Hate Hurts: Exploring the Impact of Online Hate on LGBTQ+ Young People,” Women & Criminal Justice 32, no. 1–2 (October 17, 2021): 29–48, <https://doi.org/10.1080/08974454.2021.1988034>.
- ¹⁷ Aditya Vashistha et al., “Vulnerable, Victimized, and Objectified”: Understanding Ableist Hate and Harassment Experienced by Disabled Content Creators on Social Media, 2024, <https://www.adityavashistha.com/uploads/2/0/8/0/20800650/ableism-chi-2024.pdf>.
- ¹⁸ Jaishree Kumar, “Queer Indians Confront Online Hate While Tech Platforms Stay Indifferent,” BOOM, June 26, 2024, <https://www.boomlive.in/decode/queer-indians-confront-online-hate-while-tech-platforms-stay-indifferent-25721>.
- ¹⁹ Linda Charmaraman, J. Maya Hernandez, and Rachel Hodes, “Marginalized and Understudied Populations Using Digital Media,” Handbook of Adolescent Digital Media Use and Mental Health, July 14, 2022, 188–214, <https://doi.org/10.1017/9781108976237.011>.
- ²⁰ “#ForYourPride: Celebrating TikTok’s Visionary LGBTQIA+ Community,” TikTok Newsroom, May 31, 2024, <https://newsroom.tiktok.com/en-us/foryourpride-celebrating-tiktoks-visionary-lgbtqia-community>.
- ²¹ “Regulatory Oversight – 2024 Social Media Safety Index,” GLAAD, May 21, 2024, <https://glaad.org/smsi/2024/solutions-for-all/>.
- ²² “More Transparency and Less Spin - Center for Countering Digital Hate: CCDH,” Center for Countering Digital Hate | CCDH, February 24, 2025, <https://counterhate.com/research/more-transparency-and-less-spin/>.
- ²³ Shelby Jamerson, “Media Matters and GLAAD Found 100 Meta Posts Containing Anti-Trans Slur,” Media Matters for America, August 28, 2024, <https://www.mediamatters.org/facebook/media-matters-and-glaad-found-100-meta-posts-containing-anti-trans-slur>.
- ²⁴ Sophia Chen, “The Lost Data: How AI Systems Censor LGBTQ+ Content in the Name of Safety,” Nature News, September 24, 2024, <https://www.nature.com/articles/s43588-024-00695-4>.
- ²⁵ Archie Macfarlane, “Online Anti-LGBTQ+ Hate and Its Offline Consequences - Tech against Terrorism,” Tech Against Terrorism, July 27, 2024, <https://podcast.techagainstterrorism.org/1684819/episodes/15314807-online-anti-lgbtq-hate-and-its-offline-consequences>.
- ²⁶ Jordan Kraemer, “On Social Media, Transparency Reporting Is Anything but Transparent,” Tech Policy Press, February 15, 2024, <https://www.techpolicy.press/on-social-media-transparency-reporting-is-anything-but-transparent/>.
- ²⁷ Alice Hunsberger, “A Guide to Protecting LGBTQ+ Users,” Everything in Moderation*, June 3, 2024, <https://www.everythinginmoderation.co/lgbtq-safety-pride-month/>.
- ²⁸ Tatum Hunter, “What Is Shadowbanning? Why Social Media May Be Hiding Your Posts. - The Washington Post,” The Washington Post, October 16, 2024, <https://www.washingtonpost.com/technology/2024/10/16/shadowban-social-media-algorithms-twitter-tiktok/>.

²⁹ Alice Hunsberger, “The Importance of Anti-Bias Training in Content Moderation,” Partner Hero, 2024, <https://www.partnerhero.com/blog/anti-bias-training-in-content-moderation>.

³⁰ Sophia Khatsenkova, “The EU Tells Twitter to Hire More Human Content Moderators amid Concerns of Rise of Illegal Content,” Euronews, September 3, 2023, <https://www.euronews.com/next/2023/03/09/the-eu-tells-twitter-to-hire-more-human-content-moderators-amid-concerns-of-rise-of-illegal-content>.

³¹ Matt Motyl et al., “Making Social Media Safer Requires Meaningful Transparency,” Tech Policy Press, October 2, 2024, <https://www.techpolicy.press/making-social-media-safer-requires-meaningful-transparency/>.

³² Adam Marshall Writer, “Meta Fined for Collecting Data on Gay & Transgender Facebook Users,” Tech.co, November 5, 2024, <https://tech.co/news/meta-fine-south-korea>.

³³ A 2024 FTC report found that “Large social media and video streaming companies have engaged in vast surveillance of users with lax privacy controls and inadequate safeguards for kids and teens.”

³⁴ [YouTube's algorithm pushes right-wing, explicit videos regardless of user interest or age, study finds](#)

³⁵ Private messaging apps should [implement end-to-end encryption by default](#) to protect LGBTQ people from persecution, stalking, and violence.

³⁶ Lisa Schirch, Ravi Iyer, and Lena Slachmuis, “Toward Prosocial Tech Design Governance,” Toward Prosocial Tech Design Governance, December 21, 2023, <https://techandsocialcohesion.substack.com/p/toward-prosocial-tech-design-governance>.

³⁷ See for example:

[Meta's New Policies: How They Endanger LGBTQ+ Communities and Our Tips for Staying Safe Online; Meta's Content Moderation Changes are Going to Have a Real World Impact. It's Not Going to be Good.](#)

³⁸ [GLAAD Responds To Meta's Latest Anti-LGBTQ Changes To Content Policy and DEI That Will Harm Users](#)

³⁹ The term “transgenderism” is a well-known [right-wing trope](#) intended to falsely assert that being transgender is an ideology rather than an innate identity, and the term “homosexuality” is an outdated, [intentionally pathologizing way of referring to gay and lesbian people](#).

⁴⁰ Source: [Read: Meta's memo to employees rolling back DEI programs](#)

⁴¹ Source: [Google tells employees why it's ending DEI hiring goals](#)

⁴² Source: [Google Kills Diversity Hiring Targets](#)

⁴³ Targeted misgendering is a form of hate speech that involves the intentional use of the wrong gender and/or gender pronouns when referring or speaking to a transgender, nonbinary, or gender non-conforming person. Source: <https://glaad.org/releases/glaad-responds-twitters-roll-back-long-standing-lgbtq-hate-speech-policy/>

⁴⁴ Targeted deadnaming is a form of hate speech whereby a person intentionally “reveal[s] a transgender person’s former name without their consent – often referred to as ‘deadnaming’ – [which] is an invasion of privacy that undermines the trans person’s true authentic identity, and can put them at risk for discrimination, even violence.”

Source: <https://glaad.org/releases/glaad-responds-twitters-roll-back-long-standing-lgbtq-hate-speech-policy/>

⁴⁵ “Conversion therapy” is a widely condemned practice that involves any psychological or religious intervention aimed at changing an LGBTQ person’s sexual orientation, gender identity, or gender expression. Complicating efforts to address the amplification of harmful “conversion therapy” content online, its purveyors also promote this dangerous practice under alternate labels such as “leaving homosexuality” and “unwanted same-sex attraction.” Sources: <https://glaad.org/reference/conversion-therapy>; <https://globalextrémism.org/reports/conversion-therapy-online-the-ecosystem-in-2023/>