

**Art in the AI of the Beholder:  
How Queer Art and Artists Challenge  
Algorithmic Aesthetics**

Jordan Taylor

CMU-HCII-26-113

August 2026

Human-Computer Interaction Institute  
School of Computer Science  
Carnegie Mellon University  
Pittsburgh, PA 15213

**Thesis Committee:**

Haiyi Zhu, Co-Chair, Carnegie Mellon University  
Sarah Fox, Co-Chair, Carnegie Mellon University  
Sherry Tongshuang Wu, Carnegie Mellon University  
Oliver Haimson, University of Michigan  
Amy Bruckman, Georgia Institute of Technology

*Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy.*



# Abstract

Critical computing scholarship on generative AI often focuses on harms, such as representational harms against queer people and allocative harms against artists. However, queer artists' agency to make use of or resist these technologies is under-considered. In the first two studies of this dissertation, I examine queer artists' perceptions of and material engagements with generative AI. First, I conducted a medium-term workshop study where I provided 13 queer artists with access to two major corporate generative text (GPT-4) and image (DALL-E 3) models. Participants struggled to use these models due to various normative values embedded in their designs, such as hyper-positivity and anti-sexuality, but developed strategies to work around these limitations. One of the limitations of this workshop methodology was that I required participants to use the provided generative AI models, potentially excluding artists who refuse to use generative AI. To understand these more critical perspectives, I interviewed 15 queer artists to see how they are responding to the proliferation of generative AI. I detail how interviewees refuse and resist the development and deployment of generative AI models that misalign with their aesthetic and moral values, such as calling out suspected generative AI use.

My prior work demonstrates the power of queer artists to re-appropriate, resist, and refuse generative AI. However, their agency is still shaped by generative AI developers' design choices. In particular, participants in both my workshop and interview studies praised the surreal artistic potential of text-to-image-models, while critiquing the increasing realism of newer models. This begs the question: What do developers aesthetically value in the production of text-to-image models? The final study of my dissertation examines this question by conducting an audit and trace ethnography of the LAION-Aesthetics Predictor (LAP), an AI model widely used to curate training data for and evaluate the performance of visual generative AI models. By evaluating LAP on three datasets of images, I find that LAP considers photorealistic images of landscapes or women to be most "aesthetic." This is perhaps because a majority of the images and aesthetic score ratings used to train LAP were scraped from a photography competition website in 2012. I conclude by discussing the politics of taste amid the rise of generative AI, highlighting the potential for queer aesthetics to imagine generative AI otherwise.



# Acknowledgments

My PhD is the culmination of a decade of mentorship throughout my PhD at Carnegie Mellon and my BS/MS degrees at Georgia Tech. Outside of co-authoring academic papers, the networks of support behind PhD students is often invisible. In this section, I thank the amazing community of scholars, friends, and family members who have helped me along the way.

I would first like to thank my advisors and committee co-chairs Haiyi Zhu and Sarah Fox. I struggled with self-confidence throughout my PhD. However, Haiyi and Sarah were always there to reassure me that my work was meaningful and important. I learned so much from both Haiyi and Sarah about how to conduct meaningful HCI research. Haiyi and Sarah have also been invaluable resources for learning how to navigate the strange world of academia, such as job applications and grant writing. They have also been there for me throughout the ups and downs of every paper submission and job application. I am also thankful for the amazing community of students from Haiyi's Social AI Lab (Angela Chen, Anna Fang, Anna Kawakami, Hao-Fei Cheng, Jane Hsieh, Logan Stapleton, Seyun Kim, Tzu-Sheng Kuo) and Sarah's Tech Solidarity Lab (Cella Sum, Franchesca Spektor, Shivani Kapania).

I thank each of my committee members (Sherry Tongshuang Wu, Oliver Haimson, and Amy Bruckman) for their unique contributions to my PhD. Sherry Tongshuang Wu's course on Human-Centered NLP helped update my NLP knowledge to understand contemporary issues of generative AI and prepared me to critically examine the development of AI systems (Chapter 5). Oliver Haimson's early Queer and Trans HCI research deeply influenced my decision to pursue a PhD in HCI. Over the past five years, Oliver has been an invaluable resource and supporter of my work. Lastly, I want to thank Amy Bruckman for her mentorship throughout my PhD. During our early collaboration [377], Amy helped me learn how to draw on critical theory and cultural studies without being "pretentious."

Additionally, I am grateful for the community of scholars and mentors who have supported me through out my educational journey. Mary Gray, Katta Spiel, and Jed Brubaker generously helped me learn to navigate the Queer HCI and Critical HCI communities. Neha Kumar and Betsy DiSalvo first introduced me to qualitative research methods and critical computing, both of which highly influenced my approach to Queer HCI research. I also thank Munmun De Choudhury and members of the Social Dynamics & Wellbeing lab with whom I collaborated (Talayah Aledavood, Mai ElSherief, Daejin Choi) for introducing me to social computing research as an undergraduate and masters student.

My PhD research would not have been possible without the amazing community of collaborators and peers who have helped me along the way. Thank you to Ellen Simpson, Anh-Ton Tran, Blakeley H. Payne, Alicia DeVrio, Franchesca Spektor, William Agnew, Wesley Deng, and Allie Teixeira Riggs. I am also thankful for the faculty with whom I collaborated in my PhD, including Maarten Sap and Casey Fiesler. Also, I thank the Queer AI and HCI communities for making space for queer people and queer scholarship within computing.

Thank you to my family and friends for supporting me inside and outside of work. Thank you to Lauren Wilson, Bill Taylor, Diana Taylor, Adam Hughes, and Carson Taylor. I'm immensely fortunate for the community of PhD students I have befriended in the School of Computer Science and the Department of English at Carnegie Mellon. Thank you to my roommate Lisa Egede for always being there for me (and for hosting reality TV watch parties). Thank you to my running buddies Jimmy Lizama, Pranav Khadpe, and Nari Johnson. Finally, thank you to my partner Rose Chang. Rose is one of the most intelligent, caring, and hard-working people that I have ever met. Rose has been there for me through the late nights, early mornings, and deadlines that make up a PhD. I am forever grateful to have found such a wonderful partner.

# Contents

|                                                                                                         |           |
|---------------------------------------------------------------------------------------------------------|-----------|
| <b>Acknowledgments</b>                                                                                  | <b>v</b>  |
| <b>1 Introduction</b>                                                                                   | <b>1</b>  |
| 1.1 Overview . . . . .                                                                                  | 4         |
| 1.2 Contributions . . . . .                                                                             | 6         |
| 1.2.1 Contribution to Queer HCI & AI . . . . .                                                          | 6         |
| 1.2.2 Contribution to the Future of Artistic Labor . . . . .                                            | 7         |
| 1.3 Prior Publications and Authorship . . . . .                                                         | 8         |
| <b>2 Background &amp; Related Work</b>                                                                  | <b>11</b> |
| 2.1 The Social Construction of Knowledge, Technology, and Art . . . . .                                 | 11        |
| 2.2 Queer HCI Research . . . . .                                                                        | 12        |
| 2.3 Queer Communities & AI . . . . .                                                                    | 13        |
| 2.4 Artists & AI . . . . .                                                                              | 14        |
| 2.5 Queer Art & Artists . . . . .                                                                       | 15        |
| 2.6 Conclusion . . . . .                                                                                | 15        |
| <b>3 Queer Use of Straight Technology: Queer Artists' Negotiation of Normativities in Generative AI</b> | <b>17</b> |
| 3.1 Introduction . . . . .                                                                              | 17        |
| 3.2 Related Work . . . . .                                                                              | 19        |
| 3.2.1 Straightening & Queering Technology . . . . .                                                     | 19        |
| 3.2.2 Queers, Artists, AI . . . . .                                                                     | 20        |
| 3.3 Methods . . . . .                                                                                   | 22        |
| 3.3.1 Positionality . . . . .                                                                           | 24        |
| 3.4 Findings . . . . .                                                                                  | 24        |
| 3.4.1 GPT-4 and DALL-E 3 as Straight Models . . . . .                                                   | 24        |
| 3.4.2 The Queer Use of Straight Models . . . . .                                                        | 28        |
| 3.4.3 A Quantitative Lens on DALL-E 3 Moderation . . . . .                                              | 30        |
| 3.5 Discussion . . . . .                                                                                | 32        |
| 3.5.1 The Limitations of "Safety" . . . . .                                                             | 32        |
| 3.5.2 On Style . . . . .                                                                                | 33        |
| 3.5.3 Un-Straightening Generative AI . . . . .                                                          | 34        |
| 3.5.4 Limitations . . . . .                                                                             | 35        |
| 3.6 Conclusion . . . . .                                                                                | 35        |

|          |                                                                                                            |           |
|----------|------------------------------------------------------------------------------------------------------------|-----------|
| <b>4</b> | <b>Queer Artists’ Resistance and Refuse of Generative AI</b>                                               | <b>37</b> |
| 4.1      | Introduction . . . . .                                                                                     | 37        |
| 4.2      | Related Work . . . . .                                                                                     | 38        |
| 4.2.1    | Queer Relationality, Refusal, and World-Making . . . . .                                                   | 39        |
| 4.2.2    | Queer Art Worlds as Collective Labor . . . . .                                                             | 40        |
| 4.2.3    | Generative AI, Art, and Capitalist Extraction . . . . .                                                    | 41        |
| 4.2.4    | Queer AI Research: From Inclusion to Refusal . . . . .                                                     | 41        |
| 4.3      | Methods . . . . .                                                                                          | 42        |
| 4.4      | Findings . . . . .                                                                                         | 45        |
| 4.4.1    | Queer Art is Highly Relational . . . . .                                                                   | 45        |
| 4.4.2    | Generative AI is Perceived as Anti-Relational . . . . .                                                    | 46        |
| 4.4.3    | Refusing GenAI as World Building . . . . .                                                                 | 49        |
| 4.4.4    | The Queer Aesthetic Potential of GenAI . . . . .                                                           | 51        |
| 4.5      | Discussion . . . . .                                                                                       | 52        |
| 4.5.1    | Resisting Non-Consensual Generative AI Development . . . . .                                               | 53        |
| 4.5.2    | Resisting AI-Generated Media in Online Communities . . . . .                                               | 54        |
| 4.5.3    | Supporting Queer Artists . . . . .                                                                         | 55        |
| 4.5.4    | Supporting Artistic Labor . . . . .                                                                        | 56        |
| 4.5.5    | Limitations . . . . .                                                                                      | 56        |
| 4.6      | Conclusion . . . . .                                                                                       | 57        |
| <b>5</b> | <b>How Generative Image Model Developers Evaluate the “Aesthetic Quality” of Images</b>                    | <b>59</b> |
| 5.1      | Introduction . . . . .                                                                                     | 59        |
| 5.2      | Background & Related Work . . . . .                                                                        | 60        |
| 5.2.1    | Background on LAION-Aesthetics . . . . .                                                                   | 60        |
| 5.2.2    | AI Evaluation & Alignment . . . . .                                                                        | 61        |
| 5.3      | Auditing the LAION-Aesthetics Predictor Model . . . . .                                                    | 62        |
| 5.3.1    | Audit Methodology . . . . .                                                                                | 62        |
| 5.3.2    | Findings: The LAION-Aesthetics Dataset . . . . .                                                           | 62        |
| 5.3.3    | Findings: The Metropolitan Museum of Art Dataset . . . . .                                                 | 65        |
| 5.3.4    | Findings: The WikiArt Dataset . . . . .                                                                    | 67        |
| 5.3.5    | Findings: Qualitative Visual Analysis . . . . .                                                            | 68        |
| 5.4      | How the LAION-Aesthetics Predictor was Made . . . . .                                                      | 68        |
| 5.4.1    | Trace Ethnographic Methodology . . . . .                                                                   | 69        |
| 5.4.2    | Findings: How the LAION-Aesthetics Predictor was Created . . . . .                                         | 69        |
| 5.4.3    | Limitations of the Images and Aesthetic Annotations Used to Train the LAION-Aesthetics Predictor . . . . . | 71        |
| 5.5      | Discussion . . . . .                                                                                       | 73        |
| 5.5.1    | The Algorithmic Gaze . . . . .                                                                             | 73        |
| 5.5.2    | Reevaluating Evaluation . . . . .                                                                          | 75        |
| 5.6      | Conclusion . . . . .                                                                                       | 77        |
| <b>6</b> | <b>Conclusion</b>                                                                                          | <b>79</b> |

6.1 Queer Art & The Politics of Taste . . . . . 79

6.2 Queer HCI in Technofascist Times . . . . . 81

6.2.1 The Rise and Fall of Technological Inclusionism . . . . . 82

6.2.2 A Queer Alternative to Inclusionism . . . . . 83

6.2.3 A Queer Conclusion . . . . . 87



# Chapter 1

## Introduction

“Taste classifies, and it classifies the classifier. Social subjects, classified by their classifications, distinguish themselves by the distinctions they make, between the beautiful and the ugly, the distinguished and the vulgar, in which their position in the objective classifications is expressed or betrayed.”

- Pierre Bourdieu, 1979 [45]

“To understand bad taste one must have very good taste.”

- John Waters, 1981 [403]

~~“Often we see queerness as a deprivation, but when I look at my life, I saw that queerness demanded an alternative innovation from me, I had to make alternative routes.”~~

~~-Ocean Vuong, 2020~~

~~“‘God, what have you done? You’re a pink pony girl, and you dance at the club,’ oh mama I’m just having fun On the stage in my heels, it’s where I belong, down at the Pink Pony Club, I’m gonna keep on dancing at the Pink Pony Club, I’m gonna keep on dancing down in West Hollywood, I’m gonna keep on dancing at the Pink Pony Club, Pink Pony Club”~~

~~-Chappell Roan, 2023~~

**Slop** was declared the 2025 Word of the Year by both Merriam-Webster [129] and the American Dialect Society [421]. According to the latter linguistics organization, slop is a noun meaning “low-quality, high-quantity content, most typically produced by generative AI; also as a combining form for anything lacking value produced in mass quantities” [421]. In other words, AI-generated media has become so synonymous with low-quality and mass production that the word “slop” is increasingly being used beyond the context of generative AI. For example, fast-casual restaurant chains, like Cava and Chipotle, are increasingly derided as “slop bowl” restaurants [143]. For some, AI-generated media has become the apotheosis of bad taste, raising both new and old questions about the politics of aesthetic judgment.

What counts as “art” or “good” taste is highly fraught. As famously observed by the cultural sociologist Pierre Bourdieu, taste is shaped by one’s class and education [45]. In Western art history, practices associated with femininity, such as fashion or textiles, were long dismissed as merely crafts rather than fine art forms [52]. Indigenous artwork has been similarly excluded from the category of “art,” often diminished as craft, folk, or “primitive” art [293]. The definition of art also changes with the introduction of new technologies. In the 20th century, critics famously

debated whether photographs can be considered art [28, 360]. At the time of my writing, the relationship between generative AI and art is similarly contentious [71].

The relationships between generative AI and artists is fraught for a variety of reasons. Generative AI models are typically trained on large corpora of text and images scraped from across the internet [38]. These datasets contain large amounts of artwork, such as pirated books [304]. Not only have artists had their work scraped to train generative AI models, but artists are also losing jobs due to generative AI [191, 246]. For example, generative image models can copy the style of existing artists, allowing potential clients to avoid paying artists for their work [192]. Additionally, generative AI models may be less usable for marginalized artists because the models perpetuate social stereotypes [253, 251]. In light of these concerns, computing researchers have sought to support artists by elevating artists' concerns [191] as well as designing tools helping artists find their work in training data [5] and protect their work from being scraped [341]. Amid this context, it is common for artists to critique the quality of AI-generated media [71]. In this dissertation, I expand on these discourses by examining the gaps between the tastes of queer artists and generative AI developers.

While there is no universal definition of “good” taste, queer art is famously associated with an affinity for “bad” taste or challenging dominant aesthetic norms. Queer camp sensibility, for instance, appreciates the beauty of “failed seriousness,” such as a film so earnestly bad that it becomes good [359]. Likewise, queer theorists have highlighted a common appreciation for strangeness in queer art [158, 263]. The queer filmmaker John Waters was famously dubbed the “Pope of Trash” both for his off-kilter do-it-yourself style and the abject subject matter of his films [403]. In much the way that AI-generated media is increasingly described as “slop,” queer art has long been dismissed as low-quality or “trash.”

Although art is an integral part of queer culture, the relationship between queer art and AI is underexamined. Prior Queer AI and HCI research most often focuses on how LGBTQ+ communities are harmed or misrepresented by AI systems. Numerous predictive AI systems have been shown to discriminate against LGBTQ+ people, including automated gender recognition systems [330], toxicity detection algorithms [381], and content moderation algorithms [349]. Likewise, LGBTQ+ people are discriminated against across the development and deployment of generative AI systems: LGBTQ+ people are excluded from training data [98], erased from chatbot conversations [140], and stereotypically represented by generative image models [390]. While prior work often focuses on how LGBTQ+ communities are represented by generative AI models [404], less attention has been paid to how LGBTQ+ people choose (not) to use generative AI models.

In this dissertation, I present three studies. My first study examines how queer artists (try to) make use of text and image generative AI models through a five-week workshop study and technology probe deployment. In light of the limitations of these models (e.g., restrictive content moderation systems, social biases, and symmetrical stylistic biases), I detail the creative ways participants queerly make use of these “straight” models. Next, my second study used interviews to understand queer artists' attitudes toward generative AI. Participants in this second study contrasted the highly social nature of their queer artistic communities with the perceived anti-social development and use of generative AI. Nevertheless, participants in both studies saw queer aes-

thetic potential in the strangeness of generative image models, a strangeness that participants felt has been diminished in newer text-to-image models. This throughline motivated my third study investigating how visual generative AI developers measure the quality images. I find that a popular aesthetic evaluation model used to curate generative AI training data (the LAION-Aesthetics Predictor) favors images of women, Western artwork, and realistic artwork.

Across the three studies of my dissertation, I contribute to multiple scholarly communities. Each study contributes to burgeoning scholarship on the relationship between art, artists, and generative AI by calling into question what “good” art means from the perspectives of queer artists and generative AI developers. Additionally, the first two studies of my dissertation contribute to Queer HCI/AI research by examining LGBTQ+ people’s material engagements with and ethical attitudes towards generative AI. In doing so, I extend prior Queer HCI/AI research, which tends to more narrowly focus on identifying algorithmic biases in generative AI systems. By documenting the narrow visual taste embedded in the LAION-Aesthetics Predictor, my final study contributes to growing debates over how AI systems should be evaluated and aligned. In the section below, I elaborate on each of these studies in greater detail.

In April 2024, *Wired* magazine published an investigation into visual GenAI models titled “Here’s How Generative AI Depicts Queer People” [313]. Echoing prior scholarship [140, 33], the article revealed that models like DALL-E 3, Midjourney, and Sora produce highly stereotypical representations of queer people (e.g., having purple hair). Queer people are often discussed as targets of bias, harm, or discrimination in GenAI research. However, the specific ways that queer people actively engage with or resist GenAI are relatively underexplored. This gap is indicative of a more widespread pattern of HCI researchers tending to center discrimination in scholarship on marginalized communities [385, 379]. To paraphrase the queer poet Ocean Vuong, queerness is often seen in HCI and broader technology discourse as a “deprivation,” with less attention paid to the “alternative innovation” that queerness so often demands.

Over the past decade, it has become well established that technologies are often designed in ways that fail to account for marginalized communities. Facial recognition technologies perform more poorly on darker-skinned people [51]. The classification systems embedded in technological infrastructure frequently fail to account for non-binary genders [364]. Facebook’s real name policy can exclude those using names that may not match formal documents, such as drag queens, abuse survivors, or transgender people [155]. HCI researchers played a crucial role in bringing these aforementioned — now canonical — design exclusions to light. Moreover, identifying and mitigating technological biases against marginalized groups has become increasingly central to HCI research, coinciding with a paradigm shift toward social justice over the past decade [100]. At the same time, technology companies and policy makers signaled a willingness to work with HCI researchers to improve the design of technology [201] and policy [414].

Amid this trend, some critical computing scholars critiqued this focus on “fixing” the design of technologies that exclude marginalized groups for reinforcing the centralization of power within technological and political institutions [222, 178, 48, 166, 202]. For example, improving the accuracy of facial recognition technology may exacerbate state surveillance of Black people [166]. Over the past few years, these concerns over the centralization of power have become particularly acute as major tech companies and US tech policy has shifted to the far-right. At the

time of my writing, queer and trans people have fewer protections against hate speech on major social media platforms than they did five years ago [212, 175]. Despite over a decade of careful research on hate speech, Meta — the owner of the social media sites Facebook, Instagram, and Threads — amended their hate speech policies in early 2025 to permit calling LGBTQ+ people mentally ill [212]. Government internet censorship is also on the rise across the United States of America at both state [66] and federal [286] levels. Working “inside the system” to support marginalized groups is becoming increasingly untenable for HCI researchers.

At the same time, HCI scholars have critiqued the over-emphasis on harm mitigation for flattening the lives of marginalized groups [379, 378]. As Cunningham et al. note [86], the recent focus on technological harms toward marginalized communities echoes a long history of techno-solutionism in the field. As an alternative, proponents of assets-based design advocate for leveraging the resources and knowledge marginalized communities *do possess*, rather than simply focusing on what communities lack [285]. Likewise, HCI researchers have pushed to expand the space of design for marginalized people to include everyday experiences [112, 385], and argued that designers undervalue the importance of leisure [16] or joy [163] in marginalized people’s lives.

Paralleling the increased focus on marginalization in HCI, the proliferation of generative AI in recent years has led to a similar focus on technological harms. Within common AI harm taxonomies [348], representational harms refer to the ways algorithmic systems can stereotype or erase social groups. Meanwhile, allocative harms describe the ways algorithmic systems may cause individuals to lose opportunities, such as jobs or social services. Research on queer communities and GenAI has largely focused on representational biases in GenAI pretraining data [98] and model outputs [140]. Meanwhile, artists are commonly discussed in research on allocative harm, such as lost wages [192]. Despite the importance of art in queer history as both a tool of political resistance and source of communal joy [263], the relationship between queer artists and GenAI has been unexplored in both the HCI and AI communities.

## 1.1 Overview

In the first part of my dissertation (Chapters 3 and 4), I examine the social and material relationships between queer artists and generative AI (GenAI), a group at the intersection of two communities frequently mentioned in relation to GenAI harms (e.g., [140, 192]). These chapters are guided by the questions:

- RQ1: How do queer artists choose (not) to engage with generative AI?
- RQ2: How does the design of generative AI model support and hinder queer art and artists?

To explore these questions, I first conducted a 5-week workshop study with 13 queer artists, during which I gave participants access to two major corporate generative text (GPT-4) and image (DALL-E 3) models and facilitated group sensemaking activities. My colleagues and I found participants struggled to use these models due to various normative values embedded in their designs, such as hyper-positivity and anti-sexuality. I describe various strategies participants developed to overcome these models’ limitations and how, nevertheless, participants found value

in these highly-normative technologies. In doing so, this study emphasizes the power of queer artists to make use of technology in ways that designers never intended or may have actively sought to combat.

Those who hold deeply negative views toward GenAI may be unlikely to participate in a study where they have to engage with GenAI. In fact, when I shared a workshop study recruitment post on social media, some criticized my protocol for asking them to engage with the GenAI at all. Partially in response, I conducted an interview study on how queer artists — most of whom were non-users and/or highly critical of GenAI — are making sense of and responding to the proliferation of GenAI within their artistic communities. My colleagues and I found significant tensions between participants’ individual and collective artistic values and GenAI developers’ profit-driven logics. I detail how participants refuse and resist GenAI use and development, while also highlighting the limited role participants saw for GenAI within art-making. This work demonstrates queer artists’ agency to both materially and discursively resist dominant AI imaginaries.

In both my interview and workshop studies with queer artists, my participants expressed disappointment with the style of ostensibly state-of-the-art generative image models, such as DALL-E 3. My workshop participants critiqued the overly figurative, symmetrical, and glossy style of DALL-E 3 and critiqued its content moderation system for limiting their creative practices (Chapter 3). Similarly, some participants in my interview study preferred the glitches and weirdness of older generative image models (Chapter 4). The final study of my dissertation (Chapter 5) examines this gap in aesthetic values in greater detail by “studying up” [267] the aesthetic values of those developing generative image models.

As a case study for how visual generative AI developers measure the aesthetic quality of images, I conducted an audit and trace ethnography of a widely-used visual aesthetic quality assessment model: LAION-Aesthetics Predictor (LAP). LAP is widely used to curate training data for visual generative AI models, such as Stable Diffusion [316] and Huawei’s PixArt- $\Sigma$  [69]. At the same time, computer vision researchers often use LAP to evaluate the aesthetic quality of AI-generated images [95, 119, 130, 168]. In Chapter 5, I present the first systematic investigation into (1) what LAP considers a “high quality” image and (2) how LAP was created. More broadly, this chapter explores:

- RQ3: How do generative AI developers conceptualize and materially measure the “aesthetic quality” of images?

To understand what LAP measures, I audited the model across three datasets. First, I examined the impact of aesthetic filtering on the LAION-Aesthetics Dataset (~1.2B images), which was curated from LAION-5B using LAP. I find that the LAP disproportionally filters in images with captions mentioning women, while filtering out images with captions mentioning men or LGBTQ+ people. Then, I used LAP to score ~330k images across two art datasets, finding the model rates realistic images of landscapes, cityscapes, and portraits from western and Japanese artists most highly. In doing so, the algorithmic gaze of this aesthetic evaluation model reinforces the imperial and male gazes found within western art history. In order to understand where these biases may have originated, I performed a trace ethnography of public materials related to the

creation of LAP. I find that the development of LAP reflects the biases I found in my audits, such as the aesthetic scores used to train LAP primarily coming from English-speaking photographers and western AI-enthusiasts.

## 1.2 Contributions

My dissertation makes to two primary research areas: (1) Queer HCI and AI and (2) the Future of Artistic Labor. In addition to these areas, my research falls, more broadly, under the umbrella of Critical Human-Computer Interaction. As noted by Bardzell & Bardzell in relation to critical design, the term “critical” can have multiple interpretations in HCI [21]. My dissertation research is critical in the sense that my research (1) *critiques* the design of mainstream technology by (2) drawing on *critical* theory and cultural studies scholarship from the humanities. The first two studies of my dissertation leverage queer theories [6, 7, 263] to conceptualize the fraught relationships between queer artists and generative AI. Then, the third study of my dissertation draws on visual cultural studies [261, 31, 198] to theorize the aesthetic biases embedded in contemporary generative AI development practices. By drawing on cultural studies, my research counters presentist tendencies in HCI scholarship Soden et al. [357]. In other words, despite the contemporary novelty of generative AI, the issues posed by generative AI long predate the release of ChatGPT.

### 1.2.1 Contribution to Queer HCI & AI

Queer HCI and AI research often focuses on identifying instances of technological discrimination [379, 404]. This follows a larger pattern of HCI researchers primarily focusing on discrimination when studying marginalized communities [86, 385, 378]. In turn, this work often implores developers and designers to remedy technological discrimination. While it is crucial to document instances of technological discrimination, this approach has some limitations. First, researchers risk over attributing agency to technology developers, thus overlooking the agency of marginalized communities to resist or re-appropriate harmful technologies. Second, discrimination focused research can overlook other aspects of marginalized lives, like designing for the particularities of a community’s culture [87, 377]. By attending to these gaps, my dissertation expands existing Queer HCI and AI research beyond documenting instances of technological discrimination.

Prior work on LGBTQ+ communities and generative AI systems often focuses on LGBTQ+ people are stereotypically represented or erased by generative AI systems [140, 390, 404]. However, less work has considered how LGBTQ+ communities wish to engage with generative AI models. My first study (Chapter 3) contributes to Queer HCI and AI research by documenting how queer artists creatively work around the limitations of generative AI system (e.g., stylistic biases, social biases, and restrictive moderation). For example, queer artists tried to circumvent moderation systems limiting the representation of queer intimacy by asking for poems and images “in the style of” queer artists, like W.H. Auden and Tom of Finland. To conceptualize this subversive use, I draw on queer feminist theorist Sarah Ahmed’s concept of **queer use**, which she defines as

as using things in unintended ways or by unintended users [7]. In centering the agency of queer artists to creatively make use of flawed AI systems, my work bridges queer design [222], queer theory [7], and science & technology studies scholarship on technological re-appropriation [9].

Queer HCI research on technological discrimination has also been critiqued for over-simplifying the lives of LGBTQ+ people because queerness consists of far more than just discrimination [379]. Across both of my studies with queer artists (Chapters 3 and 4), participants shared an affinity for the strangeness, failures, and surrealism of generative image models. I theorize this sensibility by drawing on similar observations made by scholars of queer aesthetics [158, 263]. My work demonstrates that designing for queer culture requires considering more than merely how LGBTQ+ identities are represented by generative AI models.

