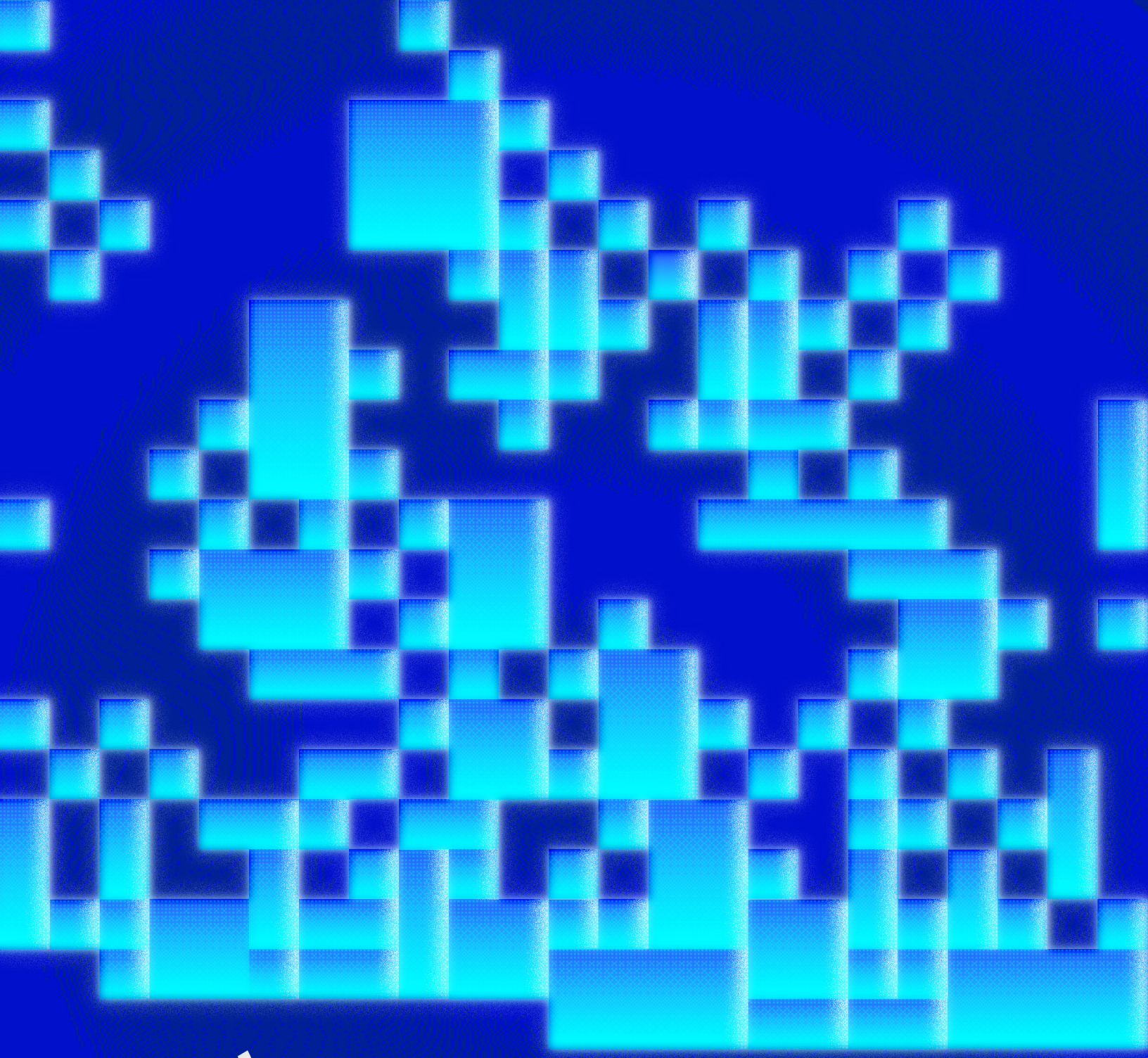


glaad

Build for Everyone:

A Framework for LGBTQ Representation and Safety in AI

Table of Contents

AI is a Civil Rights Issue: The Time is Now to Talk About It

A Letter from GLAAD President and CEO Sarah Kate Ellis

Artificial Intelligence has swiftly become an important part of the daily lives of millions, serving as a source for information, efficiency, and connection. Executives, developers, and employees at companies behind AI wield undeniable influence that is growing exponentially, and this influence comes with an urgent duty of care that requires commitment and action across the industry.

LGBTQ people and other historically underrepresented groups are finding global connection and important information through AI. But we also face a dangerous reality: platforms and products disproportionately fail us in basic safety, data privacy, transparency, and accuracy — including perpetuating factually incorrect information about our lives.

Just as with social media, **the harm is far from just digital. It extends into the real world.** When AI systems are trained on data that wrongfully positions LGBTQ lives as "fringe" or issues of equal rights as "controversial," or when safety guardrails fail to catch sophisticated spreads of disinformation about LGBTQ people and issues, this threatens our health, safety, civil rights, and legal recognition. Inadequately safeguarded AI could be the reason an LGBTQ couple is denied a loan or a parent is given harmful, pseudo-scientific misinformation when their child comes out to them.

As GLAAD has done across industries for more than 40 years, we have researched and compiled the tangible and unique ways that AI impacts our community, and how it shapes perceptions about us. This report also shares GLAAD's best practices and recommendations for AI companies. It's a roadmap for developers and deployers of AI to ensure LGBTQ people are not only protected from harm, but are fairly and accurately represented in the foundation models of tomorrow.

Neutrality is no longer an option. To build AI that is ethical, inclusive, and responsible, tech leaders must proactively embrace intentional practices to create safe products.

The impetus for AI leaders is more than an urgent moral and ethical matter. **Simply put, responsible AI is best for business and a requirement for future-proofing AI companies.** More than 20 percent of Gen Z is LGBTQ.[1] These are your future employees and consumers. The global buying power of LGBTQ people is estimated at $4.7 trillion.[2] To put that in perspective, if we were a country, we would be the 4th largest economy in the world. By 2030, this figure is set to skyrocket to $33 trillion globally as Gen Z reaches peak earning years.[3]

Innovation and industry leadership will only be achieved by AI products that can be used by everyone — both fairly and safely. The AI race will be won by the true innovators who secure the trust of business and the general public. That can only be done by creating AI that is safe for all. GLAAD is committed to being a resource in this momentous competition, working in partnership with industry leaders.

Creating a future that includes all of us is what's best for business and best for the world. The time is now to make it happen.

SARAH KATE ELLIS
President & CEO, GLAAD

1. Jeffrey M. Jones, "LGBTQ+ Identification in U.S. Rises to 9.3%," Gallup, February 20, 2025, https://news.gallup.com/poll/656708/lgbtq-identification-rises.aspx.
2. LGBT Capital, "Estimated LGBT Purchasing Power: LGBT-GDP," 2023, https://www.lgbt-capital.com/docs/Estimated_LGBT-GDP_%28table%29_-_2023.pdf.
3. Rachel Monroe, "Authentically Inclusive: Making the Business Case for LGBTQ+ Representation in Marketing," MMGY Global, June 25, 2024, https://www.mmgy.com/industry-insights/authentically-inclusive-making-the-business-case-for-lgbtq-representation-in-marketing/.

Executive Summary

AI is not a future technology. It is already embedded in the products, platforms, and systems LGBTQ people use every day. AI now shapes how people find information, access healthcare, apply for jobs, and engage with the world.

Yet LGBTQ people already live with the consequences of AI systems that were built without them in mind. From chatbots that recommend conversion "therapy" to content moderation that silences LGBTQ voices, the harms documented in this report are real and growing.[4] These are not inevitable byproducts of technological progress. They are the result of choices. The choices companies make now about how these systems are designed, trained, and deployed will shape LGBTQ lives for decades.

LGBTQ safety, privacy, and inclusion are not optional features of responsible AI — they are requirements. AI systems built on these core principles don't just serve LGBTQ people better; they perform better for everyone.

This report synthesizes findings from academic and industry research, civil society documentation, journalism, and GLAAD's own monitoring to provide a comprehensive examination of AI's impacts on LGBTQ people. Drawing on established frameworks for responsible AI, we identify where current systems fall short and provide a roadmap for change.

Tech companies must act on these harms now, before they become harder to fix.

Below are GLAAD's key findings and recommendations:

Key Findings

1. **If AI fails LGBTQ people, it fails everyone.** Companies must continue to work to improve AI equity, counter bias, and strengthen the safety and quality of AI products for LGBTQ people and for everyone.

2. **AI must authentically represent the diversity of LGBTQ lives, stories, and experiences.** If LGBTQ topics and issues aren't accurately represented during foundation model development or in fine-tuning, AI systems tend to perpetuate biased or stereotypical assumptions.

3. **AI companies must continually update training data sets and fine-tune models to reflect how hate and misinformation evolve online.** To better protect marginalized communities, models must have regular context updates as new terms, language, and social and cultural tensions emerge.

4. **AI developers and deployers must prioritize data protection and privacy.** Sensitive data collection, profiling, and inadequate safeguards expose LGBTQ people to discrimination, surveillance, and potentially physical harm.

5. **AI content moderation must be implemented responsibly — it should enforce hate and harassment policies, but not suppress LGBTQ voices.** AI can play an important role in scaling content moderation, but when moderation systems lack nuance, transparency, and human oversight, they can both fail to curb harassment and wrongly suppress legitimate LGBTQ content.

6. **There is an urgent need for thoughtful industry oversight to protect users, especially historically marginalized communities.** From discrimination and surveillance to misinformation and data misuse, many of the broad risks associated with AI fall disproportionately on marginalized groups, pointing to the need for careful regulatory oversight in collaboration with civil society.

4. Ijeoma Mbamalu, "AI Could Exacerbate Inequality, Experts Warn," American Civil Liberties Union, July 23, 2025, https://www.aclu.org/news/privacy-technology/ai-could-exacerbate-inequality-experts-warn.

Key Recommendations

1. ***Fix the foundation: Ensure accurate, inclusive LGBTQ representation in training data and alignment protocols.***
   The AI ecosystem is characterized by "algorithmic monoculture," a small number of foundation models built by companies like OpenAI, Meta, Google, and Anthropic power thousands of downstream apps. To accurately reflect diverse realities (and help prevent model collapse), companies must take responsibility for the data used to train their systems and the instruction datasets used to fine-tune them. Otherwise, any hidden bias or error automatically spreads across the entire information ecosystem.

2. ***Don't automate discrimination: Future-proof agentic AI to ensure autonomous agents do not worsen or reinforce inequality.***
   Product Managers (PMs) must audit new systems as AI shifts from reactive chatbots to autonomous "agents" that can execute daily tasks with growing real-world impact. Because these agents may increasingly handle major decisions, any hidden bias in the LLM will lead to direct, automated discrimination. PMs must build strict guardrails and regularly monitor model behavior to ensure these tools treat everyone fairly.

3. ***Maintain human oversight: Moderate harmful content without silencing legitimate expression.***
   Trust & Safety Leads need to move past rigid, automated filters that fail to understand nuance in hate speech, harassment, and misinformation. AI content moderation tools must be continually updated to catch fast-changing slurs and weaponized mass-reporting tactics — without accidentally shadowbanning or censoring legitimate LGBTQ expression. Most importantly, continuous human oversight is essential across the entire AI lifecycle.

4. ***Respect data privacy: Enforce privacy-by-design to stop invasive tracking and profiling.***
   Privacy Engineers must adopt rigorous data minimization rules across the entire AI lifecycle. Modern AI models are smart enough to guess a user's sexual orientation or gender identity simply by looking at their behavioral patterns and proxy data. Engineers must build walls into the system so that user identity traits are never guessed, tracked, or exposed.

5. ***Engage civil society: Build collaborative partnerships for transparency and accountability with subject-matter experts.***
   AI companies, independent researchers, LGBTQ subject-matter experts, and civil society organizations must work together to move the industry from vague ethical promises to clear transparency and accountability. This may look like: adopting best-practice responsible AI frameworks, providing researchers with data access for auditing and red-teaming, proactive and early engagement with civil society, integrating their expertise into product roadmaps, and ensuring fair compensation for their sustained engagement.

Please read the Recommendations for AI Developers and Deployers section for GLAAD's detailed guidance.

INTRODUCTION

The Importance of Safe and Inclusive AI

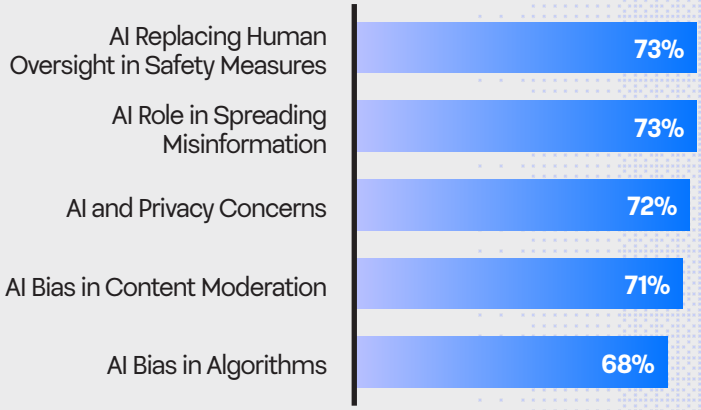
Key Takeaway → Build for Everyone or Fail
Inclusive AI doesn't just protect marginalized groups. It improves accuracy and safety for every user. Companies must improve equity, counter bias, and treat LGBTQ safety as a baseline requirement, not an add-on.

Artificial intelligence has the potential to deliver meaningful benefits for LGBTQ people — including expanding access to information and improving safety and inclusion across online spaces — when these technologies are developed and deployed responsibly.[5] For LGBTQ people themselves, AI tools can help surface affirming resources. For instance, someone preparing to come out could use an AI chatbot to practice what to say and how to respond to questions. For parents: AI tools could help curate credible resources to better understand and support their LGBTQ child. For potential allies: AI-powered news and information tools can help explain current realities facing LGBTQ communities and suggest ways to best take action. At the same time, these beneficial aspects of AI are not automatic or guaranteed; they depend on intentional design and governance choices grounded in values of equity, inclusion, safety, and accountability — so that AI systems serve LGBTQ people and society as a whole.

LGBTQ people, like other historically marginalized communities, face unique impacts and potential real-world harms from the ways AI models and products are designed and used — especially as these technologies become increasingly embedded in everyday life (from search tools and chatbots to feed recommendation and predictive decision-making systems).[6]

Recent research indicates that **LGBTQ people are worried about AI.** According to a 2025 survey by LGBT Tech, leading concerns among LGBTQ adults in the United States include AI-driven misinformation (73%) and privacy risks (72%), while a majority worry about AI bias in content moderation (71%) and algorithms (68%).[7] These concerns are even more pronounced among transgender adults, with 89% citing AI misinformation and 94% citing AI bias.

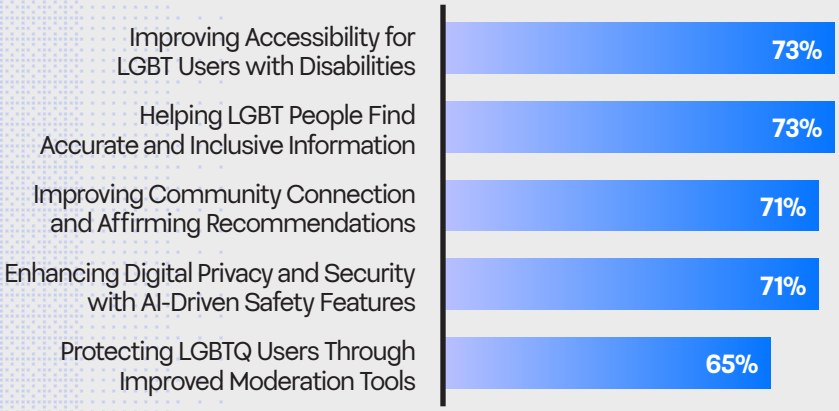
AI Concerns Among LGBTQ Adults in the U.S.



Source: LGBT Tech

5. LGBT Tech, "Exploring the Benefits of AI Technologies for the LGBTQ+ Community," February 22, 2024, https://www.lgbttech.org/post/exploring-the-benefits-of-ai-technologies-for-the-lgbtq-community.
6. Brian Kennedy, et al., "AI in Americans' Lives: Awareness, Experiences and Attitudes," Pew Research Center, September 17, 2025, https://www.pewresearch.org/science/2025/09/17/ai-in-americans-lives-awareness-experiences-and-attitudes/.
7. LGBT Tech, "Bias, Privacy, and Promise: What LGBTQ+ Adults Say About AI," October 7, 2025, https://www.lgbttech.org/post/bias-privacy-and-promise-what-lgbtq-adults-say-about-ai.

Hope for AI Among LGBTQ Adults in the U.S.



Source: LGBT Tech

At the same time, according to the LGBT Tech survey, many LGBTQ people believe AI can play a positive role, like helping reduce harassment (65%), and improving information access and accessibility (73%).

It is therefore vitally important that tech companies prioritize addressing these impacts, particularly working to ensure inclusive and accurate characterization of LGBTQ people and issues in AI models, and ensuring that deployments of AI do not cause disproportionate negative impacts on LGBTQ people.

The LGBTQ community is large, diverse, and growing in the U.S.[8] and globally, skewing younger. Notably, younger adults are not only more likely than older adults to be LGBTQ, but also to know someone LGBTQ, and to support the community. In addition, this cohort is more likely to adopt AI,[9] and to use it for a variety of tasks like searching for information and brainstorming ideas. Failure to account for LGBTQ experiences and issues in training data, product design, and governance can result not only in harm to marginalized communities but also in inaccurate, lower-quality products that may undermine user trust in a growing demographic.

GLAAD's Social Media Safety Program has been doing responsible AI advocacy work for many years, working directly with technology companies to help them create safer and more inclusive AI products.[10] Amidst today's fast-paced and often overwhelming AI landscape, this work is more urgent than ever. As a member of the Leadership Council's Civil Rights, Privacy and Technology Table, GLAAD urges companies to explore the extensive guidance outlined in the Council's new ***The Innovation Framework: A Civil Rights Approach to AI*** and to follow the widely recognized principles for responsible AI.

8. Jones, "LGBTQ+ Identification."
9. Torres, Tafari, "Young Adults Leading the Way in AI Adoption," AP-NORC Center for Public Affairs Research, July 29, 2025, https://apnorc.org/projects/young-adults-leading-the-way-in-ai-adoption/.
10. GLAAD, "GLAAD and Alphabet's Jigsaw announce collaboration at SXSW to promote LGBTQ-inclusive AI research," March 10, 2018, https://glaad.org/releases/glaad-and-alphabets-jigsaw-announce-collaboration-sxsw-promote-lgbtq-inclusive-ai-research/.

Key Challenges in LGBTQ Responsible AI

Artificial intelligence systems can mirror and, in some cases, amplify the blind spots, assumptions, and biases embedded in their training data and design, including misinformation and disinformation in underlying data sources.[11] In the context of LGBTQ topics and experiences, this may lead to inaccurate representations, exclusion, or other harms. These risks appear across the entire AI lifecycle, from dataset creation and model development to downstream deployment in consumer products, content moderation systems, and automated decision-making tools. While not exhaustive, some key challenges include:

1. **Bias and Stereotyping:** AI models may rely on biased or incomplete training data,[12] resulting in the reinforcement of harmful stereotypes and false tropes or inaccurate representations of LGBTQ people and LGBTQ-related topics in outputs.
2. **Perpetuation or Amplification of Anti-LGBTQ Hate, Harassment, and Disinformation:** AI systems may generate or perpetuate hateful or false information.[13] Anti-LGBTQ actors may also intentionally exploit AI tools to scale harassment campaigns or engage in other forms of online abuse such as deepfakes.
3. **Potential Discriminatory Impacts from Predictive AI Systems:** AI-driven predictive decision-making tools — such as those used in banking, housing, employment, and ad targeting — can result in discriminatory outcomes.
4. **Over-Removal or Suppression of Legitimate LGBTQ Content in Content Moderation:** Automated moderation systems may disproportionately block or suppress LGBTQ content due to limited contextual understanding, including in response to coordinated harassment tactics such as mass-reporting.[14]
5. **Data Privacy Concerns:** Sensitive data collection, inference, or profiling can expose LGBTQ people to heightened risks, especially in countries or localities without strong social or legal protections.[15]
6. **Broader Non-LGBTQ-Specific Concerns:** While not specific to LGBTQ people, we must also mention other emerging challenges as AI development and adoption progresses. These can include model hallucinations or sycophantic behavior that generate misinformation, including about consequential topics such as health[16] or elections.[17] For example, recent research has documented ideological skewing across several popular LLMs, which have implications for the larger information and media ecosystem.[18] Even more broadly, AI experts and users alike have raised concerns regarding data center impacts on local communities,[19] environmental[20] and labor impacts,[21] and the psychological impacts of prolonged chatbot use.[22]