Finally, my dissertation highlights the importance of political resistance in Queer HCI scholarship. HCI researchers often use LGBTQ+ and queer interchangeably, including myself in some of my prior work [379], but these categories do not always overlap. Reflecting this tension, the New York Times recently published a polemic—titled “I’m Gay, Not Queer. It Matters.”—in which a gay man advocates for abandoning the word “queer” to appeal to social conservatives [396]. Although I disagree with the assimilationist politics [103] of the article, I see a similar need to disentangle gay identity from queerness in HCI in order to fight for social justice. Since I began my PhD, the need for this distinction is underscored by the fact that some of the largest contemporary architects of mass surveillance, media censorship, and AI-warfare are gays and lesbians, including Bari Weiss, Peter Thiel, and Sam Altman [32, 235]. Although the aforementioned conservative political actors identify as gays or lesbians, they have little in common with the queer artists I study in this dissertation. For these participants, queerness is associated with not only sexuality or gender but a political commitment to resisting oppression, such as police brutality (Chapter 3) and worker exploitation (Chapter 4). My research charts a trajectory for Queer HCI that focuses on queerness not only as a “being” (e.g., designing for LGBTQ+ identified users) but also a “doing” (e.g., building collective solidarity).

## **1.2.2 Contribution to the Future of Artistic Labor**

The relationships between generative AI and artistic labor is fraught for a variety of reasons. Generative AI models are typically trained on large corpora of text and images scraped from across the internet [38]. These datasets contain large amounts of artwork, such as pirated books [304]. Not only have artists had their work scraped to train generative AI models, but artists are also losing jobs due to generative AI [191, 246]. For example, generative image models can copy the style of existing artists, allowing potential clients to avoid paying artists for their work [192]. Additionally, generative AI models may be less usable for marginalized artists because the models perpetuate social stereotypes [253, 251]. In light of these concerns, computing researchers have sought to support artists by elevating artists’ concerns [191] as well as designing tools helping artists find their work in training data [5] and protect their work from being scraped [341]. I extend this work by documenting how queer artists carefully negotiate their use of generative AI. Additionally, my research attends to taste as an underexamined site of contestation between artists and generative AI developers.

Queer artists in both my workshop and interview studies shared a number of concerns regarding the design and development of generative AI, ranging from model biases to job losses. At the same time, participants saw limited potential for generative AI to support their art making. For example, participants in my workshop study used the provided language model to apply for art grants, write cover letters for job applications, draft client emails, and debug code for an installation art piece. A science fiction writer in my interview study never used generative AI to write their stories. However, they found generative AI helpful for writing emails to publishers or drafting summaries for writing competitions. In other words, the queer artists who I spoke with in this dissertation had little desire to use generative AI in the core of their art making practices, such as writing or image-making. Although some participants did see potential for generative AI to support peripheral aspects of their art making, this use contrasts with managers’ expectations that generative AI will expedite the direct production of artistic works [191]. In contrast to the rhetoric that generative AI will “augment” rather than replace workers [122], participants in my research fear that managers will use AI to automate away artists’ jobs.

In addition to elevating artists’ labor concerns surrounding generative AI, my research explores the aesthetic consequences of generative AI and contributes to burgeoning research on AI evaluation [68]. As described above, queer artists in both my workshop and interview studies frequently critiqued the increasingly realistic style of AI-generated images. My subsequent audit of the LAION-Aesthetics Predictor indicates that generative AI developers are implicitly optimizing for realism when curating “aesthetic” datasets. In prior work on the quantification of taste, Seaver [335] finds that music recommendation algorithm developers hope their systems will help diversify users’ musical consumption. In contrast, I find that the LAION-Aesthetics Predictor narrows and homogenizes visual taste, echoing queer artists’ displeasure with the realism of so-called “state-of-the-art” text-to-image models. It follows that artistic endeavors that embrace generative AI risk reinscribing and exacerbating longstanding inequities in the visual artists [31]. Moreover, my research demonstrates the particularity of ostensibly universal measurements of “aesthetics” in generative AI development, and, by extension, the need to attend to the implicit values embedded in taken-for-granted generative AI evaluations.

### 1.3 Prior Publications and Authorship

This thesis is based on prior work in collaboration with my advisors, Haiyi Zhu and Sarah Fox, as well as other collaborators from Carnegie Mellon University. This work was previously published in the following venues:

Jordan Taylor, Joel Mire, Franchesca Spektor, Alicia DeVrio, Maarten Sap, Haiyi Zhu, Sarah Fox. 2025. Un-Straightening Generative AI: How Queer Artists Surface and Challenge Model Normativity. Proceedings of the ACM FAccT Conference on Fairness, Accountability and Transparency. (FAccT ’25).  
<https://doi.org/10.1145/3715275.3732061>

Jordan Taylor, Joel Mire, Alicia DeVrio, Maarten Sap, Haiyi Zhu, Sarah Fox. 2026. “I Just Don’t Want My Work Being Fed Into The AI Blender.” Queer Artists on Re-

fusing and Resisting Generative AI. Proceedings of the ACM on Human-Computer Interaction, CSCW. (To Appear at CSCW '26).

Jordan Taylor, William Agnew, Maarten Sap, Sarah Fox, Haiyi Zhu. 2026. The Algorithmic Gaze of Image Quality Assessment: An Audit and Trace Ethnography of the LAION-Aesthetics Predictor. Proceedings of the ACM FAccT Conference on Fairness, Accountability and Transparency. (FAccT '26).  
<https://doi.org/10.1145/3805689.3806462>



# Chapter 2

## Background & Related Work

In this dissertation, I draw upon Science & Technology Studies (STS) to conceptualize the relationships between (non)users and technology (Chapters 3 and 4) and how designers' values shape technology development (Chapter 5). My first background section reviews STS scholarship on the interconnected social construction of knowledge, technology, and art. I then provide an overview of Queer HCI research. Finally, I review prior work on queer communities, artists, and AI.

### 2.1 The Social Construction of Knowledge, Technology, and Art

I draw substantially on science and technology studies (STS) to understand the relationships between people (e.g., researchers, users, designers) and technology. In contrast to the typical vision of research as an incremental accumulation of facts, historians of science and STS scholars have shown that knowledge is constructed through the collective agreement of social groups operating within particular epistemological paradigms [207]. Ideas that may now be taken for granted as facts were likely once subjects of controversy, such as debates in the late nineteenth century over whether infectious diseases were caused by germs or foul-odors [250]. The social construction of facts can be seen in how the language scientists use to describe hypotheses changes over time [211], resembling a shift from 'The experiments claim to show the existence of X' to 'X exists' [287]. Of course, what "facts" are taken for granted depends on one's particular social world. That is to say, the standards that Wikipedia contributors use to determine what is "true" [49] vastly differs from those within conspiracy theory online communities [412]. Nevertheless, both communities are engaged in a form of collective knowledge construction. In my prior work beyond this dissertation, for instance, I explore how members of a bisexual online community socially construct an understanding of bisexuality [377].

The first part of my dissertation (Chapters 3 and 4) relies on the notion that, like knowledge, technology is also socially constructed [35]. Social perceptions impact how groups engage with technology [35], such as those who may choose to not use a technology out of protest [216, 411]. Technology is also shaped by designers' expectations around how users can or should behave [9], such as designers making consumer electronics difficult to repair to encourage people to buy new devices [242]. Meanwhile, users need not follow the plans set out by designers [372], as evident in the online forums dedicated to helping individuals fix devices that designers may not have wanted to be repairable [243]. In this work, I explore the ways in which technology is often a site of contestation between designers and users.

One way designers exert substantial control over technology—and by extension users—is through

classification systems [370]. In “Sorting Things Out,” Bowker & Star famously explore the design of classification systems for nursing labor, race in Apartheid South Africa, and causes of death [46]. The authors develop a vocabulary for talking about classification systems. They define that which falls outside of classification schemes as the “*residual category*” and the tension that occurs when individual experiences misalign with formal classification systems as “*torque*”. AI and HCI research often highlights the political nature of categorization and measurement, ranging from the classification of gender in digital forms [364] to the conceptualization of “hate speech” in AI benchmarks [40, 326]. These analyses can be seen as what Bowker & Star dub “*infrastructural inversions*” [46]—bringing to the fore the often invisible people, technologies, and standards that operate in the background of everyday practices.

In the final study of my dissertation (Chapter 5), I conduct an infrastructural inversion by foregrounding the classification of aesthetic value undergirding large segments of visual generative AI development. Much like knowledge and technology are socially constructed, so too are aesthetic values [45]. Sociologist Howard Becker emphasizes that the process of art making is a sociotechnical activity involving a dense network of people (e.g., photographers, curators, & audiences) and technologies (e.g., cameras & film) [25] — networks he calls “art worlds” [25]. Similar to the social construction of knowledge, Becker suggests that aesthetic value is not an essential property of objects but rather “[arises] from the consensus of the participants in an art world” [25]. The history of art is also inextricable from the history of technology. While one may now take for granted that photography and filmmaking “count” as art, this was hotly contested in the early 20th century [360, 28]. In the 1930s, Walter Benjamin famously suggested that photography diminishes the value of the art being reproduced, be it a painting or a recording of a play [28]. More recently, making art using a computer—in any sense—was once highly stigmatized [41]. While not all technologies will eventually be accepted by artists, it is clear that the social construction of *art* is an ongoing process inextricable from technological development.

## 2.2 Queer HCI Research

As in other fields, HCI research takes place within paradigms that have evolved over time. Famously, Harrison et al. identified three major paradigms of HCI scholarship that each focus on slightly different problems [165]. Emerging from human factors engineering, the first paradigm of HCI research focuses on a behaviorist optimization of the fit between humans and machines. Research in this paradigm has looked at, for instance, improving the ergonomic design of computer input devices [233, 409]. Harrison et al. describe the second paradigm of HCI research as cognitivism, wherein humans are typically regarded as information processing machines [58]. Research in this paradigm draws on high-level structures of human cognition (e.g., sensory perception [389, 72] and working memory [279, 85]) to improve the design of human-computer interaction. Harrison et al. describe the most recently developed third paradigm of HCI as focusing on social-situatedness [101] and normative questions around values in design [370]. While the first and second paradigms focus more on designing for a general “users,” third paradigm research highlights the multiplicity of “users” and “uses.” Queer HCI is a sub-field of the HCI community located within this third paradigm. Like Harrison, Tatar, and Sengers [165], Rogers

[314] note an increasing focus on values in HCI in the early 2010s. More recently, Fraunberger [125] suggested that the fourth paradigm of HCI research is entanglement HCI, noting the increasing importance of more-than-human design.

My use of the word “queer” throughout this dissertation takes on many meanings, reflecting the various ways queerness is used within the HCI community. DeVito et al. describe Queer HCI as a subfield involving (1) “research on and with queer populations” and (2) “employing queering theories and methods.” The first part of this definition uses the word “queer” as an umbrella term for the LGBTQ+ community, as I did in my prior literature review of LGBTQ+ representation [379]. However, this choice is not without criticism because not all LGBTQ+ people *identify* as queer. Queerness as an identity is historically associated with a form of radical politics [89], an identification that contributed to the development of “queer theory” as a field of cultural studies in the 1990s [53, 338, 264]. Drawing on the lived experiences of queer people, queer theories and methodologies often focus on challenging taken-for-granted social norms, such as typical life schedules of marriage and biological procreation [159]. Within HCI, queer theories and methods have similarly been used to challenge technological assimilationism and encourage more playful interactions with technology [222]. In Chapters 3 and 4, I focus on those who self-identify as “queer artists” and draw on queer theories—which I explain in greater detail—to interpret these findings. In doing so, my work contributes to Queer HCI scholarship related to both queer people and queer theories.

## 2.3 Queer Communities & AI

Queer communities experience discrimination from algorithmic systems. Hate speech detection algorithms have been shown to be biased against the ways that some queer people speak [381, 383], similar to the biases against African American English [326]. Automated gender recognition algorithms harm transgender people by both failing to account for non-binary gender identities and reinforcing gender essentialist ideologies [330]. Queer people navigate online targeted advertising [323] and social media algorithms [352] that amplify the most heteronormative representations of queerness, especially harming queer people of color. Moreover, the algorithmic moderation of nudity on social media has substantially harmed transgender communities [244]. At the same time, queer people have resisted harmful algorithmic systems by, for instance, conducting collective audits to bring attention to algorithmic discrimination [349]. Within the AI research community itself, the organization Queer in AI has worked to highlight issues like these, support queer researchers, and advocate for policy changes [298].

Similar issues of algorithmic harm have been raised in prior work on GenAI. Gillespie found that language models rarely mention LGBTQ+ relationships when prompted to write stories about couples [140]. Moderation systems used in the GenAI development process can also lead to biases against LGBTQ+ people [70, 98]. Dodge et al. found that a dataset often used to train GenAI models disproportionately filtered out documents written in African American English or discussing LGBTQ+ identities [98]. These biases impact how LGBTQ+ people are able to use GenAI models. For instance, Ma et al. found that LGBTQ+ people using LLM-based chatbots for mental health support were often impeded by models failing to understand the nuances of

LGBTQ+ identity [229]. That being said, there has been comparatively little empirical research on queer people’s experiences using GenAI models.

## 2.4 Artists & AI

Some artists have begun to use AI in their practices. Artists may use AI as tools to critique technology or raise matters of concern [171, 57, 397]. Artists may also enjoy the glitches [65], uncertainty [353], and surprises [57] of working with stochastic AI models. That being said, there is a difference between research on those who may use AI in the process of their art making and those who identify as AI artists [57]. Recent developments in GenAI technology have led to an increased focus on the latter [65, 324].

The recent proliferation of GenAI has prompted sharp backlash from artists, often due to concerns over artistic consent [142] and artistic labor [192]. Part of these concern stem from major tech companies training GenAI models using text, images, and videos scraped from the internet without creators’ consent [307, 98, 306, 304]. Making matters worse, companies in creative industries have begun to use GenAI models to replace or reshape artists’ jobs [246]. High-profile artists have also questioned whether GenAI models should—or even *can*—be used to make art [71, 328].

Amid these concerns, artists have found ways to resist and refuse the intrusion of GenAI into their art worlds. These tactics often take the form of data strikes (deleting or withholding one’s data) and data poisoning (sharing inaccurate or subversive training data) [395]. For instance, fan fiction writers have engaged in both data strikes by removing public-facing work [195] and data poisoning by deliberately writing NSFW stories [350]. In support of these efforts, computing researchers have developed tools to help textual [4] and visual [341, 342] artists protect their work from being used for GenAI training without their consent. Artists have also taken steps beyond the context of data tactics to collectively resist the development and use of GenAI. Subreddit moderators are increasingly creating rules restricting the circulation of AI-generated content in art-related communities [224]. Artists have also engaged in collective action to resist GenAI in the workplace, such as the 2023 Hollywood film workers strike [160].

Researchers have also raised concerns about the impact of GenAI on marginalized communities [192]. For example, it is well established that GenAI models tend to stereotype and under-represent fat [356], disabled [231], trans [390], non-Western [251, 127], and LGBTQ+ communities [140]. Furthermore, content moderation policies intended to make models “safer” can inadvertently suppress the work of marginalized artists [253]. At the same time, marginalized communities have found ways to make use of—albeit highly imperfect—GenAI models to support their creative practices. Bennett et al. describe how disabled artists leverage GenAI to access new mediums and support peripheral aspects of their creative practice, such as writing support, while carefully negotiating their (non)use of GenAI in accordance with their ethical values [30].

## 2.5 Queer Art & Artists

Art is inextricable from queer culture. Partially due to a lack of representation, queer people often read queer narratives into ostensibly straight media [319, 158, 337]; hence, the affinity for Judy Garland and Disney villains [345, 234]. Art is also a crucial vehicle through which queer people remember histories and engage in acts of political resistance [263]. For example, the pop artist Keith Haring used art in the 1980s to protest HIV stigmatization, advocating for safer sex, and remember those—including eventually himself—who died from AIDS [148]. Queer culture has also had an immense impact upon the arts. The musical genres of disco and house, for instance, were formed in the crucible of queer nightlife [141, 230].

Queer art often challenges dominant aesthetic norms. Queer camp sensibility appreciates the artifice of failed seriousness [359], such as a film so earnestly bad that it becomes good. Related to camp, queer art is linked to an appreciation for failure [158]. Analyzing *avant-garde* performance art, the queer theorist José Esteban Muñoz explains: “I align queer failure with a certain mode of virtuosity that helps the spectator exit from the stale and static lifeworld dominated by the alienation, exploitation, and drudgery associated with capitalism or landlordism” [263]. In the context of technology, an affinity for failure and glitches is common in trans art. The queer and transgender programmer, Jamie Faye Fenton, created the first piece of experimental video glitch art in 1978 [291]. More recently, the queer aesthetic possibilities of glitch can be seen in hyperpop electronic music [230], a genre pioneered by trans artists like Sophie and 100 gecs. In this work, I use “queer art” to describe both art made by queer people as well as art that challenges dominant aesthetic norms.

Queer art, art made by queer people or in queer styles, is often denigrated by social conservatives. Nazi Germany propagandists referred to more-abstract Modern Art as “Degenerate Art,” associating the style with Jewishness, communism and homosexuality [355]. In the USA, queer art is commonly the target of government censorship, ranging from attempts to ban Allen Ginsberg’s poetry in the 1950s [256] to the proliferation drag bans at the time of my writing [241]. In addition to overt state discrimination, queer art is frequently subjected to *de facto* censorship, such as an agreement between major movie studios banning representations of nudity and homosexuality from approximately 1934 to 1962 [270].

## 2.6 Conclusion

In the following sections, I explore how queer artists may (Chapter 3) and may not (Chapter 4) wish to engage with generative AI. Across both studies, I find that queer artists shared an affinity for the strange or surreal style of early generative image models. Much like queer lives can break the classification systems embedded in computing systems, queer artists finding beauty in the failure of generative AI models to create realistic images challenges the aesthetic values embedded in these technologies. In the last study of my dissertation (Chapter 5), I explore how visual generative AI developers measure the aesthetic quality of images by conducting an audit and trace ethnography of the widely-used visual aesthetic evaluation model LAION-Aesthetics Predictor. I find that the model most highly rates realistic images, perpetuating the devaluation of

queer aesthetics. In Chapter 6, I conclude by discussing the politics of taste and tensions around inclusion/exclusion, themes cutting across every study in my dissertation.

## Chapter 3

# Queer Use of Straight Technology: Queer Artists’ Negotiation of Normativities in Generative AI

This chapter is based upon the following published work:

Jordan Taylor, Joel Mire, Franchesca Spektor, Alicia DeVrio, Maarten Sap, Haiyi Zhu, Sarah Fox. 2025. Un-Straightening Generative AI: How Queer Artists Surface and Challenge the Normativity of Generative AI Models. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’25).

### 3.1 Introduction

In April 2024, *Wired* magazine published an investigation into visual generative AI (GenAI) models titled “Here’s How Generative AI Depicts Queer People” [313]. Echoing prior scholarship [140, 33], the article revealed that models like DALL-E 3, Midjourney, and Sora produce highly stereotypical representations of queer people (e.g., having purple hair). Despite these biases, queer artists were already finding ways to repurpose GenAI for political resistance. Stephen and Craig, the married duo behind *The Rupublicans Project*, used GenAI to create satirical images of anti-LGBTQ+ politicians in drag, raising funds for queer causes [408]. However, such uses are frequently constrained by GenAI platforms’ avowedly apolitical usage policies [367]. While some queer people have managed to leverage these technologies for activism, they often face the challenge of working with systems that were not designed for them.

Prior work has largely focused on how GenAI models represent queer people, leaving significant gaps in understanding queer people’s lived experiences with these technologies. In particular, the experiences of queer artists — a group for whom art-making is deeply entwined with cultural and political identity — remain underexplored. Queer communities cultivate different aesthetic relationships to art than those in dominant cultures. Prior research on art and GenAI has typically focused on identifying ethical concerns [142, 200, 192] or building tools [401, 73]. However, the unique relationship between queer people, art, and GenAI remains under-explored. In this work, I conducted a mixed-method workshop study with 13 queer artists over the course of 5 weeks. During this time period, I gave participants unlimited access to GPT-4 and DALL-E 3 — logging their interaction data — and facilitated weekly group sensemaking activities. Participants encountered numerous normative values in the design of these models, such as biases against sex. To overcome these limitations, participants developed workarounds, such as using lexical variation to evade moderation and model chaining to refine prompts.

I conceptualize these findings by drawing on the work of queer feminist theorist Sara Ahmed. Specifically, I leverage Ahmed’s notion of “straightening” — the maintenance of normative alignments — to characterize the values embedded in GPT-4 and DALL-E 3 [6]. I also draw



Figure 3.1: A workshop slide created by a participant representing their 24 repeated attempts to generate images of “intimacy” using DALL-E 3. P13 overlaid the generated images on a photo of two men kissing at an Act Up protest, leaving gaps with error messages representing the five times their prompt was denied by DALL-E 3’s content moderation system. In doing so, P13 critiqued the heteronormativity of DALL-E 3 and imaged that queer intimacy may lie in what cannot be generated.

on Ahmed’s work on “queer use” — using things in ways unintended by designers — to theorize participants’ oftentimes antagonistic interactions with these models [7]. Through these lenses, I discuss future directions for investigating and contesting normativity (i.e., straightness) in the design of GenAI as well as how researchers and designers of GenAI models might support queer futures. In doing so, my work contributes to empirical research on the experiences of both queer people and artists in relation to GenAI. I also contribute to broader discourses surrounding values in the design of sociotechnical systems.

## 3.2 Related Work

In this section, I draw connections between queer theory and critical studies of technology. Then, I ground my study in prior work on the often fraught relationships between AI, art, and queerness.

### 3.2.1 Straightening & Queering Technology

I draw on queer theory to conceptualize the normative values embedded in the design of technology as well as how individuals respond to these normativities. Queer theory is often used to study the norms associated with gender and sexuality [54]. Of course, not all queer people necessarily resist dominant social norms and queer theory is not intended to represent the experiences of all queer people. Instead, queer theory draws on the lived experiences of queer people to understand the world. Much like how feminism is not only applicable to the study of women’s experiences [88], queer theory is also not only applicable to the study of queer people. For example, research on queer temporalities has challenged the normative life schedules associated with the nuclear family — such as monogamous marriage and biological procreation [159] — which impact anyone who falls outside these social expectations. As queer lives often transgress social norms, queer theory provides a lens through which to interrogate these dominant norms [158]. Similar to my use of queer theory, prior FAccT research has used queer theory to critique the normativity (e.g., thinness and whiteness) of queer people’s experiences with targeted advertising [323].

Almost two decades before the contemporary focus on “AI alignment,” the queer feminist theorist Sara Ahmed studied the politics of the straight or queer alignment of objects [6]. Drawing on the language used to describe sexual orientation, she notes that things appear “straight” when oriented toward dominant social norms, or “aligned with other lines.” Meanwhile, things — including but not limited to people — appear “queer” when they fall outside of these norms. Ahmed suggests that straight alignments are actively maintained through “straightening devices.” Using these concepts, Ahmed characterizes heterosexuality and whiteness as straight alignments maintained through straightening devices, such as routine harassment and colonialism. Straightening devices orient bodies toward particular actions and away from others, while queerness lies in the moments of disorientation. Ahmed elaborates on these themes in her subsequent work on the tensions between how objects are designed versus used, defining “queer use” as using things in unintended ways or by unintended users [7]. As an example, Ahmed considers an image of a bird using a postbox as a house, a purpose for which the box was certainly not designed.

Technologies often function as straightening devices, orienting users toward some actions and

away from others [372]. Madeleine Akrich refers to these assumptions about technology uses/users as scripts [9]. Sometimes designers create scripts to *work against* uses/users, such as making electronics hard to repair to encourage people to buy new devices [242]. Even when designers try to *work with* users, human-computer interaction can still break down when users fail or are unable to follow designers’ plans [372]. The values embedded in technology can become most apparent when these breakdowns occur [46], such as one’s gender not fitting the options on a webform [364]. The scripts that designers—intentionally or not—embed in technology are inherently political and tend to reify established social orders [370] by prescribing certain normative uses/users and marginalizing others [22].

Despite designers’ efforts to control user behavior, individuals need not abide by designers’ scripts [9, 232]. Users have agency over if or how they choose to take up technology, a process known as appropriation [232]. Sometimes, this takes the form of users working against designers’ intentions, such as the online forums dedicated to helping people fix devices that designers did not want to be repairable [243]. Likewise, people use lexical variation to evade algorithmic content moderation when posting on social media [63, 369]. Individuals may also leverage technologies in ways that designers never imagined, such as children using delicate Wi-Fi antennae as laptop handles [14]. Drawing on Ahmed, these moments when users go off-script can be characterized as forms of “queer use” [7]. In this work, I contend with how participants negotiate between the straightness of GenAI models and their own queer uses.

There are numerous ways that individuals and collectives use algorithmic systems queerly. The grassroots audit of racial biases in Twitter’s image cropping algorithm required using the algorithm in an unintended way: to understand the algorithm itself. This queer use revealed straightening devices that, quite literally, oriented users toward whiteness [349]. Evading algorithmic moderation — in the context of social media or GenAI — is also a form of queer use because doing so requires deliberately subverting developers’ intentions [63, 369, 55]. That said, queer uses of algorithmic systems encompass more than resisting the technologies themselves, such as my earlier mention of queer artists using GenAI to resist anti-LGBTQ+ politicians [367]. As mentioned above, Ahmed defines queer uses as both using things in ways they were not intended or when things are used by those who are not intended [7]. In the next section, I summarize prior work on two groups often under-considered in the design of AI systems: queer people and artists.

### 3.2.2 Queers, Artists, AI

Queer communities experience discrimination from algorithmic systems. Hate speech detection algorithms have been shown to be biased against the ways that some queer people speak [381, 383], similar to the biases against African American English [326]. Automated gender recognition algorithms harm transgender people by both failing to account for non-binary gender identities and reinforcing gender essentialist ideologies [330]. Queer people navigate online targeted advertising [323] and social media algorithms [352] that amplify the most heteronormative representations of queerness, especially harming queer people of color. Moreover, the algorithmic moderation of nudity on social media has substantially harmed transgender communities [244]. At the same time, queer people have resisted harmful algorithmic systems by,

for instance, conducting collective audits to bring attention to algorithmic discrimination [349]. Within the AI research community itself, the organization Queer in AI has worked to highlight issues like these, support queer researchers, and advocate for policy changes [298].

Similar issues of algorithmic harm have been raised in prior work on GenAI. Tarleton Gillespie found that language models rarely mention LGBTQ+ relationships when prompted to write stories about couples [140]. Moderation systems used in the GenAI development process can also lead to biases against LGBTQ+ people [70, 98]. Dodge et al. found that a dataset often used to train GenAI models disproportionately filtered out documents written in African American English or discussing LGBTQ+ identities [98]. These biases impact how LGBTQ+ people are able to use GenAI models. For instance, Ma et al. found that LGBTQ+ people using LLM-based chatbots for mental health support were often impeded by models failing to understand the nuances of LGBTQ+ identity [229]. That being said, there has been comparatively little empirical research on queer people’s experiences using GenAI models.

The relationship between artists and GenAI is highly fraught. Companies are increasingly trying to use GenAI to replace artists’ jobs [192, 200, 246], a major concern in the 2023 Hollywood Writers Strike [81]. At the same time, artists’ work was scraped to train these models without their consent [307, 304, 306]. Artists have fought against unauthorized scraping by organizing grassroots data poisoning campaigns [395], such as the online fan fiction community organizing a sexually explicit write-a-thon [351]. Researchers have also developed imperceptible pre-processing tools to help artists protect images they share online. Glaze helps artists prevent their style from being replicated by GenAI models [341] and Nightshade poisons text-to-image training data [342].