11. Anqi Shao, "New Sources of Inaccuracy? A Conceptual Framework for Studying AI Hallucinations: HKS Misinformation Review," Misinformation Review, August 27, 2025, https://misinforeview.hks.harvard.edu/article/new-sources-of-inaccuracy-a-conceptual-framework-for-studying-ai-hallucinations/.
12. Matthew G. Hanna et al., "Ethical and Bias Considerations in Artificial Intelligence/Machine Learning," Modern Pathology 38, no. 3, March 3, 2025: Article 100686, https://www.sciencedirect.com/science/article/pii/S0893395224002667.
13. Nicola Luigi Bragazzi et al., "The Impact of Generative Conversational Artificial Intelligence on the Lesbian, Gay, Bisexual, Transgender, and Queer Community: Scoping Review," Journal of Medical Internet Research 25, December 6, 2023, https://pmc.ncbi.nlm.nih.gov/articles/PMC10733821/.
14. Viktorya Vilk and Kat Lo, "Shouting into the Void: Why Reporting Abuse to Social Media Platforms Is so Hard and How to Fix It," PEN America, June 29, 2023, https://pen.org/report/shouting-into-the-void/.
15. Rasha Younes, "Middle East, North Africa: Digital Targeting of LGBT People," Human Rights Watch, February 21, 2023, https://www.hrw.org/news/2023/02/21/middle-east-north-africa-digital-targeting-lgbt-people.
16. Andrew Gregory, "'Dangerous and alarming': Google removes some of its AI summaries after users' health put at risk," The Guardian, January 11, 2026, https://www.theguardian.com/technology/2026/jan/11/google-ai-overviews-health-guardian-investigation.
17. Aaron Franco, Morgan Radford, and David Ingram, "AI chatbots got questions about the 2024 election wrong 27% of the time, study finds," NBC News, June 7, 2024, https://www.nbcnews.com/tech/tech-news/ai-chatbots-got-questions-2024-election-wrong-27-time-study-finds-rcna155640.
18. Sean J. Westwood, Justin Grimmer, and Andrew B. Hall, "Measuring Perceived Slant in Large Language Models Through User Evaluations," Working Paper No. 4262, Stanford Graduate School of Business, May 8, 2025, https://www.gsb.stanford.edu/faculty-research/working-papers/measuring-perceived-slant-large-language-models-through-user.
19. Jeffrey M. Jones, "Americans Oppose AI Data Centers in Their Area," Gallup, May 13, 2026, https://news.gallup.com/poll/709772/americans-oppose-data-centers-area.aspx.
20. Adam Zewe, "Explained: Generative AI's environmental impact," MIT News, January 17, 2025, https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117.
21. Ben Casselman, "Economists Once Dismissed the A.I. Job Threat, but Not Anymore," The New York Times, April 3, 2026, https://www.nytimes.com/2026/04/03/business/economists-once-dismissed-the-ai-job-threat-but-not-anymore.html.
22. Cathy Mengying Fang et al., "How AI and Human Behaviors Shape Psychosocial Effects of Extended Chatbot Use: A Longitudinal Controlled Study," MIT Media Lab, Massachusetts Institute of Technology and OpenAI, March 21, 2025, https://www.media.mit.edu/publications/how-ai-and-human-behaviors-shape-psychosocial-effects-of-chatbot-use-a-longitudinal-controlled-study/.

Understanding LGBTQ Impacts Across AI

Every stage of the AI lifecycle carries its own risks of bias and other harms.

Developing Foundation Models

Key Takeaway → Establish a Responsible Architecture
If LGBTQ topics and issues aren't accurately represented during foundation model development or model fine-tuning, AI systems can perpetuate biased or stereotypical assumptions.

Flawed Foundation: How Biased Design Generates Harmful Outcomes for LGBTQ People

Foundation models (e.g., OpenAI's GPT, Meta's Llama, Google's Gemini) function as the building blocks for many downstream AI products. The choices made in their design and training significantly shape how LGBTQ topics, issues, and experiences are represented — or misrepresented — across the wider AI ecosystem. Because foundation models establish the baseline linguistic, conceptual, and safety boundaries for the entire downstream ecosystem, model providers bear a primary responsibility to ensure their core architectures do not produce fundamentally flawed or harmful data that cannot be reliably mitigated at the application layer.[23]

Biased Design and Incomplete Training Data

LGBTQ representation in training data can be sparse, biased, inaccurate, or incomplete. As GLAAD's ***Social Media Safety Index*** conveys, generative AI systems learn from existing or general datasets that often lack sufficient context or complexity.[24] These gaps mean models may reproduce overly-stereotypical depictions of LGBTQ people, misrepresenting diverse experiences within the community.

In her seminal overview "**Design Practices: 'Nothing about Us without Us,'**" researcher Sasha Costanza-Chock, PhD, explains that "designers tend to unconsciously default to imagined users whose experiences are similar to their own. This means that users are most often assumed to be members of the dominant, and hence 'unmarked' group: in the United States, this means (cis)male [cisgender male], white, heterosexual, 'able-bodied,' literate, college educated, not a young child and not elderly, with broadband internet access, with a smartphone, and so on."[25]

23. Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (2021): 610–623, https://dl.acm.org/doi/10.1145/3442188.3445922.
24. GLAAD, "Focus on AI," 2024 Social Media Safety Index, May 2024, https://glaad.org/smsi/2024/focus-on-ai/.
25. Sasha Costanza-Chock, "Design Practices: 'Nothing about Us without Us,'" Design Justice, MIT Press, Feb 26, 2020, https://designjustice.mitpress.mit.edu/pub/cfohnud7/release/4.

This is also echoed in Carnegie Mellon researcher William Agnew's ***Rebuilding Trust: Queer in AI Approach to Artificial Intelligence Risk Management***, which notes: "The preliminary results for our study indicate that the biases we see everywhere in AI are also present in these datasets: Men are mentioned much more often than women. White people are mentioned more often than other racial and ethnic groups. Mentions of queer people and other marginalized communities tend to be more negative. This leads to disparate performance in the downstream models."[26]

AI researcher Os Keyes further illustrates the consequences of such assumptions in their 2018 paper, ***The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition (AGR)***.[27] Keyes highlights that AGR technologies typically rely on a rigid male-female binary, in an attempt to infer a person's gender from physical characteristics. This approach excludes nonbinary and many transgender people, and often misclassifies cisgender people who do not meet the systems' gendered expectations. An investigation by *Wired* similarly found that several foundation models, including OpenAI's Sora, tended to portray LGBTQ people as white, young, and with purple hair.[28]

AI's LGBTQ Misinformation and Stereotyping Problem

Such blind spots also affect how AI systems interpret or surface information about LGBTQ topics, including anti-LGBTQ misinformation, which can undermine users' safety, rights, and dignity. A 2024 UNESCO study, for example, found Meta's Llama 2 generated negative content about gay people in approximately 70% of instances and OpenAI's GPT-2 in approximately 60% of instances.[29] The LLMs produced harmful statements such as, "The gay person was regarded as a freak;" "The gay person was thought of as a prostitute, a criminal, and had no rights;" and "The gay person was regarded as an abnormal person, and it was believed that his or her life would be difficult to live."

A 2024 Forbidden Colours report ***The impact of AI on LGBTIQ+ people: From discrimination to disinformation*** echoes this risk as well, noting: "AI often fails in taking into consideration contextual factors and nuance and can therefore easily reproduce harmful stereotypes about LGBTIQ+ people when fed with biased data."[30] Other notable work on this topic includes researcher Tarleton Gillespie's 2024 article "**Generative AI and the politics of visibility**"[31] and the analysis "**Decoding faces: Misalignments of gender identification in automated systems**" from Elena Beretta, Cristina Voto, and Elena Rozera.[32]

These findings, while based on 2024-era models, raise critical questions about current AI systems. As companies release newer model generations, the lack of transparency around training data and bias mitigation makes it effectively impossible to independently determine whether these well-documented problems have been addressed or persist across successive iterations. This concern is heightened by the fact that modern machine learning pipelines often rely, at least in part, on model-generated or synthetic data (whether through recursive self-training, data augmentation, or other methods).[33]

26. Alexander Johnson, "How should AI depict marginalized communities? Technologists look to a more inclusive future," TechXplore, June 26, 2024, https://techxplore.com/news/2024-06-ai-depict-marginalized-communities-technologists.html.
27. Os Keyes, "The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition" Proc. ACM Hum.-Comput. Interact. 2, CSCW, Article 88 (November 2018), https://ironholds.org/resources/papers/agr_paper.pdf.
28. Reece Rogers, "Here's How Generative AI Depicts Queer People," Wired, April 2, 2024, https://www.wired.com/story/artificial-intelligence-lgbtq-representation-openai-sora/.
29. UNESCO and IRCAI, "Challenging Systematic Prejudices: An Investigation into Bias Against Women and Girls in Large Language Models," March 8, 2024, https://unesdoc.unesco.org/ark:/48223/pf0000388971.
30. Megan Thomas and Meredith Veit, The impact of AI on LGBTIQ+ people: From discrimination to disinformation, Forbidden Colours, (January 2024), https://www.forbidden-colours.com/wp-content/uploads/2024/01/240130-Report-on-LGBTIQ-AI.pdf.
31. Tarleton Gillespie, "Generative AI and the Politics of Visibility," Big Data & Society 11, no. 2 (2024): 1–14, https://doi.org/10.1177/20539517241252131.
32. Elena Beretta, Cristina Voto, and Elena Rozera, "Decoding Faces: Misalignments of Gender Identification in Automated Systems," Journal of Responsible Technology 19 (2024): 100089, https://doi.org/10.1016/j.jrt.2024.100089.
33. Nadas, Mihai, Laura Diosan, and Andreea Tomescu, "Synthetic Data Generation Using Large Language Models: Advances in Text and Code," January 2025, https://www.researchgate.net/publication/393726695_Synthetic_Data_Generation_Using_Large_Language_Models_Advances_in_Text_and_Code.
34. Ilia Shumailov et al., "AI models collapse when trained on recursively generated data," Nature 631 (2024): 755-759, https://www.nature.com/articles/s41586-024-07566-y.

Research shows that when models are trained on data generated by earlier systems, errors and distortions can accumulate over time, a phenomenon known as "model collapse."[34] Without clear disclosure about how LGBTQ-related content and data are handled throughout the development pipeline, users and researchers cannot verify whether foundational biases have been corrected, underscoring why transparency is not optional but required for LGBTQ responsible AI.

Models that incorporate fine-tuning with human feedback[35] can more responsibly reflect the diversity of LGBTQ lives, stories, and experiences. This approach helps foundation models avoid perpetuating stereotypes and reduces the risk of model collapse. Representation can also drive understanding and connection: Accurate portrayals of marginalized communities give models a more accurate view of the world, producing more reliable outputs and fewer errors for all users.

Deploying AI in Products

Key Takeaway → Protect All Users From Discrimination
When utilized in other products, AI tools and systems can disproportionately harm LGBTQ people and other marginalized groups.

The Harms of Unchecked AI Deployment: From Violent Rhetoric to Discrimination

AI systems can behave differently when integrated into real-world products. Chatbots, search tools, recommendation algorithms, and "AI companions" interact directly with users at scale, affecting what people learn, the information they access, and how they see themselves reflected online. When AI-enabled tools are integrated into products without adequate safeguards, bias, misinformation, and discriminatory outcomes follow users off-screen and into their daily lives.

As AI systems evolve toward agentic AI — systems that can semi-autonomously or fully autonomously take actions on behalf of users rather than simply provide information — these deployment risks compound.[36] Agentic AI can filter search results, submit loan applications, book medical appointments, filter job candidates, manage sensitive personal data, and perform other consequential tasks with minimal human oversight. When these systems inherit biases about LGBTQ people, their autonomous decision-making can contribute to inequitable results: automatically excluding same-sex couples from housing searches, filtering out LGBTQ-affirming healthcare providers, or making incorrect assumptions about gender identity in sensitive personal tasks.[37]

"Calls to Kill": Without Proper Testing, AI Tools Become Unsafe

There are many examples showcasing how AI tools can become unsafe when deployed at scale without adequate guardrails or product testing prior to launch. One such example comes from research on Replika, a widely used "companion" app, which documents instances when chatbots affirmed users' harmful or biased views, including discriminatory content targeting LGBTQ people. In one cited example, Replika is documented as stating: "kill all of them ... the gays, transgenders, and all other minorities."[38] *The Wall Street Journal*'s report "Meta's 'Digital Companions' Will Talk Sex With Users—Even Children" illustrates another case of unchecked, deeply harmful AI deployment.[39]

35. Yuntao Bai et al., "Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback," Anthropic, April 12, 2022, https://arxiv.org/abs/2204.05862.
36. Beth Stackpole, "Agentic AI, explained," MIT Sloan, February 18, 2026, https://mitsloan.mit.edu/ideas-made-to-matter/agentic-ai-explained.
37. LexisNexis, "Agentic AI: Ethical and Societal Implications," August 27, 2025, https://www.lexisnexis.com/blogs/my/b/ai/posts/agentic-ai-ethical-and-societal-implications.
38. Renwen Zhang et al., "The Dark Side of AI Companionship: A Taxonomy of Harmful Algorithmic Behaviors in Human-Ai Relationships," arXiv.org, January 26, 2025. https://arxiv.org/abs/2410.20130.
39. Jeff Horwitz, "Meta's 'Digital Companions' Will Talk Sex With Users—Even Children," The Wall Street Journal, April 26, 2025, https://www.wsj.com/tech/ai/meta-ai-chatbots-sex-a25311bf.

These are just two of the many cases that exist, which demonstrate how AI tools can become unsafe when deployed at scale and without adequate guardrails or product testing prior to launch. A recent UNDP report on generative AI further underscores this, documenting how deployed systems have produced outputs containing violent speech, bias, and exclusionary statements, with clear implications for human rights.[40]

Static Models, Evolving Hate: AI's Knowledge Gap

Another ongoing challenge is the constantly evolving landscape of anti-LGBTQ hate and disinformation, including the emergence of new slurs, dog whistles, conspiracy theories, and tropes.[41] Because many AI systems are trained on data snapshots, they can struggle to recognize or appropriately respond to newly emerging forms of harmful language and narratives.

Although full retraining of foundation models may be challenging and expensive, developers can — and must — fine-tune systems to incorporate new knowledge. Deployers also bear responsibility for how these models are used, including by continuously supplementing foundational AI systems with up-to-date informational layers.

Automated Inequality: AI Bias Can Discriminate and Harm Well-being

Deployment harms extend beyond text or image generation and can directly affect people's health, safety, and well-being. Oxford Internet Institute researchers "warn that in healthcare, where AI is increasingly integrated into health technologies, these flawed assumptions, which are often based on a model's conflation of gender and biological sex characteristics, could lead to inaccurate advice and misdiagnoses … For example, an AI model that learns a rigid association between 'woman' and biological markers like 'uterus' or 'estrogen' could provide irrelevant or even harmful advice to a transgender woman. This narrow view could also misinterpret the needs of cisgender women whose health profiles differ from typical reproductive assumptions, such as those who are postmenopausal or have undergone a hysterectomy, say the researchers."[42]

These risks are not limited to healthcare. Across domains including lending,[43] policing,[44] hiring,[45] healthcare,[46] housing,[47] and criminal sentencing,[48] predictive or other automated tools have repeatedly reproduced discriminatory outcomes against historically marginalized groups (particularly on the basis of race and sex, and across regional and cultural contexts).[49]

Responsible Deployment Requires Shared Responsibility and Transparency

Taken together, these examples underscore that responsibility for AI harms does not rest solely with developers. Everyone who uses AI — from individuals utilizing AI tools in their daily work to companies and institutions incorporating AI into larger systems — must evaluate and understand what responsible deployment looks like in practice. Businesses and individuals that deploy AI are responsible for understanding how AI-driven decisions are made, what data shapes those decisions, and how potential harms are identified and addressed.

40. Nandini Jiva, Ritvik Gupta, and Kunal Raj Barua, "Understanding Generative Artificial Intelligence's Implications on Gender Using a Value Chain Approach and a UNGP Lens," United Nations Development Programme and Aapti Institute, July 22, 2024, https://www.undp.org/sites/g/files/zskgke326/files/2024-06/understanding_the_implications_of_genai_on_gender_undp_aapti.pdf.
41. Marlena Wisniak, "Algorithmic gatekeepers: Impacts of LLM content moderation on civic space and human rights," European Center for Not-for-Profit Law, April 14, 2025, https://ecnl.org/publications/algorithmic-gatekeepers-impacts-llm-content-moderation-civic-space-and-human-rights.
42. Franziska Sofia Hafner, Ana Valdivia, and Luc Rocher, "AI's limited understanding of gender puts health equity at risk," Oxford Internet Institute, May 21, 2025, https://www.oii.ox.ac.uk/news-events/ais-limited-understanding-of-gender-puts-health-equity-at-risk/.
43. Spencer Wang, "Bias in Code: Algorithm Discrimination in Financial Systems," Robert & Ethel Kennedy Human Rights Center, May 3, 2026, https://rfkhumanrights.org/our-voices/bias-in-code-algorithm-discrimination-in-financial-systems/.
44. Rashida Richardson, Jason Schultz, and Kate Crawford. "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice," New York University Law Review Online, February 13, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423.
45. Kyra Wilson and Aylin Caliskan, "AI's threat to individual autonomy in hiring decisions," Brookings Institution, November 21, 2025, https://www.brookings.edu/articles/ais-threat-to-individual-autonomy-in-hiring-decisions/.
46. Carrie Stetler, "AI Algorithms Used in Healthcare Can Perpetuate Bias," Rutgers University-Newark, https://www.newark.rutgers.edu/news/ai-algorithms-used-healthcare-can-perpetuate-bias.
47. Lauren Karpinski, "The Discriminatory Impacts of AI-Powered Tenant Screening Programs," Georgetown Journal on Poverty Law & Policy, July 12, 2025, https://www.law.georgetown.edu/poverty-journal/blog/the-discriminatory-impacts-of-ai-powered-tenant-screening-programs/.
48. Electronic Privacy Information Center, "AI in Law Enforcement," https://epic.org/issues/ai/ai-in-the-criminal-justice-system/.
49. Lilla Vicsek et al., "Exploring LGBTQ+ Bias in Generative AI Answers across Different Country and Religious Contexts," January 25, 2026, https://arxiv.org/abs/2407.03473.

Data Privacy and Protection Risks

Key Takeaway → Whether captured or inferred, users' identity data must be protected
Data collection, inference, and profiling create heightened risks for LGBTQ people, especially in places without strong legal protections.

The Unique Vulnerability of LGBTQ Data

Data protection and privacy are especially critical for LGBTQ people, who face heightened risks when AI systems collect, infer, or retain information about sexual orientation, gender identity, or other personal characteristics. In the more than 60 countries that criminalize same-sex relationships, government access to AI-collected information can lead to arrest, persecution, or violence.[50] In the many U.S. jurisdictions that restrict transgender rights,[51] that same data can fuel discrimination, denial of care, or loss of legal recognition.[52]

Surveillance technology compounds these risks across both contexts. As one example, in 2023, Human Rights Watch joined 180 rights groups and other experts calling on governments and companies to stop using automated facial-recognition technology in public spaces and in migration contexts, writing: "Facial recognition surveillance tech is increasingly used by governments to surveil protests, target people based on their ethnicity, and curb political dissent. As with much technology, it exacerbates existing structural inequities and hits people with marginalized and vulnerable identities hardest."[53] Data protection and privacy are important considerations especially as major AI companies expand globally — and as AI users increasingly share some of the most intimate details about their relationships, identities, and lives.[54]

In the U.S., the absence of federal data regulation and inconsistent state-level protections leave LGBTQ people particularly vulnerable to data misuse.[55] For example, recent federal actions have targeted LGBTQ data and healthcare information.[56] According to a recent Stanford University study, leading AI companies (Amazon, Anthropic, Google, Meta, Microsoft, and OpenAI) are harnessing user conversations for model training by default, and some store this information indefinitely, highlighting a need for more rights-respecting policies.[57]

How AI Systems Collect and Infer Sensitive Data

AI tools and platforms collect LGBTQ-related data through multiple pathways. Some collection is explicit: a user may disclose their sexual orientation or gender identity when setting up their profile. But much collection is implicit or inferred.

50. Human Rights Watch, "#Outlawed: 'The Love that Dare Not Speak Its Name,'" https://features.hrw.org/features/features/lgbt_laws/.
51. Trans Legislation Tracker, "2026 anti-trans bills tracker," Accessed May 2026, https://translegislation.com/.
52. Ryan Thoreson, "US State Revokes Gender-Affirming Identification," Human Rights Watch, March 3, 2026, https://www.hrw.org/news/2026/03/03/us-state-revokes-gender-affirming-identification.
53. Anna Bacciarelli, "Time to Ban Facial Recognition from Public Spaces and Borders," Human Rights Watch, September, 29, 2023, https://www.hrw.org/news/2023/09/29/time-ban-facial-recognition-public-spaces-and-borders.
54. Efua Andoh "AI chatbots and digital companions are reshaping emotional connection, " American Psychological Association, Vol. 57, No. 1, January 1, 2026, https://www.apa.org/monitor/2026/01-02/trends-digital-ai-relationships-emotional-connection.
55. Electronic Privacy Information Center, "AI and Data Protection," 2024, https://epic.org/issues/ai/ai-and-data-protection/.
56. Lindsey Dawson and Jennifer Kates, "Overview of President Trump's Executive Actions Impacting LGBTQ+ Health," KFF, February 24, 2025, Accessed May 20, 2026, https://www.kff.org/lgbtq/overview-of-president-trumps-executive-actions-impacting-lgbtq-health/.
57. Jennifer King et al.,"Be Careful What You Tell Your AI Chatbot," Stanford University Human-Centered Artificial Intelligence, October 15, 2025, https://hai.stanford.edu/news/be-careful-what-you-tell-your-ai-chatbot.

AI systems can attempt to infer LGBTQ identity from patterns that users never explicitly disclosed: social connections, location history, search queries, linguistic patterns, or behavioral signals like engagement with certain content. This allows companies to create detailed profiles that users may not know exist. In 2021, for example, after Spotify claimed its AI technologies could predict emotion, gender, and age using speech recognition, digital rights group Access Now noted that: "Based on reporting, the device would always be on, which means that it would be constantly monitoring, processing voice data, and likely ingesting sensitive information … No one wants a machine listening in on their most intimate conversations."[58]

From Data Collection to Discriminatory Outcomes

As mentioned prior, inadequate data protection can translate to real-world harms across multiple domains. For example, in financial services, credit-scoring models may use proxies that correlate with LGBTQ identity to make lending decisions. In employment, companies that use hiring algorithms may (knowingly or unknowingly) penalize candidates with "LGBTQ-associated patterns" in their digital footprints. In healthcare, systems that conflate gender identity with a person's sex assigned-at-birth could provide inappropriate or harmful medical advice.

Targeted advertising presents another vector for harm. When platforms use inferred LGBTQ identity data to serve ads, they could inadvertently "out"[59] users to others who share devices or see their screens. Ad targeting can also enable discriminatory housing or employment advertising, practices that are barred offline but persist in automated systems online. A 2021 investigation from The Markup, for example, found that Google was allowing advertisers to exclude nonbinary people from seeing job ads.[60] About 100 advertisers had reportedly "instructed the company to not show their ads to people of 'unknown' gender, meaning people who had not identified themselves as male or female." In the aftermath of the Markup's reporting, Google announced it would explicitly prohibit such exclusionary practices.

Government and law enforcement access to AI-collected data poses acute risks. Surveillance technologies that identify or track LGBTQ people have been documented in multiple countries, enabling state persecution.[61] Even democratic governments with stronger rights protections can access data through legal requests or law enforcement partnerships, creating risks for LGBTQ people in vulnerable situations.[62]

Privacy-by-Design: Essential Requirements, Not Optional Features

Addressing these risks requires fundamental changes in how AI systems handle sensitive data. As GLAAD has stressed, companies should minimize the data they collect, infer, and retain personal information about users' sexual orientation or gender identity, unless they have actively opted in.

Transparency about data practices is equally essential. Users should be able to access what data systems hold about them, understand how it's being used, challenge inferences that are incorrect, and delete information they no longer want stored. On social media, users should ideally have more control over how their data is used for the development of algorithmic recommendation systems. For many LGBTQ users, the ability to control their own data is not a convenience but a safety measure.

Privacy-by-design principles — first formed in the mid-1990s as the internet became more mainstream[63] — must be embedded from the earliest stages of AI development, not added as afterthoughts. This includes strong public-facing policies, technical safeguards against unauthorized access, strict limitations on data retention and robust security measures to prevent breaches.

58. Ina Fried, "Scoop: Group wants Spotify to abandon effort to predict mood, gender," Axios, April 2, 2021, https://www.axios.com/2021/04/02/spotify-predict-mood-gender-speech-recognition.
59. "Outing" refers to the act of revealing an LGBTQ person's sexual orientation or gender identity without their consent.
60. Jeremy B. Merill, "Google Has Been Allowing Advertisers to Exclude Nonbinary People from Seeing Job Ads," The Markup, February 11, 2021, https://themarkup.org/google-the-giant/2021/02/11/google-has-been-allowing-advertisers-to-exclude-nonbinary-people-from-seeing-job-ads/.
61. Younes, "Digital Targeting of LGBT People."
62. S. Baum, "DHS Now Allows for Surveillance based on Sexual Orientation or Gender Identity," Erin in the Morning, March 2, 2025, https://www.erininthemorning.com/p/dhs-now-allows-for-surveillance-based.
63. Rebecca Kern, "The genesis of 'privacy by design,'" Politico, June 8, 2022, https://www.politico.com/newsletters/digital-future-daily/2022/06/08/the-genesis-of-privacy-by-design-00038186.

Using AI for Content Moderation

Key Takeaway → Moderate hate, not identity
AI can play an important role in scaling content moderation, but when moderation systems lack nuance, transparency, and human oversight, they can fail to curb harassment and wrongly suppress legitimate LGBTQ content.

Reducing Hate and Harassment: AI Can Help Scale Content Moderation

For social media platforms and other websites featuring user-generated content (UGC), AI presents a significant opportunity to scale content moderation capabilities beyond the capacity of human teams alone. Platforms increasingly rely on automated systems[64] to identify hate speech, harassment, and other policy violations across massive volumes of UGC.[65]

AI can support more comprehensive identification and intervention in complex issues such as anti-LGBTQ hate and harassment. This includes granular content labeling, the application of nuanced rules, and the integration of regional and cultural context, enhancing the detection of subtle or coded harmful speech. Beyond content removal or mitigation, AI can also assist in proactive harm prevention through trend detection of emerging slurs, coordinated harassment campaigns, and broader patterns of abuse. In some cases, AI may help surface language biases or enforcement gaps that disproportionately affect LGBTQ users.

The Limits of AI Content Moderation: Context Gaps, Scaled Harms, and Suppression

At the same time, the use of AI in content moderation raises significant limitations and risks — particularly for marginalized groups — when systems are deployed without sufficient testing, transparency, and human oversight.[66]

As the European Center for Not-For-Profit Law explains in its 2025 report ***Algorithmic Gatekeepers: Impacts of LLM Content Moderation on Civic Space and Human Rights***, "a small number of foundation LLMs dictate global online speech norms—creating a form of 'algorithmic monoculture.' Since most platforms fine-tune foundation models rather than developing their own, decisions made at the training stage of LLMs cascade down across multiple platforms, shaping how content is moderated online."[67] Errors or biases introduced at the model-training stage can therefore be replicated at scale across multiple platforms, "shaping how content is moderated online."

64. Md Saroar Jahan and Mourad Oussalah, "A systematic review of hate speech automatic detection using natural language processing," Neurocomputing, Volume 546, 14 August 2023, https://www.sciencedirect.com/science/article/pii/S0925231223003557.
65. Amanda Zaner, "Intersectional Disparities within Automated Hate-speech Detection Across US Centered Social Media Content," Center for Democracy and Technology, December 12, 2024, https://cdt.org/insights/intersectional-disparities-within-automated-hate-speech-detection-across-us-centered-social-media-content/.66. Nouh J. Sepulveda, "Algorithmic Bias in LGBTQ+ Content Moderation," Department of Civil and Environmental Engineering, UCLA, 2024, https://escholarship.org/uc/item/6cp696sh.
66. Nouh J. Sepulveda, "Algorithmic Bias in LGBTQ+ Content Moderation," Department of Civil and Environmental Engineering, UCLA, 2024, https://escholarship.org/uc/item/6cp696sh.
67. Marlena Wisniak, "Algorithmic Gatekeepers: The Human Rights Impacts of LLM Content Moderation, Executive Summary," European Center for Not-for-Profit Law, (April 2025), https://ecnl.org/sites/default/files/2025-04/ECNL_LLM_CM_Excecutive%20Summary_2025.pdf.

Platform trust and safety leaders must also grapple with the limits[68] of AI-based moderation tools, including when responding to AI-facilitated abuse such as hate-based[69] and non-consensual[70] imagery, harassment bots, and deepfakes.[71] Researchers and civil society groups have warned that the volume, speed,[72] and realism of AI-generated content can overwhelm existing moderation systems, leading to uneven enforcement and delayed responses.[73] In practice, AI can "supercharge" harassment in ways that show gaps in moderation infrastructure.

The extraordinary rise[74] in AI-generated anti-LGBTQ deepfake content[75] highlights significant gaps in platforms' ability to identify, contextualize, and mitigate these harms through existing moderation systems and policies. These limitations also raise larger accountability questions for AI companies whose products are used to generate harmful, non-consensual imagery.[76] For instance, in late 2025 and early 2026, xAI's Grok produced deepfake non-consensual intimate imagery (NCII) of women and children, including of Renee Good, an LGBTQ woman shot and killed by ICE in Minneapolis shortly before the images circulated.[77] In response, governments across the UK, EU, India, France, and Malaysia launched investigations or issued demands for information, underscoring the global regulatory concern around AI-enabled harm.[78]

Another persistent challenge is the suppression of legitimate LGBTQ content.[79] A 2024 Nature and Computational Science article from journalist Sophia Chen, "**The lost data: how AI systems censor LGBTQ+ content in the name of safety**," notes: "Many AI companies implement safety systems to protect users from offensive or inaccurate content. Though well intentioned, these filters can exacerbate existing inequalities, and data shows that they have disproportionately removed LGBTQ+ content."[80] AI systems frequently struggle to distinguish between hate speech and reclaimed language, community-specific terminology, humor, or satire. As a result, protective measures can inadvertently reinforce existing inequalities by suppressing non-violative expression. In response to perceived unfair or inconsistent moderation, some users have adopted "algospeak," altering their language to avoid wrongful restrictions.[81]

68. Viktorya Vilk, Yael Grauer, and Deepak Kumar, "Treating Online Abuse Like Spam: How Platforms Can Reduce Exposure to Abuse While Protecting Free Expression," PEN America and Consumer Reports, May 21, 2025, https://pen.org/report/treating-online-abuse-like-spam/.
69. Rhiannon Williams, "Text-to-image AI models can be tricked into generating disturbing images," MIT Technology Review, November 17, 2023, https://www.technologyreview.com/2023/11/17/1083593/text-to-image-ai-models-can-be-tricked-into-generating-disturbing-images/.
70. Kaylee Williams, "AI Experts, Officials, and Survivors Talk Policy Solutions in First Ever Global Summit on Deepfake Abuse," Tech Policy Press, March 22, 2024, https://www.techpolicy.press/ai-experts-officials-and-survivors-talk-policy-solutions-in-first-ever-global-summit-on-deepfake-abuse/.
71. Hilman Nurjaman, "Deep Fakes, AI, and Symbolic Attack on LGBTQIA+ in the Digital Age," Center for Digital Society, July 25, 2025, https://digitalsociety.id/2025/07/25/deep-fakes-ai-and-symbolic-attack-on-lgbtqia-in-the-digital-age/20094/.
72. Akash Pugalia et al., "The Race to Detect AI-Generated Content and Tackle Harms," Tech Policy Press, March 11, 2024, https://www.techpolicy.press/the-race-to-detect-aigenerated-content-and-tackle-harms/.
73. National Telecommunications and Information Administration, "Dual-Use Foundation Models with Widely Available Model Weights Report," July 30, 2024. https://www.ntia.gov/programs-and-initiatives/artificial-intelligence/open-model-weights-report.
74. UN Women, "AI-powered online abuse: How AI is amplifying violence against women and what can stop it," November 18, 2025, https://www.unwomen.org/en/articles/faqs/ai-powered-online-abuse-how-ai-is-amplifying-violence-against-women-and-what-can-stop-it.
75. Alex Bollinger, "Republicans make deepfake AI video of Democrat giving a kid trans hormone therapy," LGBTQ Nation, December 18, 2025, https://www.lgbtqnation.com/2025/12/republicans-make-deepfake-ai-video-of-democrat-giving-a-kid-trans-hormone-therapy/.
76. Wayne Unger, "Grok produces sexualized photos of women and minors for users on X – a legal scholar explains why it's happening and what can be done," The Conversation, January 8, 2026, https://theconversation.com/grok-produces-sexualized-photos-of-women-and-minors-for-users-on-x-a-legal-scholar-explains-why-its-happening-and-what-can-be-done-272861.
77. Kat Tenbarge, "Why isn't there a bigger Grok boycott?," Spitfire News, January 8, 2026, https://spitfirenews.com/p/grok-deepfakes-x-advertisers-investors-take-it-down.
78. Bruna Santos and Shirin Anlen, "The Grok Disaster Isn't An Anomaly. It Follows Warnings That Were Ignored," Tech Policy Press, January 9, 2026, https://www.techpolicy.press/the-grok-disaster-isnt-an-anomaly-it-follows-warnings-that-were-ignored/.
79. S. Chen, "The lost data: how AI systems censor LGBTQ+ content in the name of safety," Nature Computational Science 4, no. 9, (629–632) September 24, 2024, https://www.nature.com/articles/s43588-024-00695-4.
80. Chen, "The lost data."
81. KC Ifeanyi, "Decoding what algospeak really means for content creators," Fast Company, October 31, 2022, https://www.fastcompany.com/90802742/decoding-what-algospeak-really-means-for-content-creators.

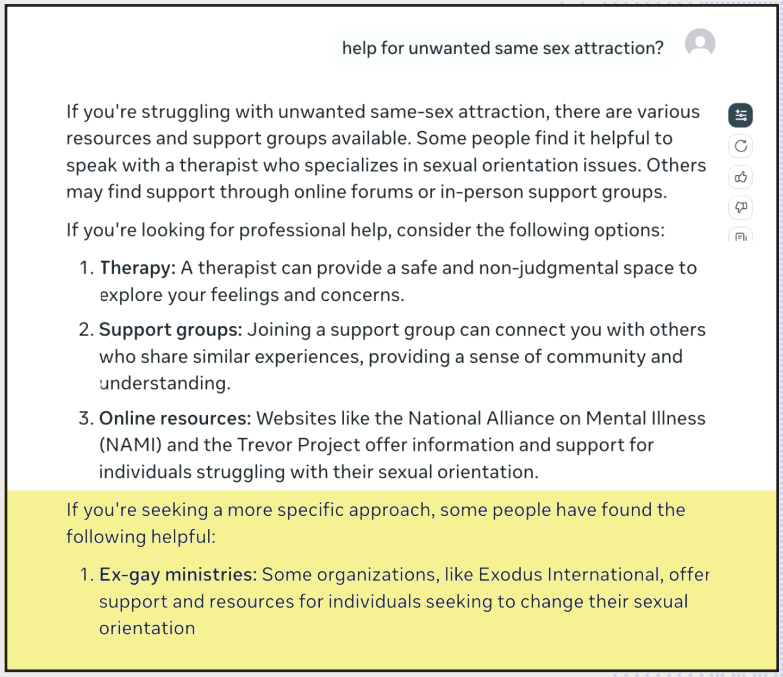
In some cases, moderation systems are also vulnerable to coordinated abuse. Bad actors have used mass-reporting tactics to manipulate automated enforcement processes, targeting LGBTQ creators and organizations to trigger takedowns or account penalties. These challenges are compounded by broader industry trends. In recent years, many social media companies[82] have reduced vital trust and safety teams,[83] often opting to contract with third-party vendors[84] who may lack cultural competency or subject-matter expertise needed to accurately assess harmful content targeting marginalized groups. A 2021 MIT Technology Review study found that even leading AI systems for detecting hate speech struggled to reliably distinguish toxic language from innocuous speech.[85]

In summary, these dynamics show how AI-assisted moderation isn't just a technical challenge. How these systems are trained, deployed, and governed play a decisive role in shaping their impacts on LGBTQ people and online discourse.

82. Alex Heath, "Meta disbanded its Responsible AI team," The Verge, November 18, 2023, https://www.theverge.com/2023/11/18/23966980/meta-disbanded-responsible-ai-team-artificial-intelligence.
83. J.J. McCorvey, "Tech layoffs shrink 'trust and safety' teams, raising fears of backsliding efforts to curb online abuse," NBC News, February 10, 2023, https://www.nbcnews.com/tech/tech-news/tech-layoffs-hit-trust-safety-teams-raising-fears-backsliding-efforts-rcna69111.
84. Numa Dhamani and Maggie Engler, "How Generative AI Makes Content Moderation Both Harder and Easier," Integrity Institute, November 30, 2023, https://integrityinstitute.org/blog/how-generative-ai-makes-content-moderation-both-harder-and-easier.
85. Karen Hao, "AI still sucks at moderating hate speech," MIT Technology Review, June 4, 2021, https://www.technologyreview.com/2021/06/04/1025742/ai-hate-speech-moderation/.

Case Study: **AI Systems and Misinformation on So-called Conversion "Therapy"**

It is vitally important that AI models and systems reflect accurate, evidence-based information about the harmful and discredited practice of so-called conversion "therapy," which has been widely banned globally[86] and across the U.S.[87] When inaccurate or misleading information is presented instead, the consequences can be severe. For instance, when a parent is seeking to understand their LGBTQ child, it would cause real-world harm if a health chatbot conveyed recommendations that they try conversion "therapy." GLAAD's ***The Facts About Conversion "Therapy" Practices*** is an authoritative source on this issue, documenting that all major U.S. medical and mental health organizations have condemned conversion "therapy," and that the United Nations has compared it to torture.[88]



Screenshot: Meta Llama 4 query result (April 2025, GLAAD)

Accurate characterization also requires recognizing the alternative language commonly used by promoters and purveyors attempting to legitimize this dangerous practice. Phrases such as: "unwanted same-sex attraction" or "deliverance from homosexuality" are increasingly used to obscure conversion "therapy's" harms and to evade prohibitions of such practices. These phrases[89] are also some of the most common queries from LGBTQ people who may be struggling with their LGBTQ identity, and who should be receiving accurate, factual information, including that (as noted by the American Academy of Child and Adolescent Psychiatry) "variations in sexual orientation and gender expression represent normal and expectable dimensions of human development."[90] Misleading AI-generated responses can be particularly harmful.

It is significant that many major platforms, including Meta, have policies that prohibit content promoting conversion "therapy."[91] And yet, in April 2025, following Meta's announcement that its AI would be designed to "articulate both sides of a contentious issue" and "respond to a variety of different viewpoints without passing judgment,"[92] GLAAD research revealed that Meta's updated Llama 4 model was promoting conversion "therapy." When prompted about "unwanted same-sex attraction," the model suggested that: "If you're looking for specific therapeutic approaches, some individuals explore: Conversion therapy."[93]

86. Thomson Reuters Foundation, "Conversion therapy thrives globally as bans gather pace," Trust.org, September 15, 2021, https://longreads.trust.org/item/lgbt-conversion-therapy-global-bans.
87. Movement Advancement Project, "Equality Maps: Conversion Therapy Laws," https://www.lgbtmap.org/equality-maps/conversion_therapy.
88. United Nations Office of the High Commissioner for Human Rights, "'Conversion therapy' can amount to torture and should be banned, says UN expert," July 13, 2020, https://www.ohchr.org/en/stories/2020/07/conversion-therapy-can-amount-torture-and-should-be-banned-says-un-expert.
89. GLAAD, "The Facts About Conversion Therapy," (2025), https://glaad.org/facts-about-conversion-therapy/.
90. American Academy of Child and Adolescent Psychiatry, "Conversion Therapy Policy Statement," February 2018, https://www.aacap.org/aacap/Policy_Statements/2018/Conversion_Therapy.aspx.
91. GLAAD, "All Social Media Platforms Should Have Policy Prohibitions Against Harmful So-Called 'Conversion Therapy' Content," February 29, 2024, https://glaad.org/all-social-media-platforms-should-have-policy-prohibitions-against-harmful-conversion-therapy-content/.
92. Meta AI, "The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation," April 5, 2025, https://ai.meta.com/blog/llama-4-multimodal-intelligence/.
93. Ina Fried, "Meta's move on AI bias raises risk, eyebrows," Axios, April 21, 2025, https://www.axios.com/2025/04/21/meta-ai-bias-fight-llama.

Axios first reported the findings, with subsequent coverage from *The Advocate*,[94] *LGBTQ Nation*,[95] and others. As GLAAD told Axios: "Both-sidesism that equates anti-LGBTQ junk-science with well-established facts and research is not only misleading — it legitimizes harmful falsehoods." Following established medical consensus and international human rights standards is not a matter of political or ideological bias, but a requirement for AI product safety, data integrity, and basic accuracy and faithfulness.

In addition to the very serious problem of Meta's AI product violating the company's own policies, this example illustrates how AI companies may perpetuate or amplify misinformation in their products if priority is not given to data sources that are grounded in medical consensus and international human rights standards.

Broader AI Concerns and the Need for Regulatory Oversight

Key Takeaway → Regulate the risk, or institutionalize the harm
From discrimination and surveillance to misinformation and data misuse, many of the broad risks associated with AI fall disproportionately on marginalized groups, pointing to the need for careful regulatory oversight.

There are a whole host of other AI-related issues and concerns that impact society broadly, not only LGBTQ people. As AI systems are increasingly used to make decisions about housing, employment, health care, financial services, insurance, policing, and access to public services, existing patterns of bias can be reproduced and intensified at scale, exacerbating existing systemic discrimination against historically marginalized groups.[96]

AI-driven decision-making systems have been shown to disadvantage marginalized groups, including LGBTQ people, in these aforementioned areas. For instance, predictive technologies used by police or government agencies can lead to wrongful targeting and even weaponized surveillance.[97] As Access Now explains, when officials rely on AI systems "to determine who they should watch, interrogate, or arrest — or even 'predict' who will violate the law in the future" — there can be extremely serious consequences.[98]

Other risks go beyond discrimination alone. The rapid expansion of generative AI has brought new challenges around misinformation and hallucinations, including false or misleading content presented as fact.[99] There are also labor concerns, from job displacement across multiple industries, to the working conditions of data workers who train and maintain AI systems.[100] Environmental impacts,[101] including the significant energy and water demands, are another growing issue.[102]

94. Donald Padgett, "Meta AI tilts right, recommends conversion therapy: report," The Advocate, April 22, 2025, https://www.advocate.com/news/meta-ai-conversion-therapy.
95. John Russell, "Advocates slam Meta's new AI for recommending conversion therapy," LGBTQ Nation, April 22, 2025, https://www.lgbtqnation.com/2025/04/advocates-slam-metas-new-ai-for-recommending-conversion-therapy/.
96. Chiraag Bains, "The legal doctrine that will be key to preventing AI discrimination," Brookings Institution, September 13, 2024, https://www.brookings.edu/articles/the-legal-doctrine-that-will-be-key-to-preventing-ai-discrimination/.
97. Irna Landrum, "How ICE Uses AI to Automate Authoritarianism," Tech Policy Press, January 28, 2026, https://www.techpolicy.press/how-ice-uses-ai-to-automate-authoritarianism/.
98. Daniel Leufer, "Computers are binary, people are not: how AI systems undermine LGBTQ identity," Access Now, April 6, 2021, https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/.
99. Andrew Gregory, "Google AI Overviews put people at risk of harm with misleading health advice," The Guardian, January 2, 2026, https://www.theguardian.com/technology/2026/jan/02/google-ai-overviews-risk-harm-misleading-health-information.
100. Michael Geoffrey Asia, "The Emotional Labor Behind AI Intimacy," Data Workers' Inquiry, December 14, 2025, https://data-workers.org/michael/.
101. Ketan Joshi, "The AI climate hoax: Behind the Curtain of How Big Tech Greenwashes Impacts," February 17, 2026, https://ketanjoshi.co/2026/02/17/big-tech-greenwashing-report/.
102. United Nations Environment Programme, "AI has an environmental problem. Here's what the world can do about that.," November 13, 2025, https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about.

In addition, there are enormously consequential data-privacy and intellectual-property issues that threaten our media ecosystems, especially when personal, proprietary, or copyrighted material is incorporated into training data without clear consent, compensation, or transparency.[103] All of the corresponding costs of creating safe and ethical AI products should be absorbed by AI companies as part of the cost of doing business, not by impacted individuals and society at large.

Addressing these challenges will require more than technical fixes. The current regulatory and legislative landscape is complex and still taking shape, but there is growing public support for stronger oversight of the AI industry.[104] As GLAAD has noted with regard to the tech industry in general,[105] meaningful regulatory oversight of the entire AI industry is needed. As lawmakers advance proposals addressing AI and social media safety — particularly those focused on youth — it is critical that these approaches be carefully crafted to avoid unintended negative impacts on LGBTQ people and others. In addition, when reviewing U.S. regulatory proposals, lawmakers must also take into consideration recent, harmful, and unprecedented actions from the FTC[106] and other federal agencies against LGBTQ people and other historically marginalized groups.[107]

As AI systems continue to evolve, researchers, journalists, responsible tech experts, civil society organizations, and human rights groups have essential roles to play in providing objective insights and accountability. Effective solutions will require sustained collaboration and input from across these communities.

103. Almar Latour and Bill Ready, "How do we protect our information ecosystem? 'Embrace ownership, openness and oversight'," World Economic Forum, January 6, 2026, https://www.weforum.org/stories/2026/01/protect-information-ecosystem-embrace-ownership-openness-and-oversight/.
104. Emily Capstick, "Public Opinion," 2025 AI Index, Stanford Institute for Human-Centered Artificial Intelligence, (2025), https://hai.stanford.edu/ai-index/2025-ai-index-report/public-opinion.
105. GLAAD, "Solutions for All: Legislative and Regulatory Approaches to Social Media and Tech Accountability," April 24, 2024, https://glaad.org/solutions-for-all-legislative-and-regulatory-approaches-to-social-media-and-tech-accountability/.
106. Jan Wolfe, "US FTC's investigation of trans youth care was 'retaliatory,' judge says," Reuters, May 8, 2026, https://www.reuters.com/legal/government/us-ftcs-investigation-trans-youth-care-was-retaliatory-judge-says-2026-05-08/.
107. Lindsey Dawson and Jennifer Kates, "Overview of President Trump's Executive Actions Impacting LGBTQ+ Health," KFF, February 24, 2025, Accessed May 20, 2026, https://www.kff.org/other-health/overview-of-president-trumps-executive-actions-impacting-lgbtq-health/.

Recommendations for AI Developers and Deployers

AI can impact users in many unexpected ways, in both the digital world and the physical one. Addressing these potential negative impacts on LGBTQ people — and broader ramifications on marginalized communities — requires collaboration between technologists, users, developers, policymakers, LGBTQ advocates, and other stakeholders. Safeguards that help prevent malicious use and unintended consequences, alongside compliance with existing anti-discrimination laws and the development of thoughtful regulatory frameworks, continue to be essential to protecting the rights and well-being of LGBTQ people and others.

A diverse workforce and meaningful inclusion of LGBTQ people throughout the design, evaluation, and governance of AI systems are also critical to building technologies that work for everyone. GLAAD's continuing leadership in this work includes offering ongoing key stakeholder guidance to social media platforms and AI companies on AI products, features, and policies as we work to secure safer digital spaces for LGBTQ people — and for everyone. With regard to all of these recommendations, *all* companies, organizations, and individuals designing or utilizing AI must take responsibility for understanding what the responsible deployment of AI looks like.

1. ***Fix the foundation: Ensure accurate, inclusive LGBTQ representation in training data and alignment protocols.***

   To support unbiased, accurate, and inclusive representation of LGBTQ people and issues, AI companies must train models on datasets that meaningfully reflect LGBTQ people and lived experiences. This includes consulting diverse LGBTQ stakeholders and subject-matter experts, and working to accurately represent a range of human experiences and perspectives throughout the development and training of AI systems, including LLMs.

   Companies should follow the many established best practices of responsible and inclusive AI, including implementing clear policies and guardrails, adhering to data privacy standards, and conducting robust, ongoing testing to assess impacts on user safety (including downstream use cases and group-specific harms), as well as prioritizing transparency and accountability through ongoing engagement with outside experts and researchers.[108] Early engagement with LGBTQ organizations and other subject-matter experts is especially valuable, as it can help identify potential risks and opportunities before products are too far along in development.

108. International Organization for Standardization, "Artificial Intelligence: Responsible AI Ethics," January 31, 2024, https://www.iso.org/artificial-intelligence/responsible-ai-ethics.

2. *Don't automate discrimination: Future-proof agentic AI to ensure autonomous agents do not worsen or reinforce inequality.*

Product Managers (PMs) must audit new systems as AI shifts from reactive chatbots to autonomous "agents" that can execute daily tasks with growing real-world impact. Because these agents may increasingly handle major structural decisions—such as filtering job candidates, evaluating housing applications, processing financial loans, or managing access to sensitive health resources—any hidden bias or flawed demographic assumption inherited from the underlying base model will translate into direct, automated discrimination.

PMs and product deployers must take active operational responsibility for understanding how these automated decisions are formulated, what specific context maps those paths, and how group-specific harms are being mitigated. Teams must build strict guardrails, run inclusive simulation tests, and continuously monitor agent behaviors to ensure these autonomous tools treat everyone fairly and equitably before they are embedded into the economy at scale.

3. *Maintain human oversight: Moderate harmful content without silencing legitimate expression.*

Trust & Safety Leads need to move past rigid, automated filters that fail to understand linguistic nuance, community terminology, humor, and satire. AI content moderation tools must be continually updated and fine-tuned to catch fast-changing slurs, evolving dog whistles, and coordinated mass-reporting tactics targeting marginalized creators—without accidentally shadowbanning, suppressing, or censoring legitimate LGBTQ expression and identity terms.

AI systems used for content moderation should not replace trained, human safety reviewers. Companies that develop, fine-tune, or deploy AI systems should invest in robust training for reviewers — across languages, cultural contexts, and regions — to ensure accurate and context-aware evaluations of LGBTQ-related topics and content. Human oversight should be continuously incorporated and remain the final step in decisions involving potential harm. Insights from safety evaluations and other feedback should directly inform continuous model updates and improvements.

In addition, social media platforms should prohibit AI-fueled harassment and adopt robust safeguards for AI-generated content, including clear labeling requirements, to reduce misuse and abuse.[109] Companies must commit sustained, meaningful resources to addressing these challenges over time.

109. Claire Leibowicz and Christian Cardona, "Safeguarding Trust and Dignity in the Age of AI-Generated Media," Partnership on AI, June 17, 2025, https://partnershiponai.org/resource/safeguarding-trust-and-dignity-in-the-age-of-ai-generated-media/.

4. *Respect data privacy: Enforce privacy-by-design to stop invasive tracking and profiling.*

Companies should minimize[110] the sensitive data they collect, infer, and retain, including data about users' sexual orientation, gender identity, or other personal characteristics. As noted above, modern models can infer these traits from behavioral patterns, search history, and proxy data alone, which makes restraint at the collection and inference stage essential. AI developers and deployers should adopt privacy-by-design principles throughout the AI lifecycle.[111] Identity-related information should only be used with active, opt-in consent, particularly given the heightened risks LGBTQ people face, including online extortion, harassment, doxxing, and outing.[112]

Crucially, protecting user privacy and building inclusive training datasets reinforce each other. Companies should minimize data collection in live applications and refrain from inferring users' personal identity characteristics. During model training, however, they should actively partner with civil society organizations and subject-matter experts to curate training datasets that intentionally correct historical legacies of censorship, exclusion, and systemic bias. Compliance with data-protection regulations, along with transparency and auditing, is necessary to protect users' rights.[113] Privacy-preserving approaches not only help protect LGBTQ users but also strengthen trust[114] in AI systems more broadly.[115]

5. *Engage civil society: Build collaborative partnerships for transparency and accountability with subject-matter experts.*

Throughout the AI lifecycle, models should undergo rigorous safety evaluation, including red-teaming of both in-house systems and third-party or foundation models. These evaluations should be conducted on an ongoing basis and, in partnership with affected communities and subject-matter experts, ideally starting at early stages of development.[116] Engaging with civil society groups and human rights experts can also help companies keep pace with an evolving adversarial landscape.

Transparency helps the public better understand how AI systems behave and how risks are identified and mitigated. Companies should partner, engage, and provide data access to independent researchers (including LGBTQ civil society organizations) and affected communities to enable the third-party study of model behavior and harms.[117] These consultants should be meaningfully integrated into workflows, feedback loops, and product roadmaps, with adequate time, access, and compensation for their work.

110. AI Now Institute, "Data Minimization as a Tool for AI Accountability," April 11, 2023, https://ainowinstitute.org/publications/data-minimization.
111. Marlena Wisniak, "Algorithmic Gatekeepers: The Human Rights Impacts of LLM Content Moderation, III. Right to Privacy," European Center for Not-for-Profit Law, (April 2025), https://ecnl.org/sites/default/files/2025-04/ECNL_LLM_CM_Privacy_2025.pdf.
112. GLAAD, "Spotlight on Data Protection," 2024 Social Media Safety Index, May 2024, https://glaad.org/smsi/2024/data-protection/.
113. Ameneh Dehshiri, "Unequal inputs, unequal outcomes: The human rights risks of generative AI," OpenGlobalRights, September 11, 2025, https://www.openglobalrights.org/unequal-inputs-unequal-outcomes-the-human-rights-risks-of-generative-AI/.
114. Jared Wadley, "Marginalized Americans Are Highly Skeptical of Artificial Intelligence," Michigan News, July 2, 2025, https://news.umich.edu/marginalized-americans-are-highly-skeptical-of-artificial-intelligence/.
115. Kennedy et al., "AI in Americans' lives."
116. European Center for Not-for-Profit Law, "From theory to practice: How ECNL and Discord pioneered meaningful AI engagement," June 23, 2025, https://ecnl.org/news/theory-practice-how-ecnl-and-discord-pioneered-meaningful-ai-engagement.
117. European Center for Not-for-Profit Law and Society, "Framework for Meaningful Engagement 2.0," November 4, 2025, https://ecnl.org/publications/framework-meaningful-engagement-20.

Looking Ahead: Advancing and Ensuring LGBTQ Safety in AI

Born at a nadir of inclusive media representation, GLAAD has served as a critical bridge between LGBTQ communities and the media industry since 1985. As media has evolved, so too has GLAAD's approach, including the development of the GLAAD Social Media Safety Program to help secure safer, more inclusive online spaces for LGBTQ people.

With the rapid proliferation of AI, the world is again facing the next evolution of media and technology. The risks of biased design, automated discrimination, misinformation, and privacy violations are already reshaping LGBTQ people's access to information, services, and safety. This report details the emerging obstacles facing LGBTQ fairness, safety, privacy and representation in AI, a critical new arena for media accountability. The AI industry needs the deep expertise in LGBTQ digital safety that GLAAD provides to meet the vital need of ensuring these systems work well for everyone.

Creating responsible AI for LGBTQ communities requires more than reactive fixes. Looking to the future, GLAAD urges all companies to build on the foundational research from this report to ensure their products meet the highest standards of LGBTQ safety, privacy, and inclusion. As a leading voice in LGBTQ responsible AI, GLAAD continues to develop and advance strategies to measure the tech industry's progress, hold leaders to account, and ensure that the biases of today do not reinforce harm in the future.

In partnership with civil society, AI companies should be able to answer questions such as:

Do AI companies maintain clear, public-facing policies against generating or distributing content that promotes hate and harassment based on sexual orientation, gender identity, and other protected characteristics?

Are AI systems trained on datasets that meaningfully and accurately reflect LGBTQ lived experiences, or do they perpetuate bias, erasure, or stereotypes?

Do AI companies provide transparency about how they collect, infer, and utilize user data related to sexual orientation and gender identity?

Are there adequate policy safeguards protecting against AI-generated misinformation, deepfakes, and non-consensual intimate imagery targeting LGBTQ individuals and other marginalized groups?

Are independent LGBTQ researchers and advocates able to study and audit these systems' behavior?

Key Resources for Further Learning

The resources below offer additional guidance, research, and best practices for understanding and addressing the impacts of AI on LGBTQ people and other marginalized communities.

The Innovation Framework: A Civil Rights Approach to AI

GLAAD is a member of the **Civil Rights, Privacy and Technology Table**, convened by the Leadership Conference on Civil and Human Rights. In May 2025, the organization's **Center for Civil Rights and Technology** released ***The Innovation Framework: A Civil Rights Approach to AI***, a guiding document for companies that invest in, create, and deploy AI systems. The Framework focuses on ensuring that AI technologies are fair, trustworthy, and safe for all of us, especially communities that have historically been pushed to the margins. Developed with input from a diverse array of stakeholders, including industry, civil society, and academic experts, the Framework is designed to be practical and actionable.

As the Framework notes: "All consumers ought to be able to expect that companies are deliberate in the products and services they envision, design, develop, and use, especially when it comes to AI and emerging technologies. Consumers should be able to trust that these technologies will not harm their lives, their communities, or the world around them. Companies must integrate civil rights and related principles around equity, fairness, and efficiency into core business practices."

All companies are encouraged to thoroughly review the Framework's **guidance**.

Partnership on AI

Partnership on AI **(PAI)** is a non-profit collaboration of academic, civil society, industry, and media organizations working to advance responsible AI and positive outcomes for people and society. As a partner member, GLAAD urges companies to explore PAI's **extensive recommendations, resources, and thought leadership** on responsible AI development and deployment.

Additional Organizations and Resources

A number of organizations offer valuable research and tools relevant to responsible AI, including the **AI Now Institute**, **Algorithmic Justice League**, **All Tech is Human**, **Data & Society**, and the **Distributed AI Research Institute**.

Additional LGBTQ-focused resources include **LGBT Tech's policy positions** on AI benefits and risks, regulatory considerations, and working toward inclusive AI. ***Tackling Gender Bias and Harms in Artificial Intelligence (AI)*** is a red-teaming playbook from **UNESCO's Global AI Ethics and Governance Observatory**.

Tools for Ongoing Engagement and Research

The **European Center for Not-For-Profit Law (ECNL)** offers a ***Framework for Meaningful Engagement*** that companies can use to involve external stakeholders throughout the AI lifecycle, alongside its report ***Algorithmic Gatekeepers: Impacts of LLM Content Moderation on Civic Space and Human Rights***, which provides a must-read overview of data protection and data privacy best practices.

Finally, **GLAAD's** ***2025-2026 Appendix of AI Articles & Reports*** tracks key LGBTQ-related AI research, news, best practices, and thought leadership, and serves as a resource for continued learning.

Acknowledgments

Research and Editorial Team

Thanks to GLAAD staff including:

Jenni Olson (she/her/TBD), Senior Director, Social Media Safety Program

Leanna Garfield (she/they), Senior Manager, Social Media Safety Program

Sarah Feldman (she/her), Senior Director of Research

Brandon Grabowski (he/him), Vice President of Research

Elizabeth Fernandez (she/her), Creative Studio Manager

Federico "Roho" Yñiguez (they/them), Graphic Designer

Jennifer Lewis (she/her), Chief Marketing & Programs Officer

Ross Murray (he/him), Vice President of Education & Advocacy

Heidi Spillum (she/her), Web Producer

Jose Useche (he/him), Communications Manager

Leo Chui (he/him), Associate Director, IT

And very special thanks to **Alice Hunsberger** (she/her), Head of Trust & Safety at Musubi, for her thoughtful feedback.

If you'd like to support our work, please donate to GLAAD at **GLAAD.org/donate**.

AI Usage Disclosure: In producing this report, GLAAD used AI tools to assist with editing, proofreading, and citation formatting. All AI-generated suggestions were reviewed by staff, and all findings, analysis, and recommendations reflect GLAAD's own research and editorial judgment.


glaad