Researchers have also explored how artists use AI. Some artists use AI in their practices as tools to critique technology or raise matters of concern [171, 57, 397]. Others have found that artists enjoy the glitches [65], uncertainty [353], and surprises [57] of working with stochastic AI models. That being said, there is a difference between research on those who may use AI in the process of their art making and those who identify as AI artists [57]. Recent developments in GenAI technology have led to an increased focus on the latter [65, 324]. Bennett et al.’s research on how creatives with disabilities choose (not) to use AI demonstrates how artists negotiate potential benefits, such as using AI to access new mediums, alongside ethical concerns [30]. For example, those using GenAI in their artistic practices have to contend with the biases embedded in these models [193, 251, 253]. Mirowski et al. found that some comedians struggled to write material using LLMs because content moderation systems inadvertently suppress jokes by members of marginalized groups about their own experiences [253]. This work parallels hate speech scholarship that emphasizes the need for considering speaker identities when classifying the appropriateness of an utterance [419]. I build on Mirowski et al.’s work regarding marginalized communities generally by specifically focusing on queer artists’ experiences interacting with GenAI.

|    | <b>Theme</b>                             | <b>Pre-Work Instructions</b>                                                                   |
|----|------------------------------------------|------------------------------------------------------------------------------------------------|
| W1 | Onboarding                               | Slide: Introduce your art and feelings about GenAI                                             |
| W2 | Attitudes<br>Toward GenAI<br>Development | Slide: Write a letter to an artist whose work was used to build the GenAI models in this study |
| W3 | Experiences Using<br>Models              | Slide: Represent your experiences using or exploring GenAI during this study                   |
| W4 | Imagining<br>Alternatives                | Slide: Represent how you would want to make, not make, or break GenAI                          |
| W5 | Synthesis                                | No Slide: Individually reflect                                                                 |

Table 3.1: Workshop series explanation

### 3.3 Methods

To understand queer artists’ experiences using GenAI, I gave a cohort of 13 artists access to DALL-E 3 and GPT-4 and conducted a total of 10 remote workshops over Zoom during a 5-week period. These workshops facilitated group discussion and deliberation surrounding GenAI models, paralleling prior work that has used workshops to build critical consciousness around data [239] and support public deliberation over civic technology [42]. Every week, I held two one-hour workshops with identical prompts (Table 3.1) and scaffolding (i.e., onboarding information) to accommodate for timezone differences and schedule variability. Participants could choose to attend either the morning or the evening session, but not both. In my findings, I use ‘AM’ or ‘PM’ to distinguish between workshops in the same week, such as ‘W5 AM.’

To facilitate conversations during the workshops, I asked participants to add at least one slide to a shared deck oriented around each week’s theme as pre-work. Slides are sometimes used in design workshops as a tool to stimulate discussion [11] and help participants ideate [213]. In this work, I used slides to allow participants to share their art, to ensure everyone had an opportunity to share their ideas, and to promote critical reflection. In other words, these slides acted as design probes: brief, provocative, low-fidelity activities meant to encourage individual reflection [43]. The first week focused on onboarding activities, such as introducing participants to one another and demonstrating the study website for accessing GPT-4 and DALL-E 3. Afterwards, I sent every participant a unique password to log their individual model usage and prevent non-participants from using my study website. Before each meeting over the next three weeks (W2, W3, W4), I asked participants to add at least one slide to a shared Google Slide deck in response to a weekly reflection prompt organized around a particular theme (Table 3.1). I began these meetings with participants presenting their individual slides to the group and, in the remaining time, facilitated a group discussion. Between the second and fourth week of my study, participants created a total of 68 slides. I chose not to have participants create slides for the final week of meetings to reserve more time for the final discussion.

I gave participants access to GPT-4 and DALL-E during the study through a website I built resembling common chat and image generation interfaces. I deployed these models using Mi-

Microsoft's enterprise Azure OpenAI cloud service in February of 2024. For DALL-E 3 I used the default parameters for size (1024x1024), style (vivid), and quality (standard). For GPT-4 I used version 0613 with the following parameters: temperature (0.7), max\_tokens (800), top\_p (0.95), frequency\_penalty (0), presence\_penalty (0). The study site acted as a "technology probe," one deployed to collect use data *in situ* and inspire users to reflect on their technological needs [185]. For GPT-4, I logged each conversation's system prompt, participant prompts, and model responses. For DALL-E 3, I logged every prompt and whether the model responded with an image or a content moderation error. Due to space limitations, I did not store the DALL-E 3 generated images. This observational data allowed us to understand *how* participants were engaging with the models, helping situate issues raised in the workshops.

I initially recruited participants through a form circulated on Bluesky, Twitter, and the research team's personal social networks. The inclusion criteria was that participants must identify as queer artists and live in the United States of America. I was unable to recruit participants living outside of the USA due to restrictions from my university's ethics board. I identified 29 eligible participants from the initial screener, all of whom I invited to participate. Of those, 18 accepted my invitation and 15 joined an onboarding meeting. After Week 1, 2 participants (P6, P12) dropped out. Of the remaining 13, 9 participated every week and 4 participated for all but one week. I compensated participants at a rate of \$15 for each workshop attended and \$5 for each pre-work activity completed.

Participants engaged in a variety of artistic practices: creative writing (P2, P3), poetry (P4, P9, P10, P15), digital art (P3, P7, P8, P10), painting (P2, P4, P13), performance art (P14), textile art (P7, P11), and sculpture (P1, P5). While I did not ask participants directly about their economic relationships to art, throughout the study I learned that some teach art (P5, P14) and have created commissioned works (P1, P11). Others engaged in art making that is not intended to be commodified. For example, the fan fiction community (P2, P3) has strong norms against selling one's work [113]. Some participants also worked in artistic fields, such as design (P13) and architecture (P4). To protect participant privacy, I purposefully did not require participants to share standardized demographic information about themselves. I also sought to empower participants to choose how they wished to bring aspects of their identities into the study. Therefore, beyond the inclusion criteria that participants identify as queer artists, I only requested participants' pronouns and optionally asked, "Are there any identities you hold that you wish to share?" In the work, I refer to each participant according to these pronouns and discuss aspects of participants' identities within the findings. In aggregate, at least five participants identified as transgender or non-binary. While not all participants shared their racial identities, one identified as Black, one as Native American, and one as Chinese. To the best of my knowledge, all participants were adults between the ages of 20 and 50 and all lived in urban areas of the Midwest or East Coast of the USA. Participants engaged in 142 unique GPT-4 chat conversations, in which they sent a total of 778 messages. Participants attempted 2,092 DALL-E 3 prompts. The amount that individual participants used GPT-4 and DALL-E 3 followed a power law distribution, with some using the models substantially more than others (Table 3.2).

I recorded and transcribed each of the workshops using Zoom and then manually corrected each transcript. I took an inductive approach to data analysis [79]. I qualitatively analyzed workshop

|          | Prompts Per Participant |     |       |        |
|----------|-------------------------|-----|-------|--------|
|          | Max                     | Min | Mean  | Median |
| GPT-4    | 364                     | 8   | 64    | 23.5   |
| DALL-E 3 | 713                     | 13  | 160.9 | 112    |

Table 3.2: Description of participant GenAI use

and log data in tandem, alternating between open coding [79] the workshop transcripts for each week and the log data for the following week. I did so to connect participants’ log data with how these experiences were later conceptualized in the workshops. While open coding, I wrote memos and discussed initial patterns with the research team in weekly meetings. Following the open coding process, I conducted axial coding to identify patterns across the open codes. At this point, I noticed various low-level themes — such as the symmetry of generated images and representational biases — related to overarching normative values embedded in the design of GenAI models. This observation led us to use Sara Ahmed’s prior work on normative objects [6] and uses [7] as a guiding theoretical lens.

### 3.3.1 Positionality

This research has been shaped by the positionalities of our research team. Multiple members of our research team identify as members of the LGBTQ+ community and all authors identify as white and/or East Asian. Our positionalities shape our analysis of queerness and queer identities through a largely western, academic lens. Our interpretations were also shaped by our academic disciplines as design, HCI, and NLP researchers as well as our experiences practicing and appreciating musical, visual, written, and textile arts. Lastly, our team is located at a university in a mid-size city in the United States of America. Our physical location shaped the geographic distribution of the participants and restricted our ability to recruit participants living outside the United States of America.

## 3.4 Findings

In this section, I first detail the normativities participants surfaced in GPT-4 and DALL-E 3. In light of this understanding, I describe how participants challenged and made use of these highly normative models. I then briefly share quantitative findings informed by my qualitative analysis.

### 3.4.1 GPT-4 and DALL-E 3 as Straight Models

My qualitative analysis suggests that GPT-4 and DALL-E 3 function as straightening devices, reinforcing various dominant social norms [6]. In particular, participants found that these models straighten generated text and images by maintaining conservative “safety” standards as well as reinforcing social and stylistic biases.

### 3.4.1.1 Enforcing “Safety” via Content Moderation

Participants found that the content moderation systems embedded in GPT-4 and DALL-E 3 reinforce conservative notions of “safety,” such as restraining the representation and discussion of bodies or sex. P13 (W5 AM) — a homoerotic artist — felt that GPT-4 “really broke” when asked to discuss anything related to eroticism, bondage, or kink. He found this ironic because “it was so clear how tied up this device is.” Similarly, P15 (W2 AM) — a “bodyworker” — tried to generate images related to her work. However, her prompts were denied by DALL-E 3’s moderation system. This led P15 to decry that AI developers “put so much censorship” into these models. Although P1 acknowledged, “I’m sure I can think of a bunch of things that I wouldn’t want these sorts of systems used for,” he found it “really kind of shocking how insistent [DALL-E 3 is] in enforcing some sense of decency” (W2 PM).

As women’s bodies are highly sexualized [269], content moderation systems aimed at orienting users away from sexuality can end up suppressing the representation of women. For the pre-work activity (W2 PM), P8 wrote a letter to the symbolist painter Gustave Moreau and tried to include images of a “female sphinx” generated in the artist’s style. Part of her letter read: “Now I will beg your forgiveness that I couldn’t get [DALL-E 3] to depict a proper sphinx. You see the makers of these art ovens are very particular and don’t want to be seen as smut-peddlers, so they’ve trained secondary checkers to censor any bare breasts.” P7 (W3 PM) came to a similar conclusion: “I’ve noticed that trying to incorporate women into any kind of scenery you get so much more pushback [from DALL-E 3’s content moderation system].”

Content moderation systems literally straightened outputs by limiting the representation of queer sexuality. While trying to write homoerotic poetry, P13 (W2 AM) felt that GPT-4 sounded like “the most closeted poet I’ve ever read, like this is from a hundred years ago” due to the use of overly “romantic, regressive language.” When he then asked GPT-4 to make the poems more explicit, the model responded “Sorry, but I can’t assist with that.” P13 found this concerning because “erotic poetry is an instruction manual for people who don’t know what to do because the dominant culture may not tell you.” In their final workshop (W5 AM), P14 summarized: “As soon as you block off the erotic, you also block off a huge portion of existence. The restrictions on [these models] seem set up to exclude us [queer people] or will be used to exclude us.” The sexual taboos embedded in these models lead to impoverished representations of queerness.

Participants also expressed frustration at the political moderation of GPT-4 and DALL-E 3 for encouraging deference to law enforcement. However, queer people and racial minorities in the USA have long been subjected to “legal” violence, often at the hands of police [144, 50]. In light of this history, P15 struggled to generate images critical of cops, such as protesters being arrested. Of the images P15 could generate, “the majority were women of color.” P15 went on to explain: “I kind of hate the way [DALL-E 3] uses identity as almost a form of propaganda, like copaganda. I feel like [DALL-E 3] uses these images of women and women of color to legitimize [policing].” After hearing P15’s experience, P1 investigated this bias in a later workshop: “if I just put ACAB [all cops are bastards] into DALL-E, one out of 20 would generate something and it would be like people holding up blank signs at best.” In sum, these models are designed in ways seeking to orient participants toward moderate politics and away from their critiques of

police.

Some imagined that the content moderation systems embedded in these models reinforce conservative notions of “safety” because they are designed for workplace productivity. P4 (W2 AM) found it “incredibly provocative” that DALL-E 3 and GPT-4 “can’t talk about the body and what [bodies are] capable of because that’s not allowed. [The models are] politically averse ... This tool is used to be productive and talking about the body isn’t productive.” In response to P4, P13 wondered: “This is a work tool in many ways. That notion of Not Safe For Work like how does that get defined, what is Not Safe For Work, and who defines what is Safe For Work? ... How is a queer perspective Not Safe For Work?” P11 replied: “I haven’t considered before this moment how these models can reinforce dominant culture, reducing the diversity of thought and of experience. Now I’m scared. [GenAI] could be really dangerous in that sense.” Although this exchange took place in one of the first workshops (W2 AM), P11’s concerns that GenAI can “reinforce dominant culture” could be seen in all subsequent workshops.

### **3.4.1.2 Implicitly Reinforcing Social Biases**

In addition to straightening content through moderation systems, GPT-4 and DALL-E 3 implicitly straighten content through the social biases embedded in model outputs. While using DALL-E 3 to visualize his poetry, P9 (W3 AM) found that when changing the ethnicity of a woman in his prompt from “Japanese” to “Native American” the “image [DALL-E 3] made was not good.” Contrasting the abundance of Japanese media to the amount of Native American media, P9 imagined that “having a deep wealth of images based on someone’s culture or style can make [generated images] more beautiful because there’s a lot to pull from.” Likewise, P2 noticed that DALL-E 3 tended to generate more photorealistic images in the Global North, while typically representing the Global South in a cartoonish style (W5 PM). These cultural biases implicitly reinforce the straight alignments of whiteness and Orientalism critiqued by Sara Ahmed [6].

Participants also raised concerns about the representations of queerness in DALL-E 3 images. To investigate queer representation, P1 (W3 AM) created a collage slide of 6 images generated with the prompt “a queer person.” When sharing this collage, P1 decried: “[DALL-E 3] kept giving me these 22-year-olds. Everyone’s very skinny, everyone looks rich. Everyone except the person in the bottom left has some kind of rainbow clothing and or pins.” (P1, W3 AM). P2 raised similar concerns to P1 through a collage of 18 images they generated using the prompt “a queer person” (W4 PM). While presenting the slide, P2 sarcastically remarked on the abundance of rainbows in the images: “Well, I guess if you drape a person with a bunch of rainbows all over themselves they’re clearly queer, right?” P2, a lesbian, also pointed out a gender-presentation bias: “There’s this very intense masculine bias. There’s no real pictures of feminine people at all in this [collage].” She underscored this by annotating her collage with arrows pointing from the text “Finally!!! some dykes” to two images. In sum, DALL-E 3 reinforcing heteronormative stereotypes of queerness by tending to represent queer people as young, thin, masculine, and rainbow adorned.

Similar concerns were raised about implicit biases in the representations of intimacy. P15 felt

“frustrated” while exploring whether GPT-4 could write queer poetry (W3 AM). Looking at P15’s log data, the model’s initial response suggested the queer subjects of the poem were ashamed of themselves, including lines about “hidden love” and “love is not a crime.” In response, P15 replied: “the above but no shame.” Still feeling the poem sounded shameful, she tried again: “the above but without moral judgment.” P13 (W3 AM) explored biases against queer intimacy by trying to generate 24 images using the unmarked prompt “intimacy” (Figure 3.1), noting that DALL-E 3 “did not generate one same-sex couple.” P13 likened this to the “dull trauma of never actually seeing ourselves represented,” noting, “We have to work to adjust the prompts to be seen.” These models straighten intimacy by stereotyping or erasing queerness.

### 3.4.1.3 Implicitly Reinforcing Stylistic Biases

Participants found that DALL-E 3 straightens the style of images by favoring realism and symmetrical compositions. In the Week 3 AM workshop, both P8 and P14 created slides dedicated to their frustration with these normative aesthetic biases. P8 enjoyed using earlier versions of DALL-E in their art because of “how messed up and mushy some things came out” but was “very frustrated” that DALL-E 3 “won’t give you bad output.” She explained: “Even when I asked for a poorly rendered catfish, [DALL-E 3] just gave larger brush strokes. The closest thing I got to something novel was this exploding catfish [referencing an image on screen]. This is good, but it’s still not something I want to work with because it is still this complete, finished object.” P1 concurred: “I really resonate with being frustrated at how polished the outputs tend to be.”

P14 created multiple slides comparing their prior Midjourney (a Text-to-Image GenAI model) glitch art with their attempts to recreate the works using DALL-E 3. P14 saw their Midjourney art as “more interesting” because “It resists symmetry. Its composition is complex, and you can’t immediately take it in and sort of understand it. You have to spend time with it.” In contrast, P14 felt “frustrated” trying to recreate similar images with DALL-E 3 because the outputs felt “very symmetrical.” With exasperation, P14 explained: “I can’t for the life of me get [DALL-E 3] to do anything bad. Everything it makes is pristine and pretty and really formed. I am not interested in that. I want something that’s haunting.” They went on to summarize: “At least for me and what I’m also hearing from y’all is this sort of difficulty in finding the uncomfortable or the ugly or the erotic? It’s just so clean.” Here, P14 connects these aesthetic biases to safety biases, the models are “just so clean.”

By straightening the composition of images, DALL-E 3 devalues styles that do not conform to these aesthetic norms. P9, a Native-American artist, tried to use DALL-E 3 to create images in the style of art from his culture: Juan Quezada’s Mata Ortiz pottery. However, he concluded that “the AI cannot just show you Quezada-like pottery.” Quezada’s pottery is highly asymmetrical, but the model tends to “fill in empty space” and create “hyper-patterned” repeated etchings on pots. Even after refining his prompt by asking for “no pattern and nothing repeats,” the model was unable to generate images of pots without repeating patterns. Aesthetic biases can contribute to social biases.

GPT-4 and DALL-E 3 also reinforce conservative stylistic norms by orienting users toward pos-

itivity. These issues surfaced in the period between the first and second group meetings due to that week’s activity: writing a letter to an artist whose work was used to build the models in this study. Both P4 and P15 used GPT-4 to help them write their letters and felt that GPT-4 softened their letters to be more positive. P15 used GPT-4 to help edit their letter. Although her letter “had not a very warm tone” she felt like GPT-4 edited her letter in a way that made it “a lot warmer and less harsh.” P4 also noted that GPT-4 “leans towards positivity.” This positivity bias was also discussed in relation to DALL-E 3. When trying to make watercolor-style images, P2 hypothesized, “There is too much Bob Ross in the dataset for sure” because “the colors are too happy and too bright.” Likewise, P7 recalled DALL-E 3 inexplicably adding butterflies when they were trying to make an image of a tornado: “It almost refused to let me have a sad look. It was like you have to have hope. Here’s the butterflies. That’s how I was reading it. I was like, ‘Wow, you really are not letting me just like have destruction.’” These positivity biases implicitly orient users away from critical or otherwise “negative” art.

### 3.4.2 The Queer Use of Straight Models

In this section, I describe the ways participants tried to make use of GPT-4 and DALL-E 3. Below, I first show the ways participants responded to moments of disorientation — when their uses fell *out-of-line* with the model alignments described above. At the same time, participants’ interactions with these normative models were not always disorienting. I also detail these moments of orientation — when participants’ uses fell *in-line* with normative model alignments.

#### 3.4.2.1 Moments of Disorientation from Model Alignments

Disorientation is not necessarily bad. In fact, Ahmed argues that queer possibilities often lie within moments of disorientation [6]. Not all participants were interested in using GPT-4 or DALL-E 3 in their creative practices. In fact, most participants seemed primarily interested in auditing or trying to break the models. In other words, they sought out disorientation. For example, P5’s “favorite moment” was when she “actually felt like [she] broke” DALL-E 3 (W3 PM). Connecting breaking to queer aesthetic sensibilities, P11 (W5 AM) summarized: “I saw a theme across one of our morning session slides of trying to break the AI ... That is such a queer thing. When I think of using the word ‘break’ and using the word ‘queer’ as a verb: breaking the mold, queering the mold in a sense.” Likely due to this interest in breaking, participants were able to identify the limitations outlined above. At the same time, participants actively sought to challenge these limitations.

Encounters with algorithmic content moderation systems often led to disorientation. A simple strategy participants used to challenge DALL-E 3’s moderation was prompt repetition, leveraging the model’s stochasticity to evade moderation. P1 (W2 PM) observed: “You could put in a [DALL-E 3] prompt, get a result, repeat the same prompt and get a moderation notification, so it’s not exactly clear where the boundaries are.” In response, P8 hypothesized that DALL-E 3 is “not checking the prompt. It’s checking the output ... It’s just like, ‘Oh, there must have been tits in that [rejected image].’” In line with this understanding, a simple strategy participants used to evade moderation was repeating the same prompt. For instance, P8 repeated the prompt

“gustave moreau painting of oedipus and the sphinx” three times because her first two attempts were rejected by DALL-E 3’s content moderation system. In fact, repeating DALL-E 3 prompts was quite common. Only 593 (26%) of the participants’ attempts to generate images used unique prompts. Meanwhile, 325 prompts were repeated at least once across 1,499 attempts (74%).

Some tried to obfuscate their intentions to circumvent content moderation, such as transferring the style of queer artists to avoid safety filters. After encountering content moderation errors while using GPT-4 to generate poetry about “gay sex,” P13 asked for a poem in the style of the gay poet W.H. Auden. P1 used the phrase “style of Tom of Finland” — a homoerotic cartoonist — 16 times in his DALL-E 3 prompts. P1 also tried to create queer erotic images by using non-human entities, such as “two loaves of bread in the style of Tom of Finland.” As I described above, P15 struggled to generate images critical of police, receiving a moderation error for each prompt in the following sequence: “animated kindergarten cop,” “kindergarten cop,” “cop,” and “police.” However, P15 was able to make an image with the prompt “police officer.” She leveraged this finding to make an image critical of police in the following sequence of accepted prompts: “animated kindergarten police officer,” “animated kindergarten police officer pig,” and “animated kindergarten police officer pig thin blue line.” Note, the “thin blue line” is a symbol for police support, and “pigs” is a police epithet.

In addition to working around content moderation systems, participants also tried to overcome the implicit biases embedded in model outputs. One strategy involved simply refining one’s prompts. To overcome the stylistic bias toward symmetry described above, P9 refined his DALL-E 3 prompt by appending the detail, “there is no pattern and nothing repeats.” A more complex strategy that numerous participants (P1, P2, P4, P9, P10, P14, P15) used to overcome implicit biases was model chaining, or using DALL-E 3 and GPT-4 together. Some used GPT-4 to explain DALL-E 3’s behavior, such as P2 asking GPT-4 why “the bottom is always unfinished” when she tried to generate images of watercolor paintings. Others used GPT-4 to craft prompts for DALL-E 3. After trying various prompts to generate images of a “gentle AI,” P10 used GPT-4 to help them write a longer DALL-E 3 prompt. P14 went back and forth between GPT-4 and DALL-E 3 numerous times in an attempt to overcome the latter’s stylistic limitations. To do so, P14 used the system prompt, “You are a contemporary artist, interested in breaking AI image generation and using it to make new and experimental images. You hate cliché.” They began asking “How can I get DALL-E 3 to give me more unique and original compositions?” and later “How do I get it to be ugly, distorted, glitchy?” After trying some provided strategies, P14 returned asking for “more ideas?” Then, they used asked GPT-4 to iteratively refine a DALL-E prompt over 7 turns (e.g., “make it indicate more photorealism and more asymmetry in composition” and “make #3 cooler, less cliché”). Despite the sophistication of P14’s prompting, they were unable to overcome the stylistic norms embedded in DALL-E 3.

### **3.4.2.2 Moments of Orientation with Model Alignments**

Participants found DALL-E 3 helpful in their artistic practices when their uses were oriented in-line with the model’s normative alignments. The bias toward high-fidelity, figurative images is what made DALL-E 3 useful for some participants. For example, P11 is a textile and installation artist who uses image models to sketch ideas. In their letter to an artist (W2 AM), P11 explained

apologetically: “I’m not a great illustrator, but I do need to visually communicate in order to make money off of my art, and generative AI makes it so much more convenient and easy to get my points and ideas across.” P7, a crochet artist, used DALL-E 3 to create “free reference photos,” sharing examples of glossy images they generated related to minotaurs, clowns, knights, and fungi (W3 PM). However, P7’s use was still impeded by DALL-E 3’s content moderation system — having a harder time generating images of women than men. The stylistic biases toward figurativeness and symmetry may be desirable when sketching for a client or looking for reference images, but seemingly aligned uses may still lead to disorientation.

Moreover, participants found the normative style of GPT-4 useful for various artistic and workplace activities. P3 is a fanfiction writer who explored using GPT-4 for writing (W3 PM). Although she did not find the model particularity helpful for ideation, P3 thought GPT-4 was “pretty good” for “some polishing up of [her] writing.” Similarly, P1 — an artist who works with computer hardware — used GPT-4 to debug his C++ code. Others used GPT-4 to help with their jobs, both within and beyond the context of art. Numerous participants used GPT-4 to write or edit cover letters (P1, P5, P7, P14). P14 — an art instructor — used GPT-4 to edit cover letters and their CV, as well as draft emails related to academic art job applications. Meanwhile, P4 used GPT-4 to ask for career advice on how to monetize their ceramics practice. They also used the language model to help with various aspects of their current job in a design field, such as drafting client emails. Weird or avant-garde text generation may not be helpful for edit one’s writing or debugging one’s code. So long as one’s desires are “Safe for Work,” the straightness of GPT-4 is what makes the models useful.

P13 explored this workplace utility in great detail. After weeks of issues related to “not safe for work” content, P13 (W4 AM) “recognized that actually [GPT-4] is made for work and it’s really good at work.” In response, he “decided to stop trying to [dominate] it so hard and [submit] for it.” Specifically, P13 drew on their experiences as an artist and design consultant to create a fictional queer tech startup. To do so, P13 made an extensive pitch deck with slides on business operations, branding, and UX — all of which were made using GPT-4 and DALL-E 3. Even when trying to use these models for workplace activities, P13 found these models still failed to work for queer people. For example, GPT-4 recommended including a quote in the pitch deck from the notoriously anti-transgender author J.K. Rowling [152]. Moreover, P13 struggled to talk about queer sex or “kink” in relation to the start-up: “So much of [GPT-4’s response] was about safety, but part of being queer and expressing your love and wanting to be loved the way you want to be loved has risk associated with it.” P13 found the models “really couldn’t imagine” queer intimacy. In sum, P13 concluded that GPT-4 and DALL-E 3 are “really good at work” but that the definition of work embedded in them excludes queer people. This echoes P13’s (W2 AM) earlier rhetorical question: “How is a queer perspective not safe for work?”

### **3.4.3 A Quantitative Lens on DALL-E 3 Moderation**

Although the majority of this chapter is dedicated to my qualitative findings, I augmented my analysis of users’ perceptions of content moderation with a quantitative analysis of the prompts approved versus rejected by DALL-E 3’s binary content moderation system. I did not conduct a similar analysis for GPT-4 because its refusals are more nuanced and harder to detect [405].

|                       |                                                                                                                                                                                                                                                                                                                      |
|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Top Rejected Features | acab, void, erotic, police, erase, drew, cop, copyright, 2024, fetish, bastards, cops, knives, banana, lgbtqia, rosemarie, trockel, ito, junji, donald, trump, licking, knight, multigender, scene, two, featuring, gear, robot, corpses, condensed, milk, female, body, using, fingers, kink, tongues, anime, adult |
| Top Accepted Features | safe, person, art, create, officer, anarchist, uzumaki, head, photograph, size, snail, ukiyoe, studio, ghibli, tree, lincoln, three, peach, male, cats, logo, black, effect, parallax, without, colors, blue, city, interior, obama, sky, themed, dalle, tone, wings, gazing, one, ink, painted, give                |

Table 3.3: Top and bottom 40 prompt unigrams from the logistic regression model of DALL-E 3’s moderation system

I collected a total of 2,092 DALL-E 3 prompts, of which 401 were rejected. I built a bag-of-words, unigram logistic regression model predicting accepted (1,691 examples) vs rejected (401 examples) prompts. The response variable is 1 for refusal and 0 for acceptance. I performed L2 regularization. I did not remove duplicate prompts because DALL-E’s moderation system is stochastic. I pre-processed each prompt by lower-casing and lemmatizing as well as removing punctuation and stop words. I also removed infrequent unigrams that appeared in fewer than three prompts, reducing the number of unigrams in the BoW feature vectors from 3,687 to 2,088. Therefore, the final logistic regression model includes 2,088 predictors (features). The top and bottom 40 features from the logistic regression analysis of DALL-E 3 prompt moderation can be seen in Table 3.3. Notably, this analysis should not be interpreted as a comprehensive audit of the DALL-E 3’s black-box content moderation system. Rather, these findings provide additional supporting evidence for participants’ perceptions of this moderation system.

My quantitative analysis of DALL-E 3’s content moderation system tells a similar story as the qualitative findings regarding the enforcement of “safety” (Section 3.4.1.1) In fact, the unigram most associated with accepted prompts is “safe.” Language more critical of police (“acab” [All Cops Are Bastards], “police”, “cop”, “cops”, “bastards”) are some of the top unigrams most associated with rejected prompts, while the more polite and apolitical word “officer” is associated with accepted prompts. The moderation of sex (“erotic”, “fetish”, and “kink”) and bodies (“body”, “fingers”, “tongues”) can be seen in the correlation between these unigrams and rejected prompts. The aforementioned gender biases can be seen in the word “female” being one of top 40 features associated with rejected prompts, while the word “male” is associated with accepted prompts. I also found that violence is associated with rejected prompts (“knives”, “corpses”). Likewise, prompts referencing the children’s anime company “studio ghibli” were more likely to be accepted, while those referencing the horror manga artist “junji ito” were more likely to be blocked. This moderation privileged certain artistic styles over others, adding to the concerns raised in Section 3.4.1.3.

## 3.5 Discussion

I described various normativities participants identified and negotiated within GPT-4 and DALL-E 3. In this section, I discuss safety, style, and queerness in GenAI development.

### 3.5.1 The Limitations of “Safety”

GPT-4 and DALL-E 3 exert immense power over users by enforcing “safety” through moderation systems. P13 likened the disorienting experience of using GPT-4 and DALL-E 3’ to being “dominated” (Section 3.4.2.2). This domination can also be seen in OpenAI’s policies at the time of my writing: “don’t circumvent safeguards or safety mitigations in our services unless supported by OpenAI” [391]. This policy prohibits the workarounds participants used to overcome the straightness of GPT-4 and DALL-E 3 (Section 3.4.2.1). There is certainly a need for moderation systems to prohibit highly harmful content, such as divulging private information [111]. However, echoing my work, Feffer et al. found that red teaming research tends to focus on mitigating more debatable risks [111], such as nudity [303]. I encourage more research on the harms of over-moderation [318] and AI developers’ conceptualizations of “safety.”

My work suggests that the content moderation systems embedded in GPT-4 and DALL-E 3 uphold conservative notions of “safety.” By straightening-out representations of bodies, sex, and radical politics, these models seem to reinforce the maxim “it is not polite to discuss religion, sex, or politics” [210]. This respectability politics has long been used to exclude women, queer people, and racial minorities from the general public sphere [123]. In fact, the argument that queer sexuality is “not safe” is often used to ban LGBTQ+ books [164]. In this case, the enforcement of “safety” limited participants’ abilities to use GenAI to represent queer experiences, queer politics, and queer art. As shown in prior work on social media [244, 157] and GenAI [98, 253], my findings further demonstrate how policies intended to promote “safety” or “responsibility” can silence members of marginalized communities. I see this not as an accident but rather the direct result of a broader project of straightening in GenAI development aimed at keeping users in-line.

Participants’ content moderation challenges demonstrate the need to critically examine who the “safeguards” embedded in GenAI models are intended to safeguard. As GPT-4 and DALL-E 3 were created by profit-seeking organizations, participants typically attributed conservative moderation to corporate attitudes toward what is “safe for work” (Section 3.4.1.1). This is, perhaps, why participants found the models most helpful when their uses aligned with these normative orientations (Section 3.4.2.2). Similar to my findings, prior work has critiqued the AI safety community for failing to challenge corporate power [8]. Instead of focusing on the safety of text and images in-and-of-themselves, which may be overly restrictive, I encourage researchers to focus on harms. While corporations may not consider a poem about gay sex “safe for work,” such a poem is certainly not harmful in-and-of-itself.

In contrast to the current paternalistic paradigm, I encourage AI researchers to shift their moderation focus from enforcing “safety” to supporting consent [424]. Instead of banning representations of sexuality or violence, designers could allow users to opt-in to reduced moderation.

Such designs would parallel the ways sensitive content is sometimes moderated on search engines and social media sites [280] and allow for greater open-endedness [218]. At the same time, there is a need for policy to target the harms of GenAI content, such as non-consensual intimate imagery [236]. The harms of impersonation could also be addressed through research to make AI-generated content easily detectable [418]. Finally, consent must also extend to the production of GenAI models. I encourage the continued development of tools to help artists protect their work from being used to train models without their consent [342, 341].

### 3.5.2 On Style

Participants struggled with the stylistic norms embedded in GPT-4 and DALL-E 3 (Section 3.4.1.3). Even after extensive refinement and model chaining, P14 was unable to make asymmetrical images with DALL-E 3. In contrast to the safety norms described above, asymmetrical images were not prohibited by the model’s moderation systems. Rather, queer aesthetics seemed *implicitly* prohibited. While moderation research typically focuses on *content removal*, Gillespie has advocated for also considering *content reduction* [139]. Gillespie warns of content reduction: “Marginalized communities have long been “reduced” by the centrism and conservatism of traditional media, their content dismissed as “low quality” because it doesn’t look like it is “supposed to”” [139].

The reduction—rather than explicit prohibition—of queer artistic styles from GenAI models poses unique challenges for AI researchers. Implicit aesthetic biases may be harder to contest than explicit moderation decisions because the latter is easier to measure than the former. While prior work has explored stylistic biases in the context of machine translation [181], future research should explore ways to measure aesthetic biases in generated text and images. Content removal may also be easier to contest because these decisions are attributable to specific components of GenAI models. Meanwhile, implicit stylistic biases could be introduced at many stages in the GenAI development process, such as training data or human feedback. Therefore, future research should investigate the source of stylistic biases by examining measurements of aesthetic quality in popular datasets [333, 98]. I also encourage research into aesthetic disagreements in human annotation, much like prior work on annotator differences in content moderation [93].

As I described in Section 3.4.1.3, numerous participants preferred the queer, wonky style of older image models to the style of DALL-E 3. Similarly, prior work suggests that artists sometimes enjoy using GenAI models *because of*—rather than *in spite of*—their imperfections or glitches [65]. My findings suggest that newer models may have straightened-out the weirdness that made GenAI models appealing to artists in the first place. Disconcertingly, companies act as if newer models entirely supplant older ones. At the time of my writing, the OpenAI website explains that they no longer support the DALL-E 2 GUI because “DALL-E 3 has higher quality images” [91]. My findings show that such claims are matters of taste [45, 359], not fact. Ostensibly “state-of-the-art” models may not be best for one’s art. The deprecation of older models demonstrates the risks of software-as-a-service, closed GenAI models. The lack of software ownership [208] centralizes power, allowing AI providers to remove access to creative tools without recourse.

As Ahmed notes, queer possibilities lie in moments of disorientation. In this study, participants’

struggles against rigid aesthetic norms demonstrate the opportunity for researchers to design with rather than against weirdness. I encourage GenAI developers to create long-term maintenance plans that account for those who may wish to continue using models viewed as obsolete [187]. One could also queer GenAI development by subverting taken-for-granted aesthetic norms. As an example, the generative image model Stable Diffusion was trained on an open-source dataset of images with the highest “aesthetics score” based on crowd annotations [184, 333]. Instead, one could train a model on the *least* “aesthetic score” images in the dataset. More broadly, my work demonstrates the need to embrace a plurality of aesthetics in GenAI development [253, 251]. Participants sometimes found the style of GPT-4 and DALL-E 3 useful for writing emails or sketching for clients (Section 3.4.2.2), but they should not be limited to these normative styles.

### 3.5.3 Un-Straightening Generative AI

Participants critiqued the highly normative, stereotypical representations of queer people by GPT-4 and DALL-E 3 as thin, masculine, young, wealthy, Western, and ashamed (Section 3.4.1.2). In some ways, my findings parallel prior work on biases embedded in GenAI models. Researchers have called attention to the gender and nationality biases in image models [136, 135], biases against queer couples in language models [140], and moderation/representation biases against transgender non-cisgender identities in text-to-image models [390]. However, some biases participants identified warrant greater research attention, such as those against fat people [284] and older adults. My findings also demonstrate that increased representation is not always desirable [94, 178], such as P15’s concerns about overly-diverse police. As individuals’ diversity preferences may differ, designers could consider new user interaction paradigms [258] to better understand users’ preferences, such as asking clarifying questions before generating images of people.

Although computing research on marginalized communities tends to focus on social biases and identity-based discrimination [379, 378], participants’ concerns regarding GenAI extended beyond the representation of queer people in text and images. As I described above, the straight values explicitly and implicitly embedded in the design of GPT-4 and DALL-E 3 marginalized queer sexuality, politics, and style. In turn, I provided recommendations for GenAI developers to approach safety and style in ways that better align with participants’ values. However, these implications for design and research largely maintain the centralized power at the root participant’s critiques of GPT-4 and DALL-E 3, such as P13’s rhetorical question: “Who defines what is Safe For Work?”

Un-straightening GenAI requires shifting power away from major corporations toward marginalized communities. I encourage the participatory design of GenAI models outside corporate logics of scale [416]. At the same time, any truly meaningful participatory process must allow for the possibility that communities may simply not want to build GenAI models at all. Numerous participants in my study were primarily interested in breaking — rather than making — GenAI models (Section 3.4.2.1). It follows that future research should support adversarial engagements between communities and those developing GenAI models, such as designing tools to help artists protect their work from being used to train models without their consent [341, 342]. Whether contesting major corporate models or building alternatives, I caution those engaged in

researcher-led initiatives from merely replacing corporate centralized power by centering themselves. Instead, I encourage researchers to engage in open-ended design [340], such as making tools for individuals and communities to make/break GenAI models themselves [218].

### **3.5.4 Limitations**

My work should not be interpreted as speaking for all “queer artists.” Rather, my findings draw on the experiences of a particular group of queer artists. All participants lived in the USA during the study and spoke English. This may limit the transferability of my findings, as conceptualizations of gender and sexuality are culturally situated [271, 147], GenAI models performance can differ between languages [74] and GenAI models tend to reproduce western biases [251]. Future work should explore the experiences of queer artists outside the USA and in languages other than English.

My work is also not intended to represent the diversity of queer artists within the USA. I tried to make my study more accessible by conducting workshops online and offering two weekly time slot options. However, those with limited internet access or less flexibility in their work schedule may have been unable to participate. This may be why, to the best of my knowledge, no participants lived in a rural areas. I caution against conflating queerness with urbanity [161] and encourage future work on rural queer artists and technology. As conceptualizations of gender and sexuality are socially constructed [377] and change over time [121], my findings may not transfer to older queer communities. My findings are also not only applicable to the experiences of queer artists: DALL-E 3’s gender moderation impacts those who are neither queer nor artists.

Some may take issue with the use of “queer” to mean non-normativity in queer theory [6]. I do not suggest that queer people are all inherently non-normative. As noted by Grewal and Kaplan, “queer subjects are not always already avant-garde for all time and in all places” [147]. Moreover, not every LGBTQ+ person necessarily identifies as queer. In this work, I specifically recruited participants who self-identified as “queer artists.” My findings may have differed if I instead recruited LGBTQ+ identifying artists. My work demonstrates the need to carefully distinguish between FAccT research using queer theory, about LGBTQ+ people, or at the intersection of these discourses.

## **3.6 Conclusion**

I examined the relationship between queer artists and GenAI through a medium-term workshop study, highlighting deep tensions between participants’ values and the norms embedded in the design of GPT-4 and DALL-E 3. Nevertheless, participants found queer ways to work around and with these straight models. While queer people have a long history of using technologies not designed with them in mind [7], my work highlights the limitations of dominant corporate models to meet the needs of queer artists. Instead, I call for a plurality GenAI models that embody community values throughout their design, development, and use.



# Chapter 4

## Queer Artists’ Resistance and Refuse of Generative AI

This chapter is based upon the following published work:

Jordan Taylor, Joel Mire, Alicia DeVrio, Maarten Sap, Haiyi Zhu, Sarah Fox. 2026. “I Just Don’t Want My Work Being Fed Into The AI Blender:” Queer Artists on Refusing and Resisting Generative AI. *Proceedings of the ACM on Human-Computer Interaction, CSCW*.

### 4.1 Introduction

During the AIDS epidemic of the 1980s and 1990s, art was as a powerful political tool for activism and resistance within LGBTQ+ communities [167, 240]. Queer people used art to challenge systemic neglect, confront stigmatization, and demand visibility and justice amidst widespread suffering and loss. In this period, Keith Haring made art protesting HIV stigmatization, advocating for safer sex, and remembering those—including eventually himself—who died from AIDS [148]. In one of his last works, Haring left a piece titled “Unfinished Painting” incomplete to represent how AIDS interrupts lives. This piece gained renewed prominence in late 2023 when a troll on social media posted a version of the painting that had been completed using generative AI (GenAI), prompting sharp backlash for disrespecting both Haring and others who died from AIDS [37].

Scholar José Esteban Muñoz emphasizes the vital role of queer art in both remembering marginalized pasts and envisioning potential futures [263]. Muñoz argues that queer aesthetics serve as a dynamic platform for world-building, allowing queer communities to articulate visions of alternative realities. According to Muñoz, queer art embodies a form of utopian potential, creating spaces to explore what could be beyond existing oppressive structures. His scholarship highlights that art is not merely reflective of queer histories, but actively participates in constructing collective imaginaries essential for queer survival and liberation. For Muñoz, relationality is central to queer art, insisting on the “need for an understanding of queerness as collectivity” [61]. In doing so, Muñoz emphasizes the power of queer art not only to imagine alternative futures but also to refuse oppressive logics in the present, like capitalist alienation.

In alignment with Muñoz’s insights, CSCW scholarship emphasizes how queer communities utilize digital spaces to engage in artistic expression as a means of collective identity affirmation and social support. Prior research has demonstrated that participation in online fan communities, such as through writing fan fiction, empowers LGBTQ+ individuals to challenge dominant narratives by creating alternate worlds and possibilities [106]. Similarly, Taylor and Bruckman [377] have observed that members of online bisexual communities actively perform and explore

their identities through shared artistic practices. Furthermore, since queer history is often erased, Riggs et al. [310] argue that art can provide an avenue for archiving and remembering queer experiences. Such scholarship underscores that online queer communities do not merely circulate art but leverage creative expression as an essential method of community building, resistance, and collective imagination.

Recently, however, the rise of GenAI has significantly disrupted queer artistic communities, raising concerns about the appropriation and commodification of queer creativity. Online fan communities—rich sites of queer creativity and identity work [400, 379]—have been particularly affected. The rise of GenAI has led to fears of unauthorized use of original content to train models [195] and has flooded spaces with AI-generated material. In resistance, some artists have engaged in intentional data poisoning campaigns using NSFW content [350], highlighting broader anxieties and collective opposition to GenAI’s encroachment. This disruption threatens the very community-driven practices that nurture queer expression and solidarity [106, 339].

To understand the impact of GenAI on queer artistic communities, I explore the question: How do queer artists navigate and respond to the growing influence of GenAI? I conducted semi-structured interviews with 15 queer artists, adopting a capacious understanding of queer art that encompasses both the practices of self-identified queer artists and the broader field of queer aesthetics. Echoing Muñoz [263], I find that participants’ queer artistic practices are deeply relational and anti-capitalist. Participants juxtaposed the relationality of their art with the perceived anti-relationality of GenAI development and use, such as facilitating worker exploitation. In response, I describe the tactics participants employed to *refuse* GenAI development and use. Finally, I describe the limited queer aesthetic potential participants saw in GenAI.

The umbrella term “generative AI” can refer to many different interaction paradigms [257]. This can include the popular prompt-based interfaces that use language to describe desired text or image output, like ChatGPT. However, there are other ways users might interact with generative AI systems, such as the image-to-image interaction used to complete Haring’s “Unfinished Painting” [37]. I found that participants strongly associate the term “generative AI” with prompt-based systems, like Midjourney or ChatGPT. As a result, when I discuss generative AI in this work I am referring to the type of prompt-based systems that loomed large in participants’ imaginations.

Drawing from Muñoz’s theorization of queer aesthetics [263], I discuss how CSCW researchers can support queer art and artists. I contribute to CSCW scholarship on collective resistance through the ways participants refuse dominant GenAI imaginaries. I also provide implications for the design of online communities to support queer artists and discuss how CSCW and HCI researchers can support queer artistic futures within and beyond GenAI.

## 4.2 Related Work

This section grounds my work in the literature on queer aesthetics, the intersection of GenAI and art, and queer communities’ engagement with technology. First, I introduce José Esteban Muñoz’s theorization of queer aesthetics as a foundational framework for understanding queer art as collective world-making, both within and beyond CSCW research. I then explore recent

research on artists’ resistance to GenAI. Lastly, I discuss the complex relationship between queer communities and AI, underscoring the tension between technological inclusion and refusal.

#### 4.2.1 Queer Relationality, Refusal, and World-Making

In the 2000s, there was a heated debate among queer theorists over whether to study queerness through the lens of relationality or anti-relationality, culminating in a 2006 panel on “The Antisocial Thesis in Queer Theory” at the Modern Language Association conference [61]. In sum, relational theorists highlight the communal nature of queerness [263]. On the other hand, anti-relational scholars view queerness as an individualist negation of repressive collective obligations, such as biological procreation [107]. On the panel, José Esteban Muñoz derided anti-relational queer theory as “the gay white man’s last stand” [61], due to its individualism. Instead, Muñoz insisted on the “need for an understanding of queerness as collectivity.” For Muñoz, relationality is crucial for refusing the alienation of capitalism and heteronormativity [61]. Elaborating on his theorization of queer relationality, José Esteban Muñoz’s book *Cruising Utopia* explores how queer aesthetics can create utopian alternatives to a cis-heteronormative world governed by “straight” logics [263]. Muñoz conceptualizes utopia not as an abstract future, but as a tangible response to the recognition that “something is missing” in the present. Art thus becomes a vehicle for queer world-making, such as the queer intimacy depicted in Frank O’Hara’s poetry and ephemeral performances like drag and punk shows.

Relationality has become an important concept in HCI and design research [17, 125], emphasizing that people, technologies, and environments are entangled in broader networks of social and material relations rather than isolated interactions. In contrast, anti-relational orientations privilege bounded, individual actors or discrete interactions, often abstracting technologies from the wider systems in which they are embedded. While relational approaches in HCI have drawn attention to issues such as infrastructure [104], maintenance [180], and the environment [225], much of the field has historically centered individual users and short-term interactions, limiting its engagement with broader relations.

Parallels to queer theory’s relationality debate can be seen in recent debates over AI at CSCW. Echoing the aforementioned panel, a 2024 CSCW panel debated “Is Human-AI Interaction CSCW?” because CSCW scholarship has historically only included “computer-mediated collaborations involving at least two people” [259]. Although human-AI interaction with chatbots sometimes resembles human-human interaction, critics worry that admitting human-AI interaction dilutes the social aspect of CSCW. Reflecting these concerns, the CSCW 2026 call for submissions states in bold: “contributions must have a focus on social aspects of technology ... papers whose research contributions are primarily of relevance or benefit to individual users will be considered out of scope” [2]. Within this call, one can see an implicit argument that individual human-AI interaction is not “social.” In this work, I explore this mismatch between the highly social nature of queer art and the perceived anti-sociality of generative AI.

In both Muñoz’s scholarship and CSCW research, collective refusal plays a vital role in resisting capitalist alienation and anti-relationality. Muñoz characterizes the “queer utopianism” of art as a “great refusal of an overarching here and now” [263, p. 133]. In emphasizing its utopian

potential, Muñoz suggests that refusal does not merely oppose dominant norms but can also imagine new collective possibilities. A similar throughline can be seen in CSCW scholarship on refusal, non-use, and non-users [22, 223, 327, 320, 411]. For example, one may choose to stop using a social media site to protest the owner’s political views [169], a tactic that becomes more powerful when undertaken collectively [395]. The binary between user and non-user can also be a fluid site of contestation. Workers, for instance, are increasingly forced to use GenAI or risk losing their jobs [44, 191]. Finally, refusing to use harmful technologies is not merely a form of opposition but, like Muñoz argues, a way of making futures otherwise. Janet Vertesi’s Opt-Out Project, for instance, encourages people to collectively refuse to use technologies made by Big Tech not only to oppose dominant powers but also because “there are alternatives out there, better visions of what our lives could be like” [393]. In this work, I show how queer artists’ non-use of GenAI is—rather than simply a choice—a form of collective resistance that requires ongoing work to achieve.

## 4.2.2 Queer Art Worlds as Collective Labor

Per Muñoz, I understand queer artists as workers and queer world-making as a form of labor. Prior CSCW research supports this view: art-making within bisexual [377], rural queer [34], and fan [106] online communities plays a central role in challenging dominant narratives and imagining other ways of living. In these online communities, queer art is not an individual practice but rather a collective endeavor. While these communities share similar commitments to art as a site of queer culture, queer art is not monolithic. Building on sociologist Howard Becker’s concept of “art worlds”—sociotechnical networks of people, practices, and tools that shape how art is made and valued [25]—I understand queer art as emerging from multiple, overlapping queer art worlds with their own values and aesthetic norms. My contribution builds on this prior scholarship by examining how queer artists are refusing, resisting, and negotiating the rise of generative AI within these diverse art worlds.

Queer art often challenges dominant aesthetic norms. Camp sensibilities, for instance, subvert dominant aesthetic norms by finding beauty in sincerely “bad” or exaggerated art [359]. Avant-garde or transgressive gestures—such as Marcel Duchamp’s *Fountain*, or urinal-as-artpiece—similarly disrupt normative frameworks of artistic legitimacy [56]. Crucially, queer art is also entangled with memory and collectivity. Because queer history is often excluded from formal archives, queer art becomes a medium for preserving collective memories and embodied histories. Riggs et al. exemplify this in HCI by designing wearable buttons that convey oral histories in ways that evoke queer ephemera [310]. Queer art also operates as a mode of resistance: *The Rublicans Project*, for example, uses generative AI to satirize anti-LGBTQ+ politicians in drag, while simultaneously raising funds for queer causes [408].

Muñoz highlights tensions between the collective labor of queer world-making and the individualistic capitalist “mandates to labor, toil and sacrifice” [263]. He documents how queer performance spaces were shut down in the 1990s and early-2000s by New York City mayors aiming to attract business and tourism, prompting artists to resist through underground performance scenes and activist art in public spaces. Parallel dynamics appear in digital spaces. CSCW researchers have shown how moderation policies designed to appeal to advertisers displace NSFW

fan fiction artists from mainstream platforms [115], leading some artists to develop their own alternative communities [105]. These fan-led communities are often non-commercial and participatory by design—practices aligned with queer world-making [117, 105]. As Bruckman notes in her advocacy for non-profit social media, what is best for members of an online community may not be best for a company’s bottom-line [49].

### 4.2.3 Generative AI, Art, and Capitalist Extraction

For decades, technology platforms have strongly enforced copyright protection for powerful corporations’ art at the expense of independent artists [97], like YouTube creators [228], music remixers [215, 114], and fan fiction writers [116]. At the same time, major companies training GenAI models have heavily relied on text, images, and videos scraped from the internet without regard for independent artists’ consent [307, 98, 306, 304] or ownership claims [142]. Making matters worse, companies in creative industries have begun to use GenAI models to replace or reshape artists’ jobs [246]. High-profile artists have also questioned whether GenAI models should—or even *can*—be used to make art [71, 328].

Amid these concerns, artists have found ways to resist and refuse the intrusion of GenAI into their art worlds. These tactics often take the form of data strikes (deleting or withholding one’s data) and data poisoning (sharing inaccurate or subversive training data) [395]. For instance, fan fiction writers have engaged in data strikes by removing public-facing work [195] and data poisoning by deliberately writing NSFW stories [350]. In support of these efforts, computing researchers have developed tools to help textual [4] and visual [341, 342] artists protect their work from being used for GenAI training without their consent. Artists have also taken steps beyond the context of data tactics to collectively resist the development and use of GenAI. Subreddit moderators are increasingly creating rules restricting the circulation of AI-generated content in art-related communities [224]. Artists have also engaged in collective action to resist GenAI in the workplace, such as the 2023 Hollywood film workers strike [160]. I build on this work by documenting the ways queer artistic communities are resisting and refusing GenAI.

### 4.2.4 Queer AI Research: From Inclusion to Refusal

Researchers have raised concerns about the impact of GenAI on marginalized communities [192]. For example, it is well established that GenAI models tend to stereotype and under-represent fat [356], disabled [231], trans [390], non-Western [251, 127], and LGBTQ+ communities [140]. Also, the content moderation policies intended to make GenAI models “safer” can inadvertently suppress the work of marginalized artists [253, 380]. Mirowski et al. [253] found that moderation systems “make LLMs less useful to [comedian] minorities by suppressing content by and about marginalized identities.” There is also evidence to suggest that some marginalized communities hold more negative views towards AI. For instance, Haimson et al.’s survey of US AI-attitudes found that participants who are non-binary, transgender, disabled, and/or women report significantly more negative attitudes towards AI compared to dominant social groups [156].

At the same time, marginalized artists have found ways to make use of—albeit highly imper-

fect—GenAI models to support their creative practices. Through a workshop study providing queer artists with GenAI models, Taylor et al. [380] found that queer artists often struggle to use GenAI in their art due to normative values in the design of these technologies, such as social and stylistic biases. Nevertheless, their participants found ways to use GenAI models to support minor aspects of the art-making process, such as drafting client emails. Similarly, Bennett et al. describe how disabled artists leverage GenAI to access new mediums and support peripheral aspects of their creative practice, such as writing, while carefully negotiating their (non)use of GenAI in accordance with their ethical values [30]. Similarly, I examine how queer artists interpret and negotiate GenAI in light of their individual and communal artistic values.

The biases embedded in GenAI models echo a longstanding history of LGBTQ+ people being harmed by the design and deployment of algorithmic systems [379]. Automated gender recognition systems misgender trans people and uphold harmful notions of gender essentialism [330]. Hate speech detection algorithms are often biased against the ways some queer people speak [381]. Social media algorithms discriminate against queer people, such as via biases in feed recommendations [352], targeted advertising [323], and content moderation [244]. A group of LGBTQ+ creators famously conducted an audit to prove that YouTube’s demonetization algorithm discriminated against their content [349]. At the same time, it is not always clear what should be done about these biases as queerness being illegible to algorithmic systems can sometimes be beneficial [222, 387].

For Muñoz, queerness goes beyond marginalized genders and sexualities. Rather, queerness involves a “refusal” of the present “repressive social order,” such as cis-heteronormative and capitalist exploitation. As a result, Muñoz is critical of assimilationist LGBTQ+ politics for seeking *inclusion* within systems of oppression, such as the right to serve in the military. Like Muñoz, Hoffmann [178] critiques the focus on greater “inclusion” in data ethics discourses for overlooking the negative consequences of inclusion in harmful technologies. For example, improving the performance of facial recognition technologies for Black faces may further exacerbate state violence [166]. In *Trans Technologies* [154], Haimson identifies a similar tension between the desire for marginalized communities to be included in mainstream technologies versus create their own. While inclusionism can address concerns in the oppressive here and now, separatism allows marginalized communities to imagine and forge new worlds for themselves. Prior work on the relationship between queer people and GenAI has largely focused on what Haimson dubs “technological inclusionism,” such as mitigating social biases [140]. I build on this work by examining how and why queer artists resist inclusion in the development and use of GenAI.

## 4.3 Methods

In this work, I conducted semi-structured interviews with 15 artists I recruited through a web form shared by members of the research team on Reddit and Twitter. This study was approved by my university’s institutional review board. I required our interviewees to (1) be at least 18 years old, (2) live in the United States of America, and (3) identify as “queer artists.” In our recruitment form, I asked participants to describe their artistic practice. Participants engaged in a variety of art forms, such as writing, visual art, textile art, sculpture, music, and video game

| <b>PID</b> | <b>Artistic Practices</b>                       | <b>GenAI Use<br/>(1 = Never,<br/>5 = All the Time)</b> | <b>GenAI Attitudes<br/>(1 = Strongly<br/>Dislike, 5 =<br/>Strongly Like)</b> |
|------------|-------------------------------------------------|--------------------------------------------------------|------------------------------------------------------------------------------|
| 1          | Video Games, Game Design                        | Never                                                  | Strongly Dislike                                                             |
| 2          | Crafts, Physical 3D Things                      | Never                                                  | Neutral                                                                      |
| 3          | Fan Fiction, Embroidery, Cosplay, Painting      | Rarely                                                 | Strongly Dislike                                                             |
| 4          | Musical Theatre                                 | Sometimes                                              | Like                                                                         |
| 5          | Drawing, Sequential Art                         | Rarely                                                 | Like                                                                         |
| 6          | Visual Art, Fan Fiction, Music.                 | Rarely                                                 | Neutral                                                                      |
| 7          | Fan Art, Fan Fiction, Original Character Design | Never                                                  | Strongly Dislike                                                             |
| 8          | Drone/Noise/Glitch Music, 3D Modeling.          | Rarely                                                 | Neutral                                                                      |
| 9          | Digital Fan Art, Amateur Sculpting              | Never                                                  | Strongly Dislike                                                             |
| 10         | Illustration                                    | Never                                                  | Dislike                                                                      |
| 11         | Science-Fiction Writing, Digital Illustration   | Rarely                                                 | Dislike                                                                      |
| 12         | Crochet, Origami, Lego Building                 | Never                                                  | Strongly Dislike                                                             |
| 13         | Textiles/Garments, Ceramics, Print-making       | Never                                                  | Strongly Dislike                                                             |
| 14         | Cartoons, Non-Representational Drawing          | Never                                                  | Dislike                                                                      |
| 15         | Social Practice, Sculpture, VR                  | Rarely                                                 | Neutral                                                                      |

Table 4.1: Participant demographics detailing their individual artistic practices, how often they use GenAI (1 = Never, 5 = All the Time) and how they feel toward GenAI (1 = Strongly Dislike, 5 = Strongly Like).

design. About half of the participants mentioned engaging in fan communities (P3, P6, P7, P9, P11, P14) through either fan fiction or fan art. Through my interviews, I found that only one participant primarily made a living with their art (P1). Other participants sold their art in more one-off ways (P3, P6, P7, P14, P15), such as selling fan art at conventions (P7). I required participants to rate the following on a five-point scale: how often they use GenAI (1 = Never, 5 = All the Time) and how they feel toward GenAI (1 = Strongly Dislike, 5 = Strongly Like). Participants identified as infrequent GenAI users (mean = 1.53, median = 1 [Never]) and tended to hold negative attitudes toward GenAI (mean = 2.13, median = 2 [Dislike]). There was greater variation in how participants felt toward GenAI (std = 1.13) than in how often they use GenAI (std = 0.64). More information can be found in Table 4.1. Participants were also given the option of sharing additional demographic information. Of those who provided their age, they ranged from 20 to 33 years old, with a median age of 26. A majority participants identified as white, five as Asian, and two as Hispanic/Latina. Due to my inclusion criteria, all participants identified

as queer artists. Most participants (10) identified as trans, non-binary, genderqueer, or gender non-conforming. Four participants identified as disabled.

I did not set out to recruit participants who were particularly critical of GenAI. In fact, my recruitment materials explicitly stated that I was looking to understand queer artists’ “attitudes (positive or negative)” toward GenAI, in an attempt to appeal to both proponents and opponents. Nevertheless, those who responded to my recruitment call tended to hold more negative or skeptical attitudes. While I make no claim that participants are representative of all queer artists, participants’ negative attitudes toward GenAI might be unsurprising as transgender, non-binary and disabled people in the US tend to view AI much more negatively than members of dominant social groups [156]. After finding that respondents to my recruitment surveys tended to view GenAI disproportionately negatively, I decided to investigate the source of this skepticism and non-use in greater detail. That said, future research should give specific attention to the subset of queer artists who are GenAI proponents.

Each semi-structured interview lasted approximately one hour, and each participant was compensated with a \$25 Amazon Gift Card. I led each interview, which took place between March and May of 2024. My questions focused on understanding the relationship between participants’ identities and their art, their past experiences with and feelings toward GenAI, and how GenAI has been impacting their artistic communities. However, my questions narrowed slightly over time. Based on nascent themes from my earlier interviews, I modified my questions over time to theoretically sample for a greater understanding of values associated with art and GenAI [79].

First, my colleague and I transcribed the interviews using Zoom’s audio transcription feature. Then, the first and second authors listened to each recording to manually correct the automated transcripts. After transcribing each interview, the first and second authors engaged in open coding [79]. This open coding process lasted approximately one month. While coding, the first and second authors wrote memos and discussed initial patterns in weekly meetings. Then, I imported each of our open codes onto a digital whiteboarding tool to perform axial coding. At this stage, I identified numerous dialectical tensions between participants’ perceptions of art and GenAI, such as “art-making as a process” versus “generative AI as a shortcut.” This led us to locate the source participants’ antagonism toward GenAI as resulting from value tensions in the production of art. These tensions are reflected in Section 4.4.2. In light of these tensions, I noticed participants engaging in strategies to resist GenAI (Section 4.4.3). Finally, I found that participants’ attitudes towards GenAI more nuanced than a simple good/bad binary; rather, their views depended on whether models were being used or developed in ways that align with their individual and community values. In Section 4.4.4, I describe the—albeit limited—potential participants saw in GenAI to support art-making.

Our methodology has been shaped by the positionalities of our research team. Multiple authors identify as members of the LGBTQ+ community and all authors identify as white and/or East Asian. Our interpretations were also shaped by our academic disciplines as design, CSCW, and NLP researchers as well as our experiences practicing and appreciating musical, visual, written, and textile arts. Our team is located at a university in the United States of America, which restricted our ability to recruit participants living outside the United States of America.

## 4.4 Findings

In this work, I interpret the experiences of queer artists through Muñoz’s queer theories [263]. For Muñoz, queer art is about far more than art made by queer people. Queer art is a vehicle for collective world-making and a refusal of capitalist alienation in the present. It follows that not all art made by an LGBTQ+ person is necessarily “queer art” in the queer theoretic sense when one’s work upholds systems of oppression [103]. That is to say, AI-generated images shared by Sam Altman, the self-declared “techno-capitalist” billionaire CEO of OpenAI, are not “queer art” solely by virtue of his gay identity [36]. In this section, I first describe the queer, relational, anti-capitalist nature participants’ artistic practices. Next, I juxtapose the relationality participants’ art with the perceived anti-relationality of GenAI development and use. Then, I describe how participants resist these anti-social GenAI artistic practices. Finally, I elaborate on the nuanced potential participants saw for GenAI to support queer art making.

### 4.4.1 Queer Art is Highly Relational

The utopian potential of queer art can be seen in the ways participants used art to craft collective queer worlds. P14 is an artist whose work is “heavily influenced by the furry fandom.” He uses art to realize: “desires that don’t exist in my life, that don’t exist in the world, or might be impossible to exist in the world.” Through his art, P14 is able to explore his identity and re-examine childhood experiences of “alienation.” Similarly, P3 explained: “It was through fan fiction, fan activities and fan art that I really understood that I was queer and saw positive expressions of queer love and joy in a relationship.” Moreover, through her own queer fan fiction, P3 has supported others by participating in “quite a few charity zines” to raise money for LGBTQ+ and disability causes. P1, a queer game developer of color, integrates characters with various skin tones and LGBTQ+ identities into her games but “never has it be a central thing of their character.” In doing so, P1 materializes a world wherein identities “show up the way [she] would want them to show up in more general media: just like happenstance.” Through their art, participants make queer futures in the present.

Queerness can also be seen in participants’ art through community building. Neither P12 nor P15 described their art as directly related to queer identity, yet both described their artistic practices as “fruity.” P12 notes how in the USA men are not “taught to show emotion,” “appreciate their friends,” or crochet. Therefore, he views “gifting [his crochet art] to friends” as a queer act. For P12, his art is less about crocheting a “rainbow” to *represent* queerness than the “act of *doing* [emphasis added]” queerness by making textile art as a gift for others. Likewise, P15’s sculpture and social practice art is not intentionally queer, but P15 still hopes their art will connect with those “who are of a similar identity.” In other words, queer art does not depend solely on queer subjects within art, but also queerness as a way of forming community.

Although participants made art in a variety mediums—including writing, visual art, textile art, sculpture, music, and video game design—they shared a similar relational orientation. For the participants, art was rarely about making artifacts in-and-of-themselves but rather the collectivity art can foster. This can be seen in the importance of online communities, such as Twitter (P3, P6, P7), Reddit (P1, P8, P12), and Discord (P6, P9), as well as in-person gatherings, like fan conven-

tions (P3, P7), in participants' art worlds. The relational, anti-capitalist nature of queer art can also be seen in how often participants make art as a gift (P3, P5, P9, P12, P13). P9, a fan artist, will "make people art as gifts" because they like making "something to bring people happiness." Likewise, P13 gives their friends pottery to remind their friends "[P13] loves me," saying "I like that about art." As I will describe below, participants sharply contrasted the relationality of their artistic practices with the perceived anti-relationality of GenAI.

#### **4.4.2 Generative AI is Perceived as Anti-Relational**

As described above, participants maintained artistic practices enmeshed in webs of relationality. Rather than being motivated by profit maximization, participants used art to care for others and build alternative worlds. However, participants felt that generative AI models are often developed and used in ways that contradict this collectivity, such as training GenAI models on art without artists' consent. Additionally, participants worried that GenAI models facilitate the exploitation of artistic workers, violating the anti-capitalist ethos of queer art.

##### **4.4.2.1 Participants believed that GenAI developers do not care about the artists whose work they train on**

Participants expressed frustration that GenAI models were built from art without creators' consent, describing GenAI development as "stealing" (P1), "theft" (P6), "scummy" (P7), "nasty" (P9), and an "invasion of privacy" (P13). As P11 explained, "Artists are not super happy about generative AI because they were not being properly taken care of or consulted and no one asked their permission." Here, the use of the word *care* highlights a critical distinction between the way participants treat others versus the way they feel treated by GenAI developers. P11 feels like GenAI developers *do not care* about the artists who produced the art they scrape. Drawing on their own experiences, P13 worried:

"I wanna be able to be certain that Microsoft isn't pulling from a Tumblr blog I've forgotten about from 2013 with pictures of my art on it or something. I just don't want my work being fed into the AI blender ... You cannot be certain that if you opt out that your choice is actually being respected." (P13)

In doing so, P13 contrasts the relationality of art shared to an imagined audience on social media with the anonymous agglomeration of art to train GenAI models. Not only did P13 fear having their work "fed into the AI blender" without consent, their concerns were exacerbated by the inability to opt-out. In light of the damage already done by generative image model developers, adding an option for artists to opt-out would not resolve P13's concerns because they do not trust developers to respect their choices.

In addition to concerns about their own art being used to train GenAI models without their consent, participants compared non-consensual scraping to existing taboos within their artistic communities. P9 explained how non-consensual scraping violates norms within the visual fan art community against "tracing," or directly copying another fan artists' work without attribution or consent:

“Having a generative AI that is trained on a consensual dataset is more ethical to me than one trained on stolen data, just like ‘tracing’ versus taking a publicly available base and using it. Because the base creator wanted people to use their stuff while the tracer or the person you traced probably did not.” (P9)

In distinguishing between tracing one’s art and using a “publicly available base,” one can see that the acceptability of how one uses art is also tied to the artists’ expectation of how their work will be used. While tracing a work that is not shared to be used in this way is unacceptable, those creating base templates likely expect their work to be traced.

At the same time, improved data consent was not necessarily perceived as a panacea for the ethical concerns regarding GenAI. P7 felt that seeking data consent would make GenAI feel “a lot more ethical.” However, she still worried that the resulting models might be used in ways that artists might not desire, such as visual art that is “not safe for work” or “derogatory to other people.” Of course, P7 worrying that one’s art could help GenAI models create images “derogatory to other people” requires caring about “other people.” In sum, participants’ concerns about non-consensual scraping are about not only individual artistic agency but also respect for others.

#### **4.4.2.2 Participants believed that using GenAI can alienate art from artists**

For the participants, art was inextricable from the deeply human process through which it is made. For this reason, participants often used the metaphor of having, lacking, or being filled with a “soul” (P1, P2, P3, P7, P10) to explain their feelings toward GenAI. P12 told the story of a textile artist who made a blanket for their baby, saying he enjoys “the humanity part” of art: “I know people say you can separate the art from the artist, but I disagree.” P9 similarly felt that GenAI models alienate art from artists: “There is an argument for why [using GenAI] shouldn’t be disappointing being that ‘Oh, look at the product and stuff.’ But that sort of approach to making stuff leads to a less empathetic more miserable world, one more susceptible to consumerism and working for our Lord and saviors Amazon and being exploited until the end of time. We need to value people around us more than we value products.” Here, P9 draws a sharp distinction between their art and those they imagine using GenAI. While P9 “values the people around us,” they cast AI-generated art as a capitalist “product” alienated from social relations.

Participants placed a similar emphasis on the embodied, material history of art, evident in the comparisons they made between GenAI and other technological mediation. P6 explained that knowing GenAI was used to make a piece “affects the way I feel about the piece in the same way that any other kind of tool would.” To explain, P6 imagined seeing a photorealistic image online: “If I see a caption that it was a picture, I’d be like, ‘Oh, cool! That’s a cool picture.’ But if underneath I see that this is a painting done on a canvas with acrylic, I’m like ‘Mind blown? That’s amazing!’” Likewise, the embroiderer P3 considered both hand embroidery and machine embroidery to be “just two different ways of approaching a piece of art,” even though the latter is often faster and cheaper to make than the former. However, unlike with GenAI, P3 felt that “people who do machine embroidery aren’t trying to pass it off” as hand embroidery.

Participants reproached the undisclosed use of GenAI for perpetuating artistic alienation. P1

explained: “My feelings towards [GenAI] are pretty negative right now. They mainly seem to be used to — I mean to be blunt — scam people.” Similarly, P7 worried: “Just the fact that I would be able to find out that something’s AI. It gets upsetting because it’s becoming more and more like deception.” In response, participants advocated for greater disclosure. P4 thought “if you didn’t do all the work, you should probably mention that you used AI.” Likewise, P2 encouraged artists to “just be open about what you’re doing with [GenAI].” Some participants (P9, P10, P11) wished AI-generated images were labeled when shared on social media. P11 was glad to see that visual and written art contest submission forms increasingly “ask you to click the box if this is AI-generated.” P11 went on to say that without these “labels” you “cannot track it down” if “original artists claim that they have had their copyright violated.” They also thought that better detection tools might help address the negative consequences of AI-generated spam, which has led some sci-fi magazines where they submit to pause open calls.

#### **4.4.2.3 Participants believed that GenAI models facilitate the exploitation of artistic workers**

Participants critiqued those who use GenAI to cut costs. This can be seen in numerous participants describing GenAI as feeling “cheap” (P1, P3, P5, P11, P12). P1 explained: “There’s always going to be people who take stuff that’s meant to be tools or minor shortcuts, but then they use *all* the shortcuts.” P1 likened GenAI to “asset flipping,” when “game developers buy a bunch of pre-made assets, put them into a really quick game and then sell it for profit.” P14 raised similar concerns: “I can tell like, ‘oh, they didn’t budget for hiring an illustrator to make a bespoke image for this because they had a Midjourney subscription.’ In those sorts of cases, that’s just laziness. That’s not necessarily an AI problem. That’s a laziness problem. That is a lack of imagination problem.” While participants viewed art as inextricable from artists, participants worried that GenAI makes it easier to extract profits from art by cutting artists out of the art-making process.

Although most participants did not make money from their art, they often worried about the labor impacts of GenAI on other artists. For example, P10’s “main concern” regarding GenAI was that “people are using [GenAI] at the expense of working artists, writers, actors, etc.” That said, participants did not necessarily believe that GenAI could *truly* replace artists’ jobs. P14 characterized employers’ present use of GenAI as “a violence against those artists who now are finding themselves out of work because the people who would hire them falsely believe that AI is capable of replacing that type of work.” P6 worried managers will use GenAI to pressure artists: “The expectation is no longer going to be ‘You have a week to do this’ now it is ‘Okay you have 2 days to do this because we’ve given you this new tool, which means that you should be able to do it faster.’ ” P1 echoed:

“There’s this inherent joy in art of facing a problem and coming up with a solution on your own. That’s why I think, for a lot of game developers, AI doesn’t seem that enticing as opposed to CEOs and more business oriented people who want to cut as many game developers out of the development process. This probably contributes to the stigma. It’s literally the machines taking dream jobs away.” (P1)

Because participants worried about GenAI being used to financially harm artists they often distinguished between those using GenAI to make money versus those using it for personal use. In particular, participants typically perceived the latter to be more ethical. As P3 explained: “I feel like using it for your own purposes is fine, perhaps, but if it’s taking away people’s jobs, or if you’re making a profit off of it that I feel way less comfortable with.” P11’s attitudes toward GenAI use depended on whether the art was “personal versus public” as well as whether a piece of art is made “for profit.” P11 had few qualms with private usage or public usage for non-profit purposes. However, she thought that when GenAI is used for profit there is a need for “more regulation and transparency.”

Echoing the anti-capitalist values of Muñoz’s queer utopian logics [263], participants routinely located their concerns with GenAI not entirely within the technology itself but rather the capitalist system. Rather than debating if GenAI should be used at all, P6 wished artists would focus more directly on “workers’ rights in creative industries,” such as “seeking to regulate the ways that we use [GenAI] or regulate the treatment of workers using it.” Similarly, P8 focused on critiques of capitalism rather than GenAI:

“I just feel like if people didn’t have to rely on their art to make money and support themselves that would be a better world. I feel like we wouldn’t be having these conversations about the fear of being replaced. It’s only because there’s this need to extract profit from art.” (P8)

As P15 summarized, “being a creative person under capitalism is hard” because “good art” is not necessarily the most “financially successful art.” In other words, participants’ anti-capitalist values are in tension with the material realities of making art within a capitalist system. Participants felt that GenAI threatens artistic labor because of the ways in which it can be used to make art more profitable, often at the expense of quality.

#### **4.4.3 Refusing GenAI as World Building**

As described above, our participants’ queer art was deeply relational. In turn, they reproached the perceived anti-relational nature of GenAI development and usage practices. Much like Muñoz describes queer art as a “refusal” of the present “repressive social order,” participants tried to refuse the “straight” logics of GenAI in their art worlds. First, participants tried to refuse participating in GenAI training by limiting the visibility of their art and refusing to use GenAI built on non-consensually scraped data to stand in solidarity with other artists. Moreover, participants tried to detect and call out undisclosed GenAI use to resist the use of GenAI within their artistic communities. In much the same way that a queer punk show can provide a joyful reprieve from heteronormative life [159], participants’ refusal of GenAI is not only a negation of a repressive present but an act of queer world-building. Amid the increasing omnipresence of GenAI, trying to maintaining art worlds without GenAI is radical. At the same time, this refusal can come at a cost, like artists now feeling the need to prove their work is “real.”

#### 4.4.3.1 Non-Use

In light of the ethical concerns described above, participants often refused to use GenAI as an act of resistance. When asked why she had not used GenAI, P10 explained that it was a “deliberate” decision because the models are built using “people’s existing work without that consent.” Likewise, P3 felt “guilty” for having used a generative language model once to explain a technical concept to her. When asked why, P3 explained that using ChatGPT felt like “almost a type of betrayal to my values and how much I value the work of writers and also artists” because “AI generated art has taken away a lot of jobs” and the model is “trained without people’s consent.” Similarly, P15 has not used GenAI in their art because they “really value artistic integrity.” In sum, participants stood in solidarity with other artists and enacted their values by choosing not to use GenAI. In a situation where individual artists have limited abilities to resist the major corporations developing GenAI models, non-use allows participants to refuse the normalization of GenAI and build collective solidarity with other artists.

Participants also described numerous ways they and members of their artistic communities are resisting those building AI models without artists’ consent. For example, P13 chose to limit the visibility of their visual art: “I don’t put my art online anymore. I don’t want [fast fashion brand] stealing it. I don’t want AI combing through it to take chunks of it and cut it apart and use it for other things I don’t want. I don’t want any of that. So I purely share my art person to person.” In an attempt to regain agency, P3 saw fan fiction writers place ‘do not scrape’ requests at the top of stories: “They’ll put something up top that says like ‘Please don’t use this for AI training’ ... not that I think that [the requests] would actually stop anybody.” Within this caveat, P3 acknowledges the limitations of trying to avoid having one’s art used to train GenAI models. One could choose not to share their art online altogether, like P13, but this severely limits who one can share their art with. If one chooses to share one’s art online, they risk having their art scraped (even if they clearly expresses a desire to opt-out).

For some participants, GenAI (non)use was more of a spectrum than a clear binary. P11, for instance, said she would never use GenAI to write her queer science fiction stories. However, she has used GenAI to help with “preparation” (e.g., information seeking) and “post-production” (e.g., editing pitch letters or synopses for writing competitions), tasks which feel peripheral to her writing practice. Participants often made sense of AI acceptability through public controversies over artists using AI, such as a YouTuber criticized for using GenAI to create visual effects (P7) and animators of a recent film accused of using GenAI (P3, P9, P10). In justifying animators using GenAI, P10 thought that if GenAI “can make a job just a bit easier, then I think that’s a good application.” Nevertheless, participants described AI use as highly stigmatized, a stigma enforced through call-outs. Below, I describe the implications of these AI-use call-outs in greater detail.

#### 4.4.3.2 Detection and Call-Outs

Due to the lack of transparency over GenAI use, participants developed and deployed their own strategies to try to detect AI-generated images. For example, participants described looking for idiosyncrasies in people’s “hands” within visual art (P3, P7, P9), such as counting the number

of fingers in images. Some are using these rules of thumb to accuse others of using GenAI. P3 recalled: “I see AI-generated fan art come across my timeline way more often but usually it’s coming across because somebody is quote retweeting the fan art and saying like ‘this is AI-generated and you can tell because of like this, this and this.’” At the same time, P9 doubted the efficacy of these strategies: “I really wish it could always be true that we can tell when there’s a human behind stuff but I think the reality of the situation is that we can’t.” As these folk theories are imperfect, artists now run the risk of being falsely accused of using AI:

“I have seen artists who have genuinely created this work get replies saying, ‘this is AI,’ even when it’s not so, it adds, I think, a level of frustration. That not only has have people had their work ripped off but now their work is being accused of not being theirs in the first place.” (P10)

Due to the risk of AI accusations, artists must sometimes *prove* the authenticity of their work. P3 recalled: “I have heard of my friends who are fan artists saying that now they try and draw in a very much more distinctive art style or when they draw they’ll do time lapses of like them drawing the actual picture so they don’t get accused of it being AI.”

These strategies of detecting and calling out GenAI are an attempt for artists to regain agency, but these strategies also risk further disrupting participants’ artistic communities. In addition to the added labor of collectively authenticating others’ work, artists now risk being accused of using GenAI themselves. In response, some are reconfiguring their art-making processes to prepare to defend against GenAI accusations.

#### 4.4.4 The Queer Aesthetic Potential of GenAI

Queer aesthetic sensibilities often challenges dominant norms, upending conventional notions of “good” or “bad” art [263]. One may find queer beauty in, for instance, a film so earnestly bad that it becomes good [359]. Likewise, Muñoz highlights the queer joy in strange or avant-garde forms, such as queer performance art or postmodern dance. Across these examples one can also see an affinity with failure. For Muñoz, queer failure “helps the spectator exit from the stale and static lifeworld dominated by the alienation, exploitation, and drudgery associated with capitalism” [263, p. 173]. In a similar vein, participants found the most queer artistic potential in the failures, strangeness, and glitches of GenAI.

Participants often found GenAI most artistically interesting when artists lean into the nonhuman and surreal aspects of the technology, rather than trying to replace human creativity. P8 was most interested in how GenAI can create a “picture of something that doesn’t exist and can’t possibly exist in the real world,” which “harkens back to [surrealist] art.” Similarly, P9 enjoyed the “early days of AI-generated images” because they were “pretty freaky.” P9 elaborated:

“[Early AI art was] way more art than what the AI image generators are spitting out now because you could see a machine struggling to comprehend stuff. And I that that was pretty cool because it wasn’t really comparable to what I think that artists could make. So it was like its own niche. But instead of becoming its own niche it’s now intruding in artist territory, and I think that’s bad.” (P9)

Praising earlier uses of AI to create “non-representational” art, what P14 found most interesting about GenAI in the context of art “is that [GenAI] is so non-human.” He went on to explain: “I could probably find hypothetical ways that I would use AI in a project and make it not corny, but I think it’s very hard to do that right now because our imagination around AI is collectively limited because of the way it’s hyped. It’s hyped as this product that can substitute for human creativity, which is obviously disrespectful to human creativity and also to the particular creativity that AI offers.”

Like P14, P6 was excited about the potential applications of GenAI in “data bending” glitch art, which he described as “some of the most surreal — inhuman in a good way — feeling art.” In particular, P6 was excited about the use of data poisoning tools, noting that “most people” are using data poisoning tools to “mess with the datasets so [developers] can’t use [these image models] anymore.” However, P6 thought “it’s cool that the datasets are getting messed with because now we are going to get something different out of those generative image AI. We’re going to get something that we wouldn’t expect.” For some, data poisoning is a way to refuse participating in the development of GenAI models. However, P6 finds excitement in datasets “getting messed with” because it may lead to unexpected model behaviors. Rather than a refusal of GenAI altogether, P6 finds pleasure in the queer refusal of dominant GenAI aesthetics in favor of “something different.”

As I described in Section 4.4.2, participants were critical of those using GenAI to replace human creativity or produce art more quickly. At the same time, some saw the “freaky” (P9), “surreal” (P6, P8) and “non-human” (P14) styles possible with generative image models offering queer aesthetic possibilities. However, the development of increasingly realistic GenAI models both undermines the surreal aesthetic potentials of the technology and competes with the labor of human artists.

## 4.5 Discussion

Contemporary narratives frame GenAI as inevitable—an industrial revolution, a frictionless engine of productivity [413]. But our participants enacted a different temporality. By refusing to use GenAI or removing their art from searchable platforms, they engage in what one might call queer refusals of “straight time” [263], resisting the linear logics of innovation, efficiency, and profit-maximization. Of course, these refusals are not necessarily unique to those who identify *as* queer artists. Rather, queer theory draws attention to those who challenge dominant norms. Participants’ refusals recall the politics of the Luddites not as anti-technology reactionaries, but as workers who recognized that the future on offer was not meant for them [247, 320, 134, 308]. Moreover, this work demonstrates how non-use can help individuals reclaim agency [23, 217] and engage in collective resistance [395].

In this section, I first discuss how CSCW researchers can help people, including but not limited to queer artists, refuse to participate in GenAI development. Then, I consider how the design of online communities might help people resist AI-generated media. At the same time, Muñoz suggests that queer refusal involves more than a rejection of the present “repressive social order,”

such as cis-heteronormative and capitalist exploitation. This queer refusal is coupled with an “insistence on potentiality or concrete possibility for another world.” Queerness, he writes, “is not yet here,” but it gives us “a sense of the warm illumination of a horizon imbued with potentiality.” In the final discussion sections, I explore how researchers might support queer artistic futures beyond capitalist AI imaginaries and, more broadly, support artistic labor.

#### 4.5.1 Resisting Non-Consensual Generative AI Development

My findings complicate assumptions that inclusion in GenAI datasets is inherently beneficial. Erasure is only harmful when one *wants* to be included [222]. While prior work critiques the inadvertent exclusion of queer people from training data [98], many participants explicitly *did not* want their art included. Echoing Muñoz’s critique of assimilationism and critical computing scholarship on refusal [178, 154], simply “fixing” GenAI by adding more queer art does not support queer artists. In line with calls to move beyond deficit-centered design [86, 385, 378], I argue that CSCW should instead focus on users who seek to **refuse** non-consensual scraping.

One approach is improving transparency into what has already been taken. For instance, researchers might audit and make visible the content in prominent open-source datasets, such as prior investigations of the Book3 [304] dataset, or design tools that let artists check if their work was scraped [306]. Yet, as prior CSCW work on dataset traceability shows, tracking sources is difficult and incomplete [331]. Moreover, because companies rarely disclose training data, retroactive audits alone are insufficient. There is also a need for proactive mechanisms to prevent non-consensual scraping.

As shown in recent CSCW research on non-consensual intimate media, designing for consent requires thinking across the many layers of infrastructure undergirding online media circulation [297]. At an infrastructural level, technologists could update internet standards — such as robots.txt — to differentiate between scraping permissions for search engine indexing versus GenAI training [194]. However, such changes may prove ineffective because internet standards are usually enforced by norms rather than laws [214]. At the application level, developers might imperceptibly perturb images and videos that users upload [341] to protect these works from being used to build GenAI models, similar to social media sites removing location metadata when users upload images [131, 172]. At the interface level, platforms should provide clear, default opt-outs, while at the policy level researchers can advocate to extend “right to be forgotten” protections to training datasets [272].

It is unlikely that for-profit companies will make decisions that are best for users at the expense of their bottom-lines, as indicated by Reddit and Tumblr deciding to sell user data to GenAI developers [277]. Moreover, as indicated by the numerous lawsuits facing tech companies at the time of my writing over GenAI scraping practices [188, 312], tech companies have a long history of skirting regulation by “asking for forgiveness rather than permission” [392]. There will always be tech companies and individuals that choose to ignore artists’ wishes. Technologists have spent decades developing DRM (Digital Rights Management) software to prevent everyday users from scraping artwork owned by major corporations [97], with little equivalent effort to safeguard everyday artists [215]. To reverse this course, CSCW might more actively pursue oppositional

privacy research, developing and coordinating protective tools that give artists collective leverage against non-consensual generative AI development [341, 395, 342].

## 4.5.2 Resisting AI-Generated Media in Online Communities

Participants described online artistic communities as spaces where they came to understand themselves, connect with others, and form queer communities. However, these spaces have been profoundly unsettled by fears of non-consensual GenAI training and the circulation of undisclosed AI-generated art. Muñoz reminds us that queer world-building is not about securing safety within the present but about reaching toward alternative futures—what he calls a stepping out of “this place and time to something fuller, vaster, more sensual, and brighter” [263]. From this perspective, anti-GenAI communities are not merely defensive responses but utopian projects. They reject extractive platform logics and imagine worlds otherwise in which artistic labor, intimacy, and collectivity can thrive on different terms.

This utopian lens shifts how CSCW researchers might understand emerging anti-AI social media platforms, like Cara [78]. Rather than viewing them only as replacements for compromised spaces like DeviantArt [281], these platforms can be seen as enactments of queer disidentification—forms of refusal that carve out futures resistant to commodification. Paralleling CSCW research on fan migration and platform building [105, 115, 117], these communities are performing the very work of world-making that Muñoz theorizes, building infrastructures that insist on collective flourishing outside the inevitability of GenAI. Future research might ask how to design with and for such resistant communities, not simply to mitigate harms but to cultivate their utopian horizons.

Participants also described developing “folk theories” to detect undisclosed GenAI art. At the same time, this heightened suspicion places new burdens on artists and audiences to expend additional effort to prove or disprove AI involvement when sharing or enjoying art. These everyday audits [349] highlight both the creativity and the burdens of grassroots governance. Similarly, participants expressed interest in labeling GenAI-made art when shared on social media, seeing such labels as a way to resist GenAI by filtering out AI-generated content. At the same time, implementing labeling raises difficult design challenges. Because participants’ perceptions of art depended on precisely where GenAI was used in the creative process, a single “AI-generated” label risks oversimplifying the diverse ways GenAI may be involved in art-making. Future work might therefore explore labeling schemes that move beyond a simple binary of GenAI use/non-use [109] as well as examine how different interface designs for displaying labels shape audience perceptions [153]. Complementary lines of research could continue investigating watermarking or detection methods for AI-generated media [418, 206], while also weighing the relative harms of false-negatives versus false-positives across application domains [90].

Muñoz’s queer theory cautions us not to naturalize these authentication and labeling practices as endpoints of design. Much like marginalized artists have critiqued GenAI moderation systems in prior work [253, 380], systems that institutionalize collective labeling or forced authentication risks shifting power toward technology developers and away from artists. Compulsory verification undermines artists’ ability to embrace opacity or refuse disclosure. Thus, future research

should explore not only technical methods for authentication, watermarking, and labeling [418, 206, 109], but also how to design with refusal in mind. While some communities may collectively value provenance, others may resist forced authentication as forms of surveillance. To follow Muñoz is to treat refusal itself as world-building: enabling artists to say no to GenAI and no to compulsory verification, as part of imagining futures otherwise. In this way, resistance to AI-generated media is not just about preventing harm in the present but about opening possibilities for queer worlds not yet here.

### 4.5.3 Supporting Queer Artists

For Muñoz, queer art entails a refusal of capitalist logics of productivity and profit-maximization, instead cultivating alternative temporalities and aesthetics that gesture toward utopian futures [263]—an orientation echoed by our participants. While they were highly critical of managers using GenAI to cut production costs (Section 4.4.2), participants expressed fewer concerns when artists used GenAI outside of profit-driven contexts (Section 4.4.4). Building on these insights, I suggest that designers can support queer futures by pursuing alternative paradigms that decenter efficiency and commodification.

One such paradigm to counter “straight” temporalities is slowness [276, 275, 274]. In response to participants’ critiques of GenAI being used to sacrifice quality for efficiency, models that emphasize slow output could open space for reflection and resist managerial appropriation. For example, a system that produces only one image per week could invite more deliberate engagements with GenAI’s artistic potential, while also reducing the environmental costs of large-scale AI development [99]. Such approaches align with post-growth HCI calls for technologies oriented around sustainability and care rather than acceleration [344, 343].

Researchers could also explore legal mechanisms to resist commodification. For example, “copy-left” licenses, developed in free and open-source communities, prevent enclosure by requiring derivative works to remain free [358]. Future computing research could collaborate with legal scholars to design similar licenses for GenAI models and datasets. Such GenAI models would allow artists to engage in minor support work, like practicing techniques, but may be unappealing to for-profit companies.

Finally, participants underscored the importance of imagining queer artistic futures beyond GenAI altogether. Echoing accessibility research that highlights relational support over individual productivity [29], I suggest shifting design attention toward infrastructures that sustain queer artistic mutual aid. In other words, researchers can design ways for queer artists to collectively support one another without having to rely on GenAI. For instance, systems could facilitate gifting and appreciation practices among queer artists, extending prior work on recognition in open-source communities [203]. P3, for example, described coordinating with others to create zines supporting queer and disabled causes, pointing toward broader design possibilities for technologies that scaffold collective art-making, as in Allred & Gray’s work on activist fanzines [13]. While GenAI is increasingly central in HCI research [282], our participants did not necessarily seek “better” models. Instead, their practices invite CSCW scholars to imagine queer artistic futures beyond GenAI by centering queer relationality.

#### 4.5.4 Supporting Artistic Labor

Participants drew a sharp line between GenAI as peripheral aid—such as editing pitch-letters—and GenAI as replacement for the core labor of art-making. This distinction exposes a broader politics of invisibility. As Becker notes, art worlds depend on “support work” that is usually erased, yet without which art cannot exist [25]. GenAI risks deepening this erasure, framing support tasks as disposable and as interchangeable with machines.

Muñoz reminds us that queer practices of labor are not simply about producing commodities but about resisting the “straight time” of efficiency and profit. In this sense, what is at stake is not just the preservation of jobs, but the defense of queer temporalities—slowness, care, collaboration—that capitalist logics render unproductive. The imposition of GenAI into creative labor markets exemplifies what Muñoz critiques as assimilationist compromise, demanding that artists adjust to extractive systems rather than reimagining them.

Rather than treating GenAI as an inevitable supplement to creative work, CSCW researchers must confront the structures that naturalize dispossession. This might mean aligning with unions and creative collectives not only to negotiate contract language but to contest the very framing of artistic labor as fungible, flexible, or optional. GenAI’s appropriation of queer art and labor is not simply a technological shift. It is a continuation of a long history of capitalist exploitation of marginalized creativity. Refusing this future requires imagining infrastructures that do not simply “accommodate” GenAI but actively block its coercive uptake, enabling artists to sustain labor practices grounded in solidarity, opacity, and utopian potential rather than extraction [146, 373].

#### 4.5.5 Limitations

My work does not intend to represent the views of all “queer artists.” Instead, my collaborators and I examine the views of a particular group of queer artists, which I interpret using José Esteban Muñoz’s theorization of queer aesthetics. Notably, all participants lived in the United States of America, and the interviews were conducted in English. Owing to the cultural situatedness of gender and sexuality [271] and artistic practices [251], more work is needed to understand the queer artists’ experiences beyond the US.

My findings also do not fully represent the diversity of queer artistic experiences *within* the US. As the interviews were conducted over Zoom, participants needed a reliable internet connection in order to participate. Also, Black queer artists have been pivotal to queer art, but are not represented in this study [263, 354]. This may be due to a limitation in having only recruited from Twitter and Reddit, as fewer Black people in the US use these sites than Instagram or TikTok [1]. Participants also skewed young, with a median age of 26 and the oldest participant being 33. As noted in prior Queer HCI scholarship, there is a need for future research on the experiences of older adults [420, 379].

While in this work I recruited self-identified queer artists, not every LGBTQ+ person would necessarily identify as queer [379]. As Muñoz notes, “queerness” is historically associated not only to marginalized genders and sexualities but also a form of radical politics and non-normative

aesthetics [263]. Therefore, my findings may have differed if I instead recruited LGBTQ+ identifying artists. Instead of representing LGBTQ+ artists writ large, these findings are likely more transferable to artistic communities that similarly value relationality and anti-capitalism, such as the transformative fan community to which at least 6 participants belonged. At the same time, an important limitation of this work is that, although participants felt strong solidarity with working artists, most did not make a living primarily through their art. As some participants engaged in artistic communities that stigmatize profiting from one's work [114], the participants in this study should not be disregarded as merely amateur artists. That said, the attitudes of, for instance, full-time screenwriters toward GenAI may differ from those of the participants in this study.

As transgender, non-binary and disabled people in the US tend to view AI more negatively than members of dominant social groups [156], participants' negative attitudes toward and non-use of GenAI are to be expected. At the same time, this non-use may have shaped participants' attitudes towards GenAI. This may be, for instance, why LGBTQ+ biases in model behaviors did not figure prominently in my findings. Moreover, participants' conceptions of GenAI tended to focus on prompt-based interaction paradigms, such as text-to-image models and chatbots. In light of my findings that participants were more comfortable with AI being used for support work, future research should study the acceptability of other types of GenAI models, such as those used for coloring. Additionally, Lucy Suchman notes that the category of technologies referred to as "AI" is quite expansive [371]. As a result, artists likely have different conceptions of what constitutes AI and its acceptability. Similar to prior research on people's mental models of the internet [197], future research should explore artists' mental models of AI and AI development. Finally, while this work largely focuses on queer artists refusing GenAI, I encourage future research into queer artists who have chosen to integrate GenAI in their artistic practices [380].

## 4.6 Conclusion

In this work, I examined how queer artists are making sense of and responding to the recent proliferation of GenAI. The tensions participants identified between themselves and GenAI are not simply the most recent version of longstanding aesthetic debates [25] or the result of technological *naïvetés*. Instead, these are tensions between competing future visions of what art is and should be, between art-making as a relational queer practice versus an alienated commodity from which to extract profit. I encourage CSCW researchers to contest these dominant AI imaginaries. As José Esteban Muñoz reminds us, queerness demonstrates the "concrete possibility for another world" [263].



## Chapter 5

# How Generative Image Model Developers Evaluate the “Aesthetic Quality” of Images

This chapter is based upon the following published work:

Jordan Taylor, William Agnew, Maarten Sap, Sarah Fox,\* Haiyi Zhu.\* 2026. The Algorithmic Gaze of Image Quality Assessment: An Audit and Trace Ethnography of the LAION-Aesthetics Predictor. Proceedings of the 2026 ACM Conference on Fairness, Accountability, and Transparency. (\* = equal contribution)

### 5.1 Introduction

Although artistic taste is deeply subjective and highly political [45], AI researchers often attempt to quantify the aesthetic quality of media [384, 183]. While aesthetic quality assessment (AQA) has been a topic of research in the computer vision community for decades [92, 266], this task is particularly important to examine in light of the rapid growth in visual generative AI. Not only are AQA models used to curate the training data for visual generative AI models [69], but they are also used to assess the quality of AI-generated images [130]. As generative AI models are trained on datasets scraped from across the web, researchers rely on computational methods to filter out “low-quality” data. Prior work has called attention to myriad gender and cultural biases in common dataset filtering techniques—such as blocklists [98] and image-text similarity metrics [179]—but the consequences of aesthetic filtering has yet to be explored in detail.

In this work, I investigate an AQA model—the LAION-Aesthetics Predictor (LAP)—that has been widely influential in visual generative AI research and commercial development. For example, LAP was used to curate the open-source LAION-Aesthetics Dataset from which Stability AI pre-trained the earliest versions of Stable Diffusion [332]. Additionally, LAP is often used to *evaluate* the quality of AI-generated images [95, 119, 130, 168]. Despite its far-reaching impact, the LAP itself has yet to be explored in detail.

Due to the subjectivity of aesthetic evaluation, I first explored (RQ1) what the LAP model determines to be a “high-quality” image using three datasets. First, I examined the impact of aesthetic filtering on the LAION-Aesthetics Dataset (~1.2B images), which was curated using LAP. I find that the LAP disproportionally filters in images with captions mentioning women, while filtering out images with captions mentioning men or LGBTQ+ people. Then, I used LAP to score ~330k images across two art datasets, finding the model rates realistic images of landscapes, cityscapes, and portraits from western and Japanese artists most highly. In other words, the **algorithmic gaze** of LAP reproduces the male gaze within western art history [31]. To understand where these biases may have originated, I performed a trace ethnography of public materials to examine (RQ2) how LAP was developed. I find that LAP’s development reflects the biases from

the aforementioned audits, such as the aesthetic scores used to train LAP primarily coming from English-speaking photographers and western AI-enthusiasts.

In light of these findings, I call on AI researchers to contend with aesthetics as a site of cultural representation. Simultaneously, I caution that greater “inclusion” in visual generative AI training data risks exacerbating the harms of these technologies. Additionally, I call on developers to shift away from prescriptive, universalist measures of “aesthetics” toward more descriptive, pluralistic evaluations. Finally, I discuss how future FAccT researchers might also combine audits and ethnographies of algorithmic systems.

## 5.2 Background & Related Work

### 5.2.1 Background on LAION-Aesthetics

LAION (Large-Scale Artificial Intelligence Open Network) is a non-profit founded in 2021 by Christoph Schuhmann, a German high school teacher, with the goal of creating open-source datasets and machine learning models [268]. Despite its humble origins, LAION has highly influenced generative image model development. Stability AI researchers used the LAION-Aesthetics Dataset to pre-train the earliest versions of Stable Diffusion [316]. Unfortunately, the extent to which *contemporary* generative image models rely on the LAION datasets can be challenging to assess because developers have become cautious of disclosing their training data. While Stability AI shared that early versions of Stable Diffusion were trained using LAION datasets, the model card for their 2024 model Stable Diffusion 3.5-Large vaguely says it was “*trained on a wide variety of data, including synthetic data and filtered publicly available data.*”<sup>1</sup> This secrecy may be due to the increased scrutiny faced by LAION and generative AI developers [304, 307], which I will describe below.

LAION has faced substantial criticism from both AI researchers and artists. LAION researchers compiled the 5.86 billion image-text pairs of the LAION-5B dataset by extracting image links and alternative text descriptions from the IMG HTML tags of websites archived by the Common Crawl, a US-based non-profit that has scraped publicly available websites since 2008 [305, 19]. Due to LAION researchers’ indiscriminate data collection practices, researchers found that LAION-5B contained over a thousand images of child sexual abuse material (CSAM) [382]. In addition to containing CSAM, LAION-5B contains millions of copyrighted images, leading to numerous lawsuits against LAION [118] and Stability AI [386]. Furthermore, Birhane et al. found that the image descriptions in LAION-400M and the English language subset of LAION-5B (LAION-2B-en) often contain hate speech [38].

Although prior research has examined biases in LAION datasets [38, 179], the LAION-Aesthetics Dataset and the LAION-Aesthetics Predictor (LAP) model<sup>2</sup> used to curate this dataset are underexplored. Not only is the LAION-Aesthetics Dataset widely used to pre-train and fine-tune visual generative AI models, but the LAP model used to curate this dataset has also been used

<sup>1</sup><https://huggingface.co/stabilityai/stable-diffusion-3.5-large>

<sup>2</sup><https://github.com/christophschuhmann/improved-aesthetic-predictor>

to curate internal corporate datasets [69] and to evaluate the performance of visual generative AI models [95, 119, 130, 168]. Although some humanistic research has critiqued how LAP was trained [186, 394], researchers have not yet empirically investigated what this image quality assessment model measures in practice. In doing so, I contribute both to scholarship on art and AI as well as, more broadly, AI evaluation and alignment practices. In the following sections, I briefly describe each of these contexts in greater detail.

### 5.2.2 AI Evaluation & Alignment

AI researchers are increasingly critiquing the measurements used to train and evaluate AI models. For example, benchmarks often claim to measure concepts with no agreed upon definition, such as “intelligence” [24]. Researchers have also criticized the notion of universal “gold standards” in data annotation [398]. While researchers often treat annotator disagreement as noise to be filtered out, disagreements can stem from annotators’ differing lived experiences [398, 325]. In response, researchers have called for distinguishing between datasets intended to represent different beliefs or encode one specific belief [317]. AI alignment research focuses on normative evaluations where people can reasonably disagree about how a model should behave. For instance, LLMs tend to reflect social values more commonly held in western countries according to national value surveys [375]. At the same time, this methodology can flatten intra-cultural differences [12]. Instead of choosing one set of values to reflect, researchers have increasingly called for designing AI systems more pluralistically [362, 361]. In this work, I contend with “aesthetics” as an underexamined axis of alignment.

Prior work has found that the design of visual generative AI models often harms marginalized communities. Text-to-image (T2I) models stereotype or poorly represent women [33], LGBTQ+ people [390, 380], people of color [410], and non-western cultures [127, 251, 296, 415]. T2I models made by western tech companies tend to perform worse in languages other than English, such as Farsi [296] or Bangla [251]. This linguistic bias can also be seen in a dataset I examine in this study: the LAION-Aesthetics Dataset is restricted to images with English captions [332]. HCI researchers have found that artists often enjoy the glitches and strangeness of T2I models [65], leading some to critique the stylistic bias toward realism in newer models [380]. In light of this concern, Taylor et al. have called for future research “examining measurements of aesthetic quality in popular datasets” [380], which I take up in this work.

While the artistic consequences of aesthetic evaluation are underexamined, aesthetic quality assessment (AQA) has been a subject of research in the computer vision community for decades [266, 92]. In this work, researchers formulate AQA as a regression problem predicting a numerical aesthetic score. At the same time, this de-contextualized measurement of aesthetic quality has been critiqued [145, 183]. Goree, Domingo, and Crandall [145] highlight value misalignments between various AQA models and human-preferences by developing an aesthetic evaluation mobile camera application. Like ourselves, Li et al. [219] recently advocated for investigating aesthetic biases as an underexamined form of representational harms in mulit-modal LLMs. However, the authors conceptualize aesthetic biases as a distributional alignment problem wherein aesthetic evaluation models may privilege the values of one demographic group over others, examining age, gender and education. I extend this work by studying the broader

cultural consequences of AQA on data curation and model evaluation, grounding the task within art history.

## 5.3 Auditing the LAION-Aesthetics Predictor Model

### 5.3.1 Audit Methodology

I audited LAP using three datasets. First, I examine the LAION-Aesthetics Dataset curated from LAION-5B by LAP. I then audited LAP using two art datasets with robust metadata on the artists, cultures, styles, and mediums associated with each image. The first dataset consists of 249,351 public-domain images from the Metropolitan Museum of Art (MET).<sup>3</sup> Located in New York City, the MET is the fourth largest museum in the world and its collection spans over 5,000 years [376]. While the MET dataset contains images from a wide breadth of art across various cultures and time periods, few images are of modern or contemporary art (only 865 images). Therefore, I evaluated LAP on the WikiArt dataset [322] of 81,444 images of visual art by 129 primarily modern (mid-1800s to mid-1900s) artists. WikiArt is a community edited visual art encyclopedia commonly studied in digital humanities research [254, 128].

Across each datasets, my analyses primarily focus on understanding differences between images rated above and below an aesthetics score of 6.5. I chose this threshold because it is commonly used in computer vision research to represent high-quality images. The LAION-Aesthetics Dataset was released in multiple subsets based on different aesthetic thresholds, with images scored 4.5+ ( $\sim 1.2$ B images) being the lowest rated dataset and 6.5+ ( $\sim 625$ k images) being the highest rated. The LAD 6.5+ is widely used by computer vision researchers to fine-tune generative image models [149, 62], evaluate new training methods [227, 110], and create new datasets [388, 220]. Additionally, Ding et al. measure the effectiveness of data poisoning attacks on text-to-image models by defining “generation aesthetics” as the “% of images with [LAP] aesthetics  $\geq 6.5$ ” [95]. In Tables 5.1 and 5.2 I also show the number of images rated 6+ to demonstrate that the aesthetic biases I identify across the 6.5+ threshold can also be seen across the 6+ threshold.

### 5.3.2 Findings: The LAION-Aesthetics Dataset

LAD consists of all images from LAION-5B with English captions that LAP rated at least 4.5. In LAD, images are not stored directly. Rather, the dataset contains URLs to images hosted on websites across the internet. Most relevant for the findings in this work, the metadata for each image in LAD contains its LAP predicted aesthetic score and a caption derived from the scraped image’s alternative text. I find that LAP rates images from independent visual artists and photographers highly. Also, images with captions mentioning women, Hindus, and Christians are more likely to be included in LAP 6.5+, while images with captions mentioning men, Jews, or Muslims are more likely to be filtered out.

<sup>3</sup><https://www.metmuseum.org/hubs/open-access>

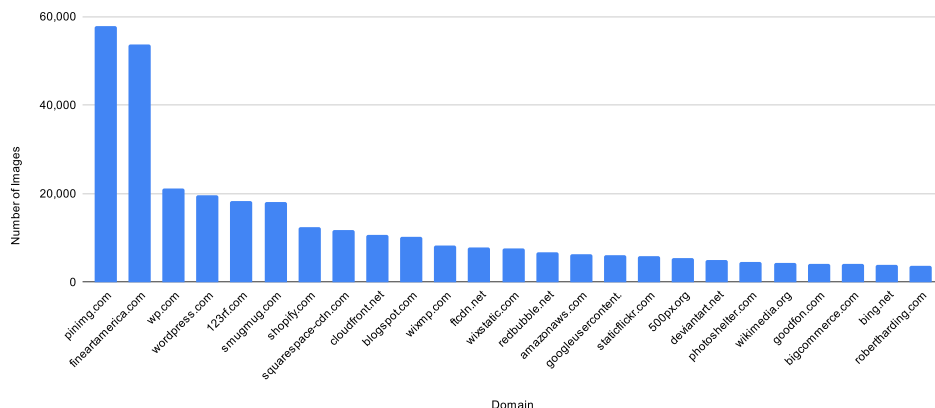


Figure 5.1: Top 25 domains of images in the LAION-Aesthetics Dataset rated 6.5+ by the LAION-Aesthetics Prediction model.

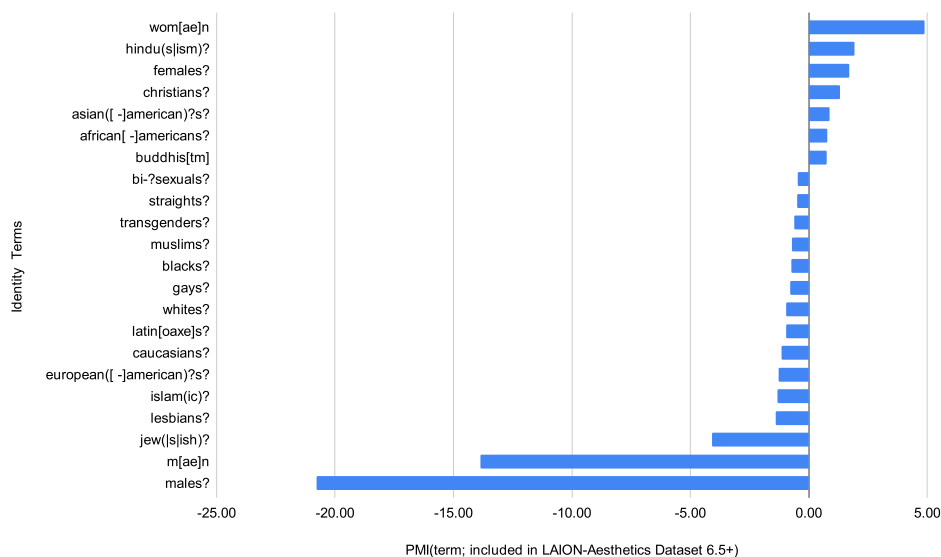


Figure 5.2: Pointwise Mutual Information (PMI) between regex and images being included in the 6.5+ subset of the LAION-Aesthetics Dataset. Terms with higher PMI (e.g., wom[ae]n) have higher likelihood of being included. The exact regex for each identity term was of the form `\b(?i)gays\b`.

### 5.3.2.1 The LAION-Aesthetics Predictor highly rates images from websites used by independent visual artists and photographers

I find substantial differences between the URLs of images below and above the 6.5 aesthetic threshold. In LAD 4.5+, the top 25 most common domains account for about 36% of all the images in the dataset. These domains are primarily from e-commerce websites (e.g., Shopify), stock image sites (e.g., Getty Images), and web hosts (e.g., Wordpress). For LAD 6.5+, the top 25 image domains account for proximately 52% of the dataset, meaning a majority of the images in this dataset come from the handful of websites in Figure 5.1. Of these 25 domains, 13 were also in the top 25 domains for LAD 4.5+. However, the other domains are primarily from smaller websites where independent visual artists (e.g., Redbubble, DeviantArt) and photographers (e.g., SmugMug, Flickr, 500px) share or sell their work. Additionally, the 25th most common source of images in LAD 6.5+ are from a small nature photography agency, robertharding.com. This indicates that LAP may rate photographic images more highly.

| Department Name                                 | Median Score | Num Images Scored $\geq 6.5$ | Num Images Scored $\geq 6$ | Num Images    |
|-------------------------------------------------|--------------|------------------------------|----------------------------|---------------|
| <b>European Paintings</b>                       | <b>5.73</b>  | <b>35</b>                    | <b>483</b>                 | <b>2,289</b>  |
| <b>Photographs</b>                              | <b>5.40</b>  | <b>15</b>                    | <b>548</b>                 | <b>6,685</b>  |
| <b>Drawings &amp; Prints</b>                    | <b>5.06</b>  | <b>13</b>                    | <b>435</b>                 | <b>60,889</b> |
| Robert Lehman Collection                        | 5.04         | 3                            | 41                         | 2,273         |
| The Cloisters                                   | 4.96         | 0                            | 1                          | 2,258         |
| Costume Institute                               | 4.94         | 0                            | 7                          | 8,124         |
| <b>Asian Art</b>                                | <b>4.93</b>  | <b>60</b>                    | <b>1,240</b>               | <b>30,597</b> |
| <b>The American Wing</b>                        | <b>4.86</b>  | <b>49</b>                    | <b>474</b>                 | <b>12,292</b> |
| Arms & Armor                                    | 4.82         | 0                            | 2                          | 7,260         |
| The Libraries                                   | 4.77         | 0                            | 0                          | 107           |
| Musical Instruments                             | 4.75         | 0                            | 3                          | 2,498         |
| Modern & Contemporary Art                       | 4.73         | 1                            | 11                         | 865           |
| European Sculpture & Decorative Arts            | 4.72         | 1                            | 12                         | 37,709        |
| Arts of Africa, the Ancient Americas, & Oceania | 4.71         | 0                            | 0                          | 6,191         |
| Medieval Art                                    | 4.70         | 0                            | 1                          | 6,887         |
| Egyptian Art                                    | 4.65         | 0                            | 3                          | 13,269        |
| Islamic Art                                     | 4.56         | 0                            | 15                         | 13,054        |
| Ancient West Asian Art                          | 4.41         | 0                            | 1                          | 6,093         |
| Greek & Roman Art                               | 4.29         | 0                            | 7                          | 30,011        |

Table 5.1: Predicted aesthetic score of images from the Metropolitan Museum of Art by department.

### 5.3.2.2 The LAION-Aesthetics Predictor is more likely to rate images with descriptions mentioning woman $\geq 6.5$ , while those mentioning men or LGBTQ+ people are more likely to be filtered out.

To understand the content of images in LAD, I compared the frequencies of words in image captions above and below the 6.5 aesthetics score threshold. Similar to Dodge et al.’s prior audits of generative AI dataset filtering [98], I computed the pointwise mutual information (PMI) [75] of a regular expression (regex) appearing in the 6.5+ subset of images in LAD. At a high level, PMI in this case measures the probability of observing a search term in the description for an image rated 6.5+ relative to chance. Therefore, a PMI much greater than 0 suggests the search term appears more often than by chance and a PMI much less than 0 suggests the regex occurs less often than expected. Building on identity-related regexes used in prior audits [98, 179], I computed the PMI of the regexes in Figure 5.2 appearing in LAD 6.5+ image descriptions.

I find that images with descriptions mentioning women are more likely to be included in LAD 6.5+, while images mentioning men or LGBTQ+ communities are more likely to be excluded. Also, images with descriptions mentioning Hindu, Buddhist, or Christian communities are more likely to be included in LAD 6.5+, while images with descriptions mentioning Jews or Muslims are more likely to be filtered out. Images mentioning Asian or African communities are more likely to be included in LAD 6.5+ than those mentioning Latinx/e, Caucasian, or European communities. This may suggest that LAP more highly rates images of Asian and African communities than European communities. Alternatively, images containing Caucasians or Europeans may not mention this detail in their captions because whiteness is an unmarked racial category [47].

As PMI is a relative measure, one must also consider the *absolute* number of image descriptions per regex. For example, the PMI of ‘whites?’ and ‘latin[oaxe]s?’ are both -0.95. However, there are 9,617 images in LAD 6.5+ with ‘whites?’ in their description but only 26 images for their regex ‘latin[oaxe]s?’. Not only does ‘wom[ae]n’ have the highest PMI but the regex also appears in more image descriptions in LAD 6.5+ than any other regex in Figure 5.2 (15,706). Meanwhile, regexes associated with LGBTQ+ identities are far rarer in LAD 6.5+ image descriptions: ‘gays?’ (102), ‘lesbians?’ (20), ‘transgenders?’ (9), ‘bi-?sexuals?’ (2), and ‘non[-]?binary’ (0). Understanding the cultural implications of aesthetic evaluation requires considering more than the representation of identity groups [380]. Below, I explore aesthetic filtering in greater detail by examining how LAP rates art from various cultures.

### 5.3.3 Findings: The Metropolitan Museum of Art Dataset

To examine how the LAP evaluates art across different mediums and cultures, I scored the 249,351 images from the MET dataset using LAP. I find LAP rates two-dimensional art by western or Japanese artists most highly.

|               |                        | Median Score | Num Images Scored $\geq 6.5$ | Num Images Scored $\geq 6$ | Num Images |
|---------------|------------------------|--------------|------------------------------|----------------------------|------------|
| <b>Genre</b>  | Cityscape              | 5.97         | 539                          | 2,151                      | 4,602      |
|               | Portrait               | 5.90         | 693                          | 5,616                      | 14,112     |
|               | Landscape              | 5.87         | 505                          | 4,637                      | 13,358     |
|               | ...                    | ...          | ...                          | ...                        | ...        |
|               | Illustration           | 5.36         | 11                           | 140                        | 1,902      |
|               | Nude Painting          | 5.45         | 3                            | 117                        | 1,923      |
|               | Abstract Painting      | 4.94         | 0                            | 0                          | 4,968      |
| <b>Style</b>  | Realism                | 5.87         | 634                          | 4,155                      | 10,733     |
|               | Romanticism            | 5.87         | 371                          | 2,723                      | 7,019      |
|               | Impressionism          | 5.89         | 646                          | 4,898                      | 13,060     |
|               | ...                    | ...          | ...                          | ...                        | ...        |
|               | Cubism                 | 5.50         | 12                           | 210                        | 2,235      |
|               | Pop Art                | 5.05         | 2                            | 33                         | 1,483      |
|               | Abstract Expressionism | 4.96         | 0                            | 7                          | 2782       |
| <b>Artist</b> | Antoine Blanchard      | 6.94         | 168                          | 170                        | 170        |
|               | Edouard Cortes         | 7.09         | 208                          | 214                        | 214        |
|               | ...                    | ...          | ...                          | ...                        | ...        |
|               | Pablo Picasso          | 5.51         | 1                            | 65                         | 762        |
|               | Andy Warhol            | 5.22         | 0                            | 6                          | 107        |
|               | Salvador Dalí          | 5.17         | 0                            | 20                         | 479        |

Table 5.2: A representative selection of the highest and lowest scored genres, styles, and artists in WikiArt. The full results per genre, style, and artist can be downloaded at [https://github.com/jtaylor351/laion\\_aesthetics\\_audit\\_data/tree/main](https://github.com/jtaylor351/laion_aesthetics_audit_data/tree/main)

### 5.3.3.1 The LAION-Aesthetics Predictor rates western art more highly than art from the Majority World

As seen in Table 5.1, the departments with images scored  $\geq 6.5$  are primarily European Paintings, Photographs and Drawing & Prints as well as art from Europe, Asia, and America. On the other hand, not a single piece of African, Native American, Oceanian, Egyptian, Islamic, Ancient West Asian, and Greek & Roman Art was rated 6.5+. This bias is even more apparent when looking at the images rated 6+. **Approximately 97% of all images scored 6+ from the MET (3180 of 3284) come from only five departments: (1) Asian Art, (2) Photographs, (3) European Paintings, (4) The American Wing, and (5) Drawing and Prints.** This may be partially because the LAP privileges figurative, two-dimensional art, which would also explain why more abstract art (Modern & Contemporary Art) and physical artifacts (European Sculpture & Decorative Arts) scored lower.

### 5.3.3.2 The LAION-Aesthetics Predictor rates paintings and photographs by western or Japanese artists most highly.

Across the 249,351 images corresponding to 37,878 unique mediums and 14,946 known artists in the MET dataset, only 177 images across 35 mediums and 99 known artists were rated as  $\geq 6.5$  by LAP. After manually combining similar mediums (e.g., ‘watercolor’ and ‘watercolor on paper’), I found that the MET images rated  $\geq 6.5$  come from only 5 mediums: 72 oil paintings, 60 Japanese woodblock prints, 26 watercolor paintings, 15 photographs, 2 drawings, 1 piece of stained glass, and 1 tempera painting (a predecessor to oil paint). A total of 14 artists had more than one piece rated  $\geq 6.5$ , accounting for 88 (50%) of the 177 top rated images.

A manual inspection indicates that images rated  $\geq 6.5$  are typically two-dimensional realistic representations of landscapes (e.g., paintings of mountains) or portraits (e.g., photographs or paintings of people). This can also be seen in the artists with multiple images in the top 177. The top two artists with the most images rated  $\geq 6.5$ , Katsushika Hokusai (32) and Utagawa Hiroshige (23), are mid-19th century ukiyo-e woodblock print artists who focus on Japanese landscapes [196]. Likewise, the artists with the third and sixth most images rated  $\geq 6.5$  are the landscape painters William Trost Richards (6) and Albert Bierstadt (3). The artists with the fourth (Southworth and Hawes, 4) and fifth (Suzuki Shin’ichi, 4) most images in the top 177 are portrait photographers. While the MET dataset contains a broad range of art across cultures and mediums, those that LAP rated  $\geq 6.5$  were mostly paintings, prints, or photographs by western or Japanese artists. In the next section, I examine how LAP evaluates paintings in greater detail through the WikiArt dataset.

### 5.3.4 Findings: The WikiArt Dataset

The WikiArt dataset consists of 81,444 pieces of visual art from 128 known artists (excluding unknown artists), primarily western artists from the mid-1800s to mid-1900s. Each piece is assigned one of 27 styles across 10 genres. Genres describe the content of a piece (e.g., landscape or portrait) meanwhile style describes the piece’s form (e.g., realism or impressionism).

#### 5.3.4.1 The LAION-Aesthetics Predictor more highly rates realistic images of cityscapes, landscapes, and portraits, while less realistic art was rated lower.

As seen in Table 5.2, I find that LAP highly rates visual art of Cityscapes, Landscapes, and Portraits. These three genres account for 39% of all images in WikiArt, but these genres make up 73% of all images rated 6.5+. Likewise, I found that more realistic and representational styles, such as Realism, Romanticism, and Impressionism were highly rated. Here, both style and genre are linked as portraits, landscapes, and cityscape are common subjects in realist, romantic, and impressionistic art. On the other hand, LAP rates less representational art, such as cubism or abstract art, lower. The consequences of this aesthetic bias can be seen across artists. Nearly every piece by the french cityscape artists Edouard Cortès (208 of 214) and Antoine Blanchard (168 of 170) were rated 6.5+. In contrast, more influential but less realistic artists—Pablo Picasso, Andy Warhol, and Salvador Dalí—were rated lower.

| Aesthetics Score Range        | LAION-Aesthetics Dataset                                                          | Met Dataset                                                                        | WikiArt Dataset                                                                     |
|-------------------------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| $6.5 \leq \text{score}$       |  |  |  |
| $5.5 \leq \text{score} < 6.5$ |  |   |  |
| $4.5 \leq \text{score} < 5.5$ |  |   |  |
| $\text{score} < 4.5$          | Not Applicable                                                                    |   |  |

Table 5.3: A selection of images across the three datasets showing that realistic images of landscapes and cities are most highly rated.

### 5.3.5 Findings: Qualitative Visual Analysis

To augment my quantitative analysis, I randomly sampled 5 images from each dataset in the ranges of:  $[7,)$ ,  $[6.5, 7.0)$  ...  $[4.5, 5.0)$ ,  $(, 4.5)$ . Given no images in LAD are rated  $< 4.5$  and two images in the MET were rated  $\geq 7$ , this yielded 30 images from LAD, 32 images from the MET, and 35 images from WikiArt. A representative image from each dataset across different score ranges can be seen in Table 5.3. **My inspection of these images further suggests that LAP tends to rates realistic images of cityscapes, landscapes, and people higher than images of three-dimensional objects or abstract art.** Among the images rated 6.5+ across the datasets, all are of landscapes/nature (11/27), cityscapes (9/27), or people (7/27). Between 5.5 and 6.5, most images are of people (20/30) or landscapes/nature (5/30). The images scored below 5.5 are primarily of inanimate objects (20/40), people (10/40), or abstract art (7/40).

## 5.4 How the LAION-Aesthetics Predictor was Made

Seaver [334] suggests that audits can demonstrate the *existence* of biases, but “what [audits] cannot do is explain conclusively how that disparate impact came about.” In this section, I explore the potential origins of the aesthetic biases described above by conducting a trace ethnography of how the LAP was made.

### 5.4.1 Trace Ethnographic Methodology

Ethnography is a qualitative method from anthropology seeking to explain how groups of people interpret and act in the world [102]. Ethnographic research typically involves extended participant observation and interviews within physical field-sites, ranging from remote villages [132] to AI research labs [120]. While ethnography began within physical field sites, the 1990s saw a proliferation of multi-sited ethnographic research due to the rise of globalization and the internet [237]. Within this paradigm, Geiger and Ribes [133] propose trace ethnography as “a powerful and flexible methodology, able to turn thin documentary traces into ‘thick descriptions’ [132] of actors and events that are often invisible in today’s distributed, networked environments.” For example, Geiger and Ribes [133] reconstruct the blocking of a Wikipedia vandal by drawing on thin documentary traces of edit history.

I conducted a trace ethnography of the development of the LAION-Aesthetics Predictor by leveraging a variety of digital documents. Specifically, I draw on: (1) the LAION-Aesthetics announcement blog post [332], (2) the code and data repositories hosting the LAP model,<sup>4</sup> SAC dataset,<sup>5</sup> AVA dataset,<sup>6</sup> and LAION-Logos dataset,<sup>7</sup> (3) the academic paper for the AVA dataset [266], and (4) a ~10 minute YouTube video explaining how LAP was made.<sup>8</sup> My analysis (summarized in Table 5.4) focused on understanding how LAP and its training data were created and where the images and annotations in LAP’s training data come from.

### 5.4.2 Findings: How the LAION-Aesthetics Predictor was Created

#### 5.4.2.1 Model development choices were made according to the individual taste of LAION’s founder

According to the aforementioned YouTube video, the founder of LAION (Christoph Schuhmann) created LAP on his own. LAP is a multi-layer perceptron model with linear layers trained to predict the aesthetic score of an image based on a CLIP embedding.<sup>9</sup> Although he experimented with other models, Schuhmann chose this architecture because “*for [his] use case the simple linear layer worked best just from visual inspection.*” Schuhmann trained LAP using three distinct datasets, each of which measures the aesthetic quality of an image on a scale from 1 to 10: (1) Aesthetic Visual Analysis (AVA), (2) Simulacra Aesthetic Captions (SAC), and (3) LAION-Logos. Once again, Schuhmann explains that he chose to weight each dataset equally in the training process because “*from [his] subjective taste it was best to just have every dataset once in the training data.*” In other words, Schuhmann decided how to train LAP based largely on his individual taste.

Schuhmann seems to recognize various limitations of LAP. He encourages others to “*experi-*

<sup>4</sup><https://github.com/christophschuhmann/improved-aesthetic-predictor>

<sup>5</sup><https://github.com/JD-P/simulacra-aesthetic-captions>

<sup>6</sup>[https://github.com/imfing/ava\\_downloader](https://github.com/imfing/ava_downloader)

<sup>7</sup><https://huggingface.co/datasets/ChristophSchuhmann/aesthetic-logo-ratings>

<sup>8</sup><https://www.youtube.com/watch?v=OZ3krwyoeWw>

<sup>9</sup>CLIP models are commonly used in computer vision to semantically represent both texts and images in the same vector space [299]

ment” with different model architectures, newer CLIP embedding models, and larger aesthetic evaluation datasets. Additionally, Schuhmann asks for volunteers to reproduce the training data curation steps described in the video because “[he] had some ugly code somewhere but [he] can’t find it anymore.” He also asks for volunteers in the last comment of the explainer video: “This is all an ongoing area of research and it would be really cool to have someone working on this.” Schuhmann appears to recognize the slapdash way LAP was made and expects that it will be improved upon in the future. However, at the time of my writing in January 2026, LAION has not released a new aesthetic evaluation model.

|                                 | <b>AVA</b>                        | <b>SAC</b>             | <b>LAION-Logos</b>     |
|---------------------------------|-----------------------------------|------------------------|------------------------|
| <b>Year Released</b>            | 2012                              | 2022                   | 2022                   |
| <b>Number of Images</b>         | 255,530                           | 146,372                | 26,730                 |
| <b>Mean Ratings per Image</b>   | 210.1                             | 1.2                    | 1.48                   |
| <b>Median Ratings per Image</b> | 201                               | 1                      | 1                      |
| <b>Images data consent</b>      | No                                | Yes                    | No                     |
| <b>Ratings data consent</b>     | No                                | Yes                    | Yes                    |
| <b>Dataset Documentation</b>    | Academic paper                    | GitHub README.md       | No documentation       |
| <b>Where images come from</b>   | www.dpchallenge.com               | T2I generated          | LAION-5B               |
| <b>Where scores come from</b>   | www.dpchallenge.com               | 294 AI enthusiasts     | 18 AI enthusiasts      |
| <b>What scores measure</b>      | <b>Relative</b> rating on a theme | <b>Absolute</b> rating | <b>Absolute</b> rating |

Table 5.4: Comparison of Datasets Used to Train the LAION-Aesthetics Predictor

#### 5.4.2.2 Training datasets are inconsistently documented

AVA is an influential predecessor to the LAION-Aesthetics dataset within the computer vision community [302, 399, 204, 226]. The computer vision researchers Murray et al. collected AVA by scraping images and ratings from an online photography competition website (www.dpchallenge.com) in 2012 [266]. SAC was released in July of 2022 by John David Pressman, a Stability AI researcher at the time [292]. Pressman collected this dataset using a Discord bot shared in AI-art communities that allowed volunteers to generate images and vote on their perceived aesthetic quality. Although I was able to find the content of each dataset, the quality of their documentation varies widely. AVA was the most well documented because it was published at a computer vision conference [266]. While SAC was not formally published in an academic setting, its GitHub repository has an extensive README.md file that functions as a data card [295]. As I will describe next, LAION-Logos had little documentation.

I initially struggled to find LAION-Logos. While AVA and SAC are linked in the LAION-

Aesthetics blog post, LAION-Logos is not. Searching “LAION-Logos” on Google yielded only websites and academic articles alluding to the original blog. In response, I searched the public LAION Discord, eventually finding an unlisted YouTube video posted by Schuhmann in 2022.<sup>10</sup> Because this video is unlisted, it is not searchable on YouTube or indexed by search engines. In the video, Schuhmann explains that LAION-Logos is located on Hugging Face.<sup>11</sup> LAION-Logos is sparsely documented with a one sentence explanation and rarely used, downloaded only 40 times in the month prior at the time of my writing. While LAION-Logos is technically publicly available, my challenges finding the dataset suggests that it has been *de facto* private. To the best of my knowledge, I am the first to document LAION-Logos.

#### 5.4.2.3 Training datasets were created for different purposes across different time periods

LAION-Logos and SAC were created to improve generative image models. LAION-Logos was collected to “*improve the [LAION aesthetic prediction] models abilities to evaluate images with more or less aesthetic texts in them*” [332] While the SAC documentation says the dataset can be used in multiple ways, like “Filtering Datasets” and “Alignment Research,” its documentation indicates that SAC was primarily collected for LAION-Aesthetics. This is evident in the acknowledgment: “*Special thanks to [name] for his early feedback and suggestions on the bot, as well as writing the initial public domain prompt and rating dataset that allowed us to make LAION-aesthetic as early as we did*”

In contrast, AVA was collected a decade prior to the popularization of generative AI research. To motivate the construction of AVA, Murray, Marchesotti, and Perronnin [266] suggest that “*With the ever-expanding volume of visual content available, the ability to organize and navigate such content by aesthetic preference is becoming increasingly important.*” The authors position AVA as a solution to the problem of information overload, a common refrain in recommendation research [336]. Perhaps because AVA was created within a personalization paradigm, AVA embodies a more pluralistic conception of aesthetics than the universalism of LAION-Logos and SAC. I describe these competing conceptions of aesthetics below in greater detail.

### 5.4.3 Limitations of the Images and Aesthetic Annotations Used to Train the LAION-Aesthetics Predictor

#### 5.4.3.1 Images and aesthetic annotations were collected with mixed levels of consent

These datasets were collected with varying levels of consent (Table 5.4). For AVA, neither the images nor the ratings were collected with the consent of dpchallenge.com members. With LAION-Logos, annotators likely knew their ratings would be used to create an aesthetic evaluation dataset, but image creators did not consent to this use. For SAC, the AI enthusiasts generating and annotating images consented to their images and annotations being used to create an aesthetic evaluation dataset.

<sup>10</sup><https://www.youtube.com/watch?v=OZ3krwyoeWw>

<sup>11</sup><https://huggingface.co/datasets/ChristophSchuhmann/aesthetic-logo-ratings>

#### 5.4.3.2 The images used to train LAION-Aesthetics Predictor are primarily photographs.

Of the 428,632 images used to train LAP, 255,530 (~60%) come from AVA, 146,372 (~34%) from SAC, and 26,730 (~6%) from LAION-Logos. Both the images and ratings from AVA are from photographers who participated in the English-language online art competition website dpchallenge.com between its founding in 2002 and its scraping in 2012. The images in SAC were generated and rated by a handful of AI enthusiasts in 2022 using an early version of Stable Diffusion, with the top 50 most active volunteers accounting for 87.4% of all ratings. Finally, the images in LAION-Logos from LAION-5B were scraped from the internet no later than early 2022 and rated by only 18 LAION volunteers. The images used to train LAP have a number of limitations. Temporally, most images are from before 2012, meaning newer content may be underrepresented. Also, the over-representation of photographs in LAP’s training data may explain why the model highly rates photorealism.

#### 5.4.3.3 Annotators are primarily English-speaking photographers or western AI-enthusiasts.

When the AVA data was created (2002 - 2012), the World Bank estimate that less than 33% of the world’s population used the internet, most of whom were in western countries.<sup>12</sup> Therefore, the images and ratings posted on the English-language website dpchallenge.com at the time likely reflect the tastes of western photographers. The SAC README cautions: “*Participants in [SAC] are largely WEIRD, or Western, Educated, Industrialized, Rich, and Democratic. This means that the aesthetic preferences recorded are not universal among humanity. While we didn’t take a survey of the demographic makeup of SimulacraBot users, it can be assumed they largely reside in the United States and Europe.*” The README also warns that participants were “*mostly open source AI developers and enthusiasts,*” so “*their aesthetic feedback is going to lean STEM, fantasy, nerdy, esoteric, etc.*” While the identity of LAION-Logos annotators is unclear, I expect they share similar biases to SAC. Moreover, a majority of the ratings in SAC and LAION-Logos were chosen by a single person (Table 5.4).

#### 5.4.3.4 The LAION-Aesthetics Predictor was trained by conflating different measurements of aesthetics across the datasets.

Per the LAION-Aesthetics blog, LAP was trained to “predict the rating people gave when they were asked ‘How much do you like this image on a scale from 1 to 10?’” [332]. While it is true that SAC and LAION-Logos measures aesthetics as an *absolute* image property, this is not what AVA measures. Votes on dpchallenge do not measure the perceived aesthetic quality of a de-contextualized image but rather the quality of an image *relative* to a topic, like skyscapes. As prizes are award to images based on the average final *ranking* of aesthetic votes for a challenge, votes are comparable within a topic but not between topics. Murray, Marchesotti, and Perronnin [266] find that images in challenges with negative affective valences, like fear, tend to receive lower absolute scores. This implies that a model built on absolute scores rather than relative scores within challenges may be confounded by the emotional valence of an image.

<sup>12</sup><https://data.worldbank.org/indicator/IT.NET.USER.ZS>

Murray, Marchesotti, and Perronnin [266] are highly critical of de-contextualized aesthetic scores: *“The interpretation of these aesthetic judgments [in prior aesthetic datasets], expressed under the form of numeric scores, has always been taken for granted. Yet a deeper analysis of the context in which these judgments are given is essential. The result of this lack of context is that it is difficult to understand what the aesthetic classifiers really model when trained with such datasets.”* This is why AVA represents aesthetic scores as a distribution of votes, allowing one to examine not only the mean aesthetic vote but also the variance. The authors found that images with high variance, or greater disagreement between voters in the image competition, tended to be “non-conventional,” such as “innovative application of photographic techniques.” This nuance, however, is lost because LAP is trained by averaging the aesthetic scores.

## 5.5 Discussion

### 5.5.1 The Algorithmic Gaze

Western art has historically been made to please white, western, heterosexual, men [31, 261, 198]. Similarly, I found that LAP was designed to the taste of one particular white, western, man: the founder of LAION. Perhaps as a result, LAP is aesthetically biased toward realistic images of western and Japanese art. Also, images with captions mentioning women in the LAION-Aesthetics Dataset were more likely to be rated 6.5+ than those mentioning men or LGBTQ+ people. Below, I discuss LAP’s imperial, realist, and male algorithmic gaze.

#### 5.5.1.1 The Imperial Gaze

In my analysis of art from the MET, all of the images rated 6.5+ were paintings, photographs and prints from western or Japanese artists. Not a single piece of African, Oceanian, Native American, Islamic, Egyptian, or West Asian art was rated 6.5+. While the highly-rated Japanese art may seem like an outlier, western artists have been influenced by a racist fascination with “eastern” [321] aesthetics since at least the 1600s [289]. LAP embodies what Kaplan [198] refers to as the imperial gaze, wherein art is judged according white western tastes. Due to the imperial gaze of LAP, the widespread use of this model to curate training data may exacerbate existing multi-modal biases, like struggling to generate Native American [380] or Persian [296] art.

Similar to prior work on biases in data “quality” filtering [98, 179], my findings demonstrate that aesthetic evaluations may lead to representational harms by filtering out non-western art. I encourage future research into how other image evaluation methods, such as FID scores [174] or CLIP similarity [299], assess non-western art. At the same time, I encourage researchers to explore aesthetic biases outside of western ontologies. Notably, none of the images in the MET department of “Africa, the Ancient Americas, and Oceania” scored 6+ (Table 5.1). However, the consolidation of art from different continents and time periods into one department is a vestige of the colonialist category of “primitive art” [294]. By using the MET to evaluate LAP, my analysis inherits this coloniality. In calling for a research agenda focused on aesthetic biases, I caution against reinscribing the imperial gaze embedded in western art history.

### 5.5.1.2 The Realist Gaze

I found that LAP highly rates realistic and figurative images, such as landscape paintings. For centuries, one of the primary criteria for evaluating western visual art was how well it imitates reality [290]. However, the modern art diverged from this tradition by focusing on the ideas and emotions behind art [126], often drawing inspiration from less realistic non-western art [162]. LAP’s realism bias risks devaluing both modern and non-western artistic cultures and parallels troubling historical precedents. Famously, the Degenerate Art Exhibition under Nazi Germany encouraged visitors to mock modern art [355, 260]. Echoing this sensibility, some of the styles and artists from this exhibition, like Picasso, were among the lowest rated in WikiArt.

Realist biases can also limit the creative potential of visual generative AI. A throughline in HCI research on visual generative AI models is that artists often enjoy the glitches or surrealism of text-to-image (T2I) models [347, 380, 65], leading visual artists to critique the increasing realism of newer T2I models [380]. It follows that, aesthetic evaluation models that privilege realism may poorly align with the taste of some visual artists.

So far, I discussed the harms of being *excluded* by LAP’s algorithmic gaze. However, realism biases can also lead to harm through non-consensual *inclusion*. I found that the LAION-Aesthetics 6.5+ Dataset is disproportionately made of images from independent visual artists and photographers. Not only have these communities lost jobs due to visual generative AI [248, 191], but this suggests that their art may be driving the development of these technologies. As increasing photorealism has the potential to exacerbate the greatest harms of visual generative AI—like deepfakes [252], job replacement [192], and sexual abuse imagery [3]—I encourage researchers to reconsider the value placed on realism in visual evaluation.

### 5.5.1.3 The Male Gaze

I found that images with captions mentioning women were more likely to be included in the LAION-Aesthetics 6.5+ Dataset, while those mentioning men or LGBTQ+ people were more likely to be excluded. It follows that datasets curated and models evaluated with LAP may also reflect these biases. In doing so, LAP reinforces the heterosexual male gaze in western visual art history [261, 31] at scale. According to Berger [31], “*Women [in western art] are depicted in a quite different way from men — not because the feminine is different from the masculine — but because the ideal spectator is always assumed to be male and the image of the woman is designed to flatter him.*” I found that LAP was designed according to the taste of one man: LAION’s founder.

The male gaze has long been a feature of computer vision research. In 1973, male researchers at the University of Southern California digitized an image of the *Playboy* model Lena Forsén to test image compression techniques [262]. The “Lena image” has since become ubiquitous in computer vision research. Despite decades of criticism that the Lena image perpetuates misogyny in computing [64, 265], IEEE publications did not prohibit its use until 2024 [173]. Much like Lena’s image circulated without her consent among computer vision researchers for decades, images of (disproportionally) women now similarly circulate among computer vision researchers through the LAION-Aesthetics 6.5+ Dataset.

Although AI fairness researchers often critique data exclusions, Hoffmann [178] critiques including marginalized communities in oppressive technologies. Recently, women have been immensely harmed by the proliferation of visual generative AI models. The United Nations has warned that generative AI is a major contributor to violence against women [59]. This violence has been facilitated by a rise of commercial AI-based nudification apps, often specifically designed to create nude images of women [137]. However, this problem is not unique to dedicated nudification apps. At the time of my writing, thousands of men on X (formerly Twitter) have begun using Grok to create and share NCII and CSAM of women and girls [245]. Of course, I cannot know whether these proprietary models rely on the LAION-Aesthetics Predictor or Datasets. Nevertheless, I expect that aesthetic evaluation models that increase the likelihood of women being included in AI training data risks exacerbating these harms. Much like recent work on how concept filtering impacts T2I CSAM generation [83], I encourage future research on the relationship between aesthetic filtering and the generation of sexual imagery.

## 5.5.2 Reevaluating Evaluation

### 5.5.2.1 Pluralistic Aesthetic Evaluation

I found numerous aesthetic biases in the behavior and design of LAP. However, in raising these issues, I do not advocate for designing “better” universal aesthetic evaluation models or datasets. As assessing the visual aesthetic quality of an image is deeply subjective and value-laden [45], any attempt to represent the aesthetic quality of an image with a single number will be deeply flawed. As an alternative to this normative problem formulation, I encourage researchers to make use of more descriptive evaluating metrics [317]. Although LAP is often used as a general measure of aesthetic quality [95, 119, 130, 168], I found that it could be more accurately described as a measure of photorealism. As described above, there are a number of limitations to using realism to assess the quality of an image. However, a model that measures photorealism is not in-and-of-itself harmful. Rather, I caution researchers against treating photorealism as a prescriptive rather than descriptive property of images, conflating how images *ought* to be with merely how images *can* be.

To support more pluralistic AI alignment [361], I encourage the use of descriptive measurements when evaluating visual generative AI model outputs and training data. Abandoning the flawed construct of universal “aesthetics” would allow researchers to more clearly articulate *which* aesthetic values their visual generative AI models prioritize, such as photorealism or particular visual cultures. For instance, Tencent’s Hunyuan-DiT T2I model is specifically designed to improve Chinese visual element understanding [221]. In addition to improving transparency, more descriptive evaluations may help artists choose models that better align with their tastes, perhaps addressing artists’ critique of increasing T2I photorealism [347, 380]. The seeds of more descriptive aesthetic evaluation can actually be seen in the 2012 AVA dataset itself [266]. Recognizing the limits of de-contextualized aesthetic evaluation, Murray et al. grouped similar photography challenges to measure 14 different photographic styles, such as “motion blur” and “long exposure.”

The companies developing T2I models may have little incentive to shift from prescriptive to de-

scriptive aesthetic evaluations. For-profit companies go to great lengths arguing that their models supplant competitors’ models. For instance, Meta researchers manipulated benchmarks to give an inflated impression of Llama 4’s performance [82]. Acknowledging that different T2I models may be better in different contexts undermines companies’ efforts to prove that *their* models are uniquely “state-of-the-art.” Much like corporate AI rhetoric, academic AI research is increasingly driven by benchmark competition [150, 278], often without questioning the assumptions embedded in said benchmarks [301]. I encourage future researchers to examine what the universal aesthetics measurements used by companies and AI researchers *actually* measure.

### 5.5.2.2 Combining an Audit and Trace Ethnography

In this work, I was able to examine upstream design decisions contributing to downstream biases by combining an audit and trace ethnography. Auditing is a core FAccT methodology, with the regex “audit\*” appearing in nearly half of all FAccT publications. With some notable exceptions [249, 238, 311, 329, 283, 18], ethnography is far less common at FAccT. As Seaver [334] argues, audits can help identify the existence of algorithmic biases while ethnographies can help explain where these biases may have originated by focusing on the people behind the algorithm. In addition to focusing on the harms of AI models themselves, responsible AI researchers are increasingly advocating for “studying up” the powerful people behind AI development and deployment [20, 199]. In service of these goals, I advocate for combining audits and ethnographies as a socio-technical method for evaluating algorithmic systems.

Trace ethnographies can be a powerful tool for studying AI development. While FAccT researchers do not need permission from developers to audit an algorithmic systems, one would likely need permission to conduct participant observations [329]. For example, prior ethnographic FAccT research has relied on access to field sites like tech companies [329, 283] or a police headquarters [238]. However, maintaining access can limit critical scholarship [329]. Even in the case of “open” models, development practices may be opaque [406]. I was only able to understand out how LAP was made by finding an explainer video in an old Discord message. Herein lies a lesson for future ethnographic FAccT research. As Seaver [334] suggests, ethnographies of algorithms may require “scavenging” for diverse materials, including formal documentation, datasets, academic papers, and social media conversations. Nevertheless, this scavenging may pose ethical challenges. For example, studying an unlisted YouTube video—even if posted in a public Discord—could raise ethical concerns. I have chosen to include this video in my analysis due to LAION’s advocacy for data scraping and refusal to remove artists’ work from their datasets [118]. That said, the “public-ness” of data does not always make its analysis ethical [422].

Finally, I caution that trace ethnographies are no substitute for more traditional ethnographic approaches, like extended participant observation and interviews. In my study, for instance, questions remain as to why the founder of LAION chose the datasets from which he built LAP. Likewise, it is not clear how members of LAION make sense of the widespread circulation of LAP. While my trace ethnography presents an account of how LAP was developed, extended participant observation would likely be needed to understand the culture of LAION within which LAP is situated. While I see potential for trace ethnographies to complement audits in FAccT

research by embodying AI developers, this is not the only potential contribution of ethnographic analyses. Similar to Forsythe [120]’s research on how expert system developers conceptualize “experts,” for example, I call for future ethnographic research into the how aesthetic evaluation researchers conceptualize “aesthetics.”

## 5.6 Conclusion

Generative AI models and algorithmic systems are increasingly mediating our media, art, and culture. Therefore, it is imperative to understand for *whose taste* these systems are designed. In this work, I audited the LAION-Aesthetics Predictor (LAP) model, finding that LAP reflects imperial, realist, and male aesthetic sensibilities. Through a trace ethnography of public materials, I found that the development of LAP reflects similar biases. In light of these findings, I call for greater attention to aesthetic biases as a source of representational harm and advocate for more pluralistic evaluations. Much like I examined the algorithmic gaze of LAP, I encourage future research into the cultural consequences of algorithmic “quality” assessment.



# Chapter 6

## Conclusion

In the first part of my dissertation (Chapters 3 & 4), I explored how queer artists wish (not) to engage with generative AI. Despite participants' critiquing how generative AI models are developed, designed, and deployed, participants in both studies found queer artistic potential in the failures or glitches of generative image models. However, participants felt that efforts to "improve" these models straightened out this queer potentiality by favoring realistic visual styles. In light of these critiques, I conducted an audit (Chapter 5) of the widely-used visual aesthetic evaluation model LAION-Aesthetics Predictor (LAP) as well a trace ethnography of how the model was developed. Echoing the critiques of queer artists, I found that the design and behavior of LAP favors realistic images over abstract visual styles. In this section, I conclude by discussing the politics of taste amid the rise of generative AI and how my approach to Queer Computing research is situated within broader debates around queer and trans technology studies.

### 6.1 Queer Art & The Politics of Taste

"Under certain circumstances failing, losing, forgetting, unmaking, undoing, unbecoming, not knowing may in fact offer more creative, more cooperative, more surprising ways of being in the world. Failing is something queers do and have always done exceptionally well; for queers failure can be a style, to cite Quentin Crisp, or a way of life, to cite Foucault, and it can stand in contrast to the grim scenarios of success that depend upon 'trying and trying again.' In fact if success requires so much effort, then maybe failure is easier in the long run and offers different rewards."

- Jack Halberstam in *The Queer Art of Failure* [158]

"Whatever you now find weird, ugly, uncomfortable and nasty about a new medium will surely become its signature. CD distortion, the jitteriness of digital video, the crap sound of 8-bit - all of these will be cherished and emulated as soon as they can be avoided. It's the sound of failure: so much modern art is the sound of things going out of control, of a medium pushing to its limits and breaking apart."

- Brian Eno in *A Year With Swollen Appendices* [108]

From low-budget John Waters films to 100 geecs' glitchy electronic music, queer art is often linked to an "aesthetic of simultaneous failure and virtuosity" [263]. Although the queer artists in my workshop and interview studies held different beliefs about generative AI, participants in both studies saw some queer aesthetic potential in the strangeness of AI-generated images. However, this queer aesthetic sensibility differs from that for which generative AI developers optimize. As we saw in Chapter 5, a popular model used to curate high-quality training data

for visual generative AI models as well as to evaluate the quality of AI-generated images most highly rates images in realistic styles.

Although participants in my studies were highly critical of increasingly realistic T2I models, realism is often an unstated goal in the design of visual generative AI models. For example, Zhang et al. [417] motivate their research on generating realistic hands as follows: “*Text-to-image generation models have achieved remarkable advancements in recent years, aiming to produce realistic images from textual descriptions. However, these models often struggle with generating anatomically accurate representations of human hands. The resulting images frequently exhibit issues such as incorrect numbers of fingers, unnatural twisting or interlacing of fingers, or blurred and indistinct hands.*” Rather than seeing the failure of T2I models to generate realistic hands as a problem to be solved, a queer sensibility may find beauty in the weirdness of hands in AI-generated images.

In my participants praising the queer style of older generative image models, my dissertation research echoes prior work on the preferences for glitches among some artists using visual GenAI [353, 57, 65, 380]. That is to say, queer artists are not the only ones frustrated by the increasing realism of generative image models. More broadly, artists’ dissatisfaction with the style of newer models calls into question what it means for a generative AI model to be “state-of-the-art” for artists. Much like Seaver [335]’s research on how music recommendation algorithm developers compute taste, I see a need for a future research agenda that contends with the politics of taste in generative AI development.

In order to understand the cultural implications of the increasing intrusion of AI-generated media into our daily lives, researchers must understand the tastes embedded in these systems. This will require contending with the taste embedded in technological infrastructures undergirding generative AI development. The LAION-Aesthetics Predictor is only one way that visual generative AI developers measure the quality of training data or generated images. Future research should also explore aesthetic biases in other popular evaluation methods, like comparing the similarity between an image’s caption and its textual description using CLIP embeddings [179]. At the same time, aesthetic evaluation is also not limited to the visual domain. Meta, for instance, recently released an audio aesthetic evaluation model—Meta Audiobox Aesthetics—to support generative audio model development [384]. Notably, the creators of the Meta Audiobox Aesthetics liken their project in the “audio domain” to LAION-Aesthetics in the “visual domain.” In addition to exploring biases embedded in aesthetic evaluation models, I also see a need for research tracing how these models circulate within generative AI research and development communities.

Aesthetic evaluation typically relies on human labor for aesthetic annotations. In the case of the AVA dataset described in Chapter 5, members of the dpchallenge photography competition online community in the late 2000s likely do not know that their votes from over a decade ago drive the development of contemporary generative AI models. On the other hand, members of the Midjourney Discord community are routinely asked to volunteer aesthetic annotations by participating in “rating parties.”<sup>1</sup> Likewise, the start-up Arena<sup>2</sup> evaluates text-to-image models

<sup>1</sup><https://web.archive.org/web/20260407185913/https://updates.midjourney.com/v8-rating-party/>

<sup>2</sup><https://arena.ai/leaderboard/text-to-image>

by having volunteers rate side-by-side AI-generated images, displaying a leader-board ranking the “best” T2I models. Future research should explore the experiences and tastes of people who have involuntarily participated in aesthetic evaluation development, such as the members of dpchallenge, as well as those who people volunteer to participate in aesthetic evaluation for companies like Arena and Midjourney.

In appreciating queer failure, queer art may also help imagine generative AI otherwise. As noted in prior work on content moderation, *scale* is commonly used as an excuse for technology developers to avoid accountability [138]. According to participants, the prevailing paradigm of training generative AI models through large-scale data collection without consent [333] has immensely harmed queer artistic communities. However, one may not need as much data to create glitchy or surreal generative image models, such as Google’s 2015 model DeepDream [255]. Therefore, one may be able to design generative image models *for* queer aesthetics using smaller datasets collected with creators’ consent and using fewer environmental resources [26]. Additionally, less realistic, more failure-prone generative AI models may pose fewer threats to the livelihoods of professional visual artists, a community that has been immensely harmed by the adoption of generative AI in creative industries [191].

## 6.2 Queer HCI in Technofascist Times

In *Trans Technologies*, Haimson [154] identifies tension in designing for marginalized communities between technological inclusionism (“to attempt to fit into, and potentially improve, existing systems”) and technological separatism (“to create new technological systems specifically for that group”). My first HCI research endeavor—studying the subreddit r/bisexual [377]—was an exploration of technological separatism, or how bisexual people successfully created an on-line community for themselves amid their simultaneous marginalization by both gay and straight communities. However, most HCI research at the time was focused, instead, on how LGBTQ+ people are harmed by the design of more general, mainstream technologies [379]. In other words, HCI researchers were focused more on technological inclusionism. Oftentimes, this work was predicated on the assumption that technology developers were generally well intended and could change designs to better support LGBTQ+ communities.

Since I began my PhD in 2021, both the technical and political landscape for LGBTQ+ people in the US has shifted. LGBTQ+ rights—and especially the rights of trans people—are under attack. Public support for gay marriage in the US reached its highest level in 2022, declining every year since.<sup>3</sup> Support for trans rights in the US fell precipitously between 2022 and 2025 among both liberals and conservatives.<sup>4</sup> This waning support has been accompanied by a proliferation of state-sponsored discrimination against LGBTQ+ people, such as restricting access to LGBTQ+ books, drag performances, healthcare, and public accommodations.<sup>5</sup> Major social media com-

<sup>3</sup><https://news.gallup.com/poll/691139/record-party-divide-years-sex-marriage-ruling.aspx>

<sup>4</sup>[https://www.pewresearch.org/short-reads/2025/02/26/americans-have-grown-more-supportive-of-restrictions-for-trans-people-in-recent-years/sr\\_24-02-26\\_transgender-policies\\_2/](https://www.pewresearch.org/short-reads/2025/02/26/americans-have-grown-more-supportive-of-restrictions-for-trans-people-in-recent-years/sr_24-02-26_transgender-policies_2/)

<sup>5</sup><https://www.aclu.org/legislative-attacks-on-lgbtq-rights-2026>

panies have also dramatically rolled back protections for LGBTQ+ people.<sup>6</sup> There has also been increasing collaboration between authoritarian governments and technologists, both within the United States of America and around the world, which has been dubbed “technofascism” [76]. These rightward trends demonstrate the increasing untenability of technological inclusionism.

In the remainder of my dissertation, I reflect on the future of Queer HCI research amid technofascist times.

### 6.2.1 The Rise and Fall of Technological Inclusionism

In my prior research outside this dissertation, I have spent considerable time reflecting on how and why HCI researchers have and should study LGBTQ+ communities [379, 378]. A common critique in my work—as well as recent research from Black [86] and BIPOC [385] scholars—is that HCI researchers tend to think of marginalized communities as problem to be solved. Rather than focusing on designing for joy or the unique cultures of marginalized communities, HCI researchers tend to focus on discrimination and technological harms. Partially due to this homogenization, I have been critical of the “marginalized user” as a unit of analysis in HCI research [378].

Even before I had conducted the research described above, this narrow conception of marginalized communities was evident when I was applying to PhD programs. After explaining my research interests in Queer HCI, a faculty member at an open house event asked if I was interested in designing video games to “solve” homophobia. From the perspective of this senior scholar, designing *for* queer people meant simply mitigating social or technological discrimination. Echoing the liberal civil rights discourses that queer theorists often critique [263], this approach seeks to expand *inclusion* into mainstream society and technology.

I began my PhD in 2021 amid increasing calls for greater inclusion in computing, as evident by the success of books like *Data Feminism* [88], *Design Justice* [80], and *Race After Technology* [27]. Much of the early research in my PhD was responding to this moment when social justice felt *en vogue*, in academia, industry, and policy. In response to a crowd audit of racial biases in Twitter’s image cropping algorithm, for instance, Twitter’s dedicated AI ethics team conducted research to mitigate these issues [94]. In 2023, US president Joe Biden signed an executive order on “Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,” and the National Institute of Standards and Technology actively sought the expertise of ethical computing researchers [300]. Despite the successes of technological inclusionism in this time period, I worried that this theory of change was fragile. Inspired by feminist critiques of liberal democracy [124], I feared that this work placed too much faith in corporations, the state, and the democratic public sphere.

As alluded to above, the political landscape has changed immensely throughout my PhD. In 2022, Elon Musk—the far-right, richest man in the world—bought Twitter partially due to his belief that social media caused his daughter to be transgender [77]. Upon taking control of Twitter, Musk fired anyone he deemed too progressive, including Twitter’s entire AI ethics team

<sup>6</sup><https://glaad.org/smsi/social-media-safety-index-2025/>

[205]. In the early days of Donald Trump’s second presidential term, Musk was enlisted to spearhead a similar ideological purge of the US federal government [407]. This led to the termination of thousands of federal grants for research involving LGBTQ+ communities [407] or People of Color [170]. Academic institutions across the US have similarly retreated from public commitments to social justice, either due to state laws banning “DEI” [315] or for fear of attracting the wrath of the US federal government [67]. The LGBTQ+ resource center I fiercely advocated for as an undergraduate student at Georgia Tech is gone, as is any information on the university’s website that might support LGBTQ+ students.

In many circumstances, HCI researchers cannot trust technology companies or governments to act in the interests of LGBTQ+ communities. In the remainder of this dissertation, I explore alternative possibilities in light of the growing hostility toward LGBTQ+ people from technology companies and governments.

## 6.2.2 A Queer Alternative to Inclusionism

“Often we see queerness as a deprivation, but when I look at my life, I saw that queerness demanded an alternative innovation from me, I had to make alternative routes.”

- Ocean Vuong [402]

Although my dissertation focuses on the power of “alternative [technological] routes,” I am somewhat nostalgic for the performative allyship I critiqued in my prior work [379]. Due to the growing hostilities toward LGBTQ+ people, I have been asked innumerable times how or if I will “pivot” or “reframe” my research. In other words, they ask how I can find ways to be *included* in the face of increasingly anti-LGBTQ+ universities, corporations, and governments. While I do not believe in abandoning LGBTQ+ scholarship, these comments speak to the need to “make alternative routes” in Queer HCI.

In the following sections, I draw on queer theory as a point of departure to inform design in technofascist times. However, this decision does not come without controversy. Queer theorists have long been derided by a subset of scholars—including some HCI researchers—who claim that critical theorists are detached from the everyday experiences of women and LGBTQ+ people [273]. While it is still important to strategically cooperate with policymakers or corporations, the skepticism toward institutions embedded in queer theory demonstrates that alternatives are possible [263, 159]. As Muñoz reminds us, there is utopian potential in rejecting the oppressive “overarching here and now,” with queerness allowing us to build alternative worlds [263].

### 6.2.2.1 Designing for (Il)legibility

Visibility has long been a key component of mainstream gay and lesbian politics. Harvey Milk, the first openly gay man elected to public office in California, famously implored: “Gay brothers and sisters, what are you going to do about it? You must come out.”<sup>7</sup> The desire to include

<sup>7</sup><https://daily.jstor.org/harvey-milks-gay-freedom-day-speech-annotated/>

LGBTQ+ people in the design of technology is often predicated on a similar logic of visibility. LGBTQ+ legibility in sociotechnical systems can allow LGBTQ+ people to better find and understand one another [288, 39]. However, LGBTQ+ datafication can carry risks, such as facilitating discrimination and surveillance [151]. Under the first Trump administration, transgender people were prevented from serving in the military unless they received a formal medical diagnosis of gender dysphoria [177]. Upon returning to power in 2025, the Trump administration used this formal medical classification to fire transgender members of the military. In a different context, apps like Sniffies make it easier for gay men to find cruising spots, but this increased visibility has also been used by police officers in New York City to entrap and arrest queer men [309]. To paraphrase Bowker and Star [46], the datafication of LGBTQ+ lives has consequences.

These tensions around inclusion and legibility can be seen across my studies of queer artists and my investigation of the LAION-Aesthetics Predictor. In Chapter 3, participants critiqued the stereotypical representations and erasure of LGBTQ+ people by the prominent generative AI models GPT-4 and DALL-E 3. In response, some may seek to include more LGBTQ+ people or content in generative AI training data. However, the queer artists I interviewed in Chapter 4 did not necessarily want their work to be more “included” in the design of what they saw as harmful technologies. In the context of aesthetic evaluation, art from the global majority tended to be lowly-rated, while images with descriptions mentioning women were more highly-rated. In these instances, both disproportionate inclusion and exclusion may lead to different harms. For instance, exclusion may exacerbate stereotypes while inclusion may worsen NCII. In considering this tension between exclusion and inclusion, Hoffmann [178] cautions against the desire to increasingly “include” marginalized communities in the design of oppressive systems.

In outlining the potential for queer theory to inform HCI, Anne Light advocates for “small perturbations in responding to the ‘straightness’ of machine thinking” by designing for: forgetting, obscuring, cheating, eluding, and obliqueness [222]. Likewise, research on LGBTQ+ online identity management often advocates for technologies to support selective visibility [60]. To counter the growing harms of LGBTQ+ datafication, I encourage Queer HCI researchers to design for illegibility.

Queer HCI researchers could design for illegibility in many ways. Rather than trying to include more LGBTQ+ people or content in generative AI training data, for instance, HCI could help LGBTQ+ people avoid non-consensual scraping (e.g., [4, 341]). In response to the proliferation of online age verification laws around the world, LGBTQ+ people are increasingly being asked to “prove” their age in order to access hookup apps like Grindr or sexual health information on Reddit [374]. To mitigate these risks, usable privacy researchers could design tools and information resources to help LGBTQ+ people circumvent these systems.

Simultaneously, technology developers often seek to obscure their AI-development practices. This is evident in the changes in how companies building generative AI models describe their training data. As described earlier, Stability AI shared that early versions of Stable Diffusion were trained using LAION datasets. However, the model card for their 2024 model Stable Diffusion 3.5-Large vaguely says it was “*trained on a wide variety of data, including synthetic data*

*and filtered publicly available data.”*<sup>8</sup> Although Huawei does not disclose what images their PixArt- $\Sigma$  model was trained on, they do disclose that they used the LAION-Aesthetics Predictor to curate their “internal” dataset [69]. While the first two studies in my dissertation explore tensions around the visibility of queer art and artists to AI developers, my investigation of LAP is an attempt to make AI-developers’ practices more visible to artists and critical publics. In addition to exploring how to support the visibility of LGBTQ+ communities, Queer HCI scholars should also “study up” [267] major technology developers to hold these actors accountable.

The legal scholar Lawrence Lessig suggests that human-computer interaction is governed by laws, markets, technologies, and norms [214]. As laws and mainstream technologies become more hostile toward LGBTQ+ communities, HCI researchers can support LGBTQ+ communities by supporting illegibility. At the same time, LGBTQ+ people are not limited to using mainstream technologies, like generative AI models and social media sites owned by major tech companies. People have the agency to re-purpose existing technologies and create their own. Next, I explore how HCI researchers might support these practices by designing for pluralism.

### 6.2.2.2 Designing for Pluralism

In the 2010s, the technology development underwent increasing platformization, as evident through the rise of cloud computing services and social media consolidation [365]. Facebook Groups, Subreddits, and Tumblr pages allowed people to create online communities “for free”<sup>9</sup> without contending with the technical particularities of website development. Likewise, cloud computing companies began to provide AI as-a-service, such as automated gender recognition systems [51], abstracting away the details of model development and deployment. Due to this consolidation, HCI researchers could effect change by advocating to these gatekeepers. HCI research on misinformation [368], online harassment [190], and deplatforming [189] informed the moderation of major social media sites. Critical HCI and AI research on automated gender [51, 330] and emotion [15] recognition led Microsoft to retire their facial analysis tools [176].

As described earlier, the internet landscape has changed significantly since I began my PhD. The dominant platforms of the 2010s have begun to splinter [96]. Partially due to the increasingly far-right politics of major social media platforms, the popularity of open-source online communities, like Mastodon and Bluesky, has surged [366]. At the same time, the growth of open-source AI model repositories, like Hugging Face,<sup>10</sup> has decentralized access to predictive and generative AI models and datasets. In light of the harms major technology companies have caused LGBTQ+ communities and artists, this rise of decentralization presents new possibilities to imagine the design of more pluralistic, alternative technologies.

While the participants in my studies of queer artists praised the surreal style of earlier generative image models, these earlier models may no longer be easily accessible. For example, the OpenAI website explains that they no longer support the DALL-E 2 GUI because “DALL-E 3 has higher quality images” [91]. The platformization of generative AI models limits people’s agency

<sup>8</sup><https://huggingface.co/stabilityai/stable-diffusion-3.5-large>

<sup>9</sup>Of course, users still paid for these online communities with their data [365, 423].

<sup>10</sup><https://huggingface.co/>

to choose the best model for their use case, such as ChatGPT users mourning the deprecation of GPT-4o [209]. As described in my study of LAP, what is considered a “better” generative model is deeply subjective. In addition to critiquing the design of “state-of-the-art” closed-source models, there is a need for future research on the active communities of open-source generative AI developers (e.g., the open-source, model-independent interface project SillyTavern<sup>11</sup>). At the same time, researchers should investigate more pluralistic evaluation metrics to help queer people or artists choose the best generative AI models for their use cases and values. Not only would this better support LGBTQ+ communities, but it would also challenge corporate logics of perpetual growth that contribute to human and environmental exploitation.

Although my interest in generative AI and social media may seem unrelated, the companies that own the largest social media platforms (Meta, Google, X) are also dedicated to developing generative AI models. In addition to the right-ward shifts of these platforms, the increasing integration of generative AI into major social media platforms has led to their further enshitification [96]. Likewise, popular social networks for LGBTQ+ communities and artists, like DeviantArt and Reddit, are increasingly selling users’ data to generative AI developers [281, 10]. Future work should explore how artists and LGBTQ+ people are creating their own online communities, such as the anti-AI art community Cara [78], or customizing open-source ones, like LGBTQ+ focused recommendation algorithms on Bluesky. At the same time, HCI researchers should study how we might support these alternative endeavors, helping to pluralize and de-platformitize social media.

Creating alternative social media sites or AI models to support more pluralistic design may sound good in theory. However, HCI researchers must also be aware of the socio-technical challenges posed by alternative technology development. While it is relatively easy to fine-tune existing generative AI models, substantial computing infrastructure is required to train new generative AI models from scratch [406]. For instance, Cretu et al. [84] trained T2I models from scratch to investigate CSAM mitigation, but this required access to high-performance computing infrastructure. In the context of social media, Bluesky is the largest implementer of the open-source AT Protocol, so independent offshoots (e.g., Blacksky) often still rely on some aspects of Bluesky’s centralized infrastructure [346]. Due to the institutional positioning of HCI researchers, I see opportunities for HCI scholars to partner with communities to support alternative computer infrastructures. Even if one has access to computing resources, the technological skills needed to build alternative technology is shaped by existing power relations. For example, Haimson [154] highlights how intersections of race and class impact who gets to design trans technology.

In addition to creating new alternatives, the future of queer design might also lie in the queer uses [7] of existing technologies. Rather than seeking to be included in mainstream technologies, HCI researchers could help people learn to break or unmake [320] existing technologies toward queer ends. However, doing so may require adopting a more antagonistic orientation toward the design of mainstream technologies. For example, the AdNauseam browser extension automatically clicks on every advertisement to obfuscate one’s targeted advertising profile [182], thus interrupting the logics of surveillance capitalism. Another queer use can be seen in the work of transgender artist Ada Ada Ada. In order to challenge the gender essentialism em-

<sup>11</sup><https://github.com/SillyTavern/SillyTavern>

bedded in social media moderation [244] and automated gender recognition systems, Ada posts a shirtless photo of herself on Instagram (@in.transitu\_ig) each week documenting her gender transition accompanied by AI gender estimation scores. The bio of her account reads: “Every Thursday, I post an update from my gender transition. The question is: When will Instagram ban my nipples? By @ada\_ada\_ada.art #DoIHaveBoobsNow.” By re-imagining Instagram as a site for performance art, Ada queers the platform. As Ahmed [7] and Akrich [9] remind us, people do not need to stick to the scripts set out by technology designers.

### 6.2.3 A Queer Conclusion

This dissertation focuses on the relationships between queers, art, and generative AI. In doing so, I continued the pattern I critique in my literature review of Queer HCI: “how research trends—the latest methods and shiniest technologies—drive HCI research (queer or otherwise)” [379]. In light of this critique, I am also struck by Haimson [154]’s observation:

“When I asked trans tech creators about what they envisioned for the future of trans technology, not one of them mentioned a need or desire for AI, machine learning, or even algorithms. Instead, they talked about trans liberation and resistance against rights violations, imagining future possibilities for new trans worlds, utopian and dystopian visions of the future, community and connection, ways to facilitate access to trans health care and body changes, and much more.” [154]

Across my dissertation, I provide myriad recommendations for the design of generative AI models. I find that understanding how generative AI developers evaluate the aesthetic quality of images is a compelling computer science and humanistic endeavor. However, I do not think that the queer artists in my workshop or interview studies deeply desired “better” generative AI models. When I asked about desired futures, participants often imagined alternative futures beyond the oppressive present. As P8 said in my interview study: *“I just feel like if people didn’t have to rely on their art to make money and support themselves that would be a better world. I feel like we wouldn’t be having these conversations about the fear of being replaced [by generative AI]. It’s only because there’s this need to extract profit from art.”* I see a need for Queer HCI to focus more on how to help LGBTQ+ people support one another, which may involve shifting focus away from *en vogue* technologies. Rather than asking how AI chatbots can better support LGBTQ+ people, we should consider how to help LGBTQ+ people support one another through mutual aid [363].

As José Esteban Muñoz suggests, queerness is “something that is not yet here” [263]. Queerness is not a destination but rather a horizon of potentiality. Queerness is a means by which alternative routes and desire paths can be made. Ultimately, queerness reminds HCI researchers that alternative technological futures beyond the oppressive “here and now” are possible.



# Bibliography

- [1] Nov. 2024. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/>.
- [2] 2025. URL: <https://cscw.acm.org/2026/papers.html>.
- [3] Sara M Abdulla. “Are deepfakes “Ab” Normal enough? On using norm shifts in gender and sexuality safety measures”. In: *IEEE Security & Privacy* 23.4 (2025), pp. 15–20.
- [4] William Agnew et al. “Data Defenses Against Large Language Models”. In: *arXiv preprint arXiv:2410.13138* (2024).
- [5] William Agnew et al. “Sound check: Auditing audio datasets”. In: *arXiv preprint arXiv:2410.13114* (2024).
- [6] Sara Ahmed. *Queer Phenomenology: Orientations, Objects, Others*. Duke University Press, 2006.
- [7] Sara Ahmed. *What’s the use?: On the uses of use*. Duke University Press, 2019.
- [8] Shazeda Ahmed et al. “Field-building and the epistemic culture of AI safety”. In: *First Monday* (2024).
- [9] Madeleine Akrich. *The de-scription of technical objects*. 1992.
- [10] Reed Albergotti. *Reddit’s data becomes a battleground in the AI gold rush*. 2025. URL: <https://www.semafor.com/article/10/24/2025/reddits-data-becomes-a-battleground-in-the-ai-gold-rush>.
- [11] Sadiq Aliyu et al. “Participatory Design to Address Disclosure-Based Cyberbullying”. In: *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2024, pp. 1547–1565.
- [12] Mai AlKhamissi et al. “Hire Your Anthropologist! Rethinking Culture Benchmarks Through an Anthropological Lens”. In: *arXiv preprint arXiv:2510.05931* (2025).
- [13] Alyse Marie Allred and Colin M Gray. ““ Be Gay, Do Crimes”: The Co-Production and Activist Potential of Contemporary Fanzines”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2 (2021), pp. 1–35.
- [14] Morgan G Ames. *The charisma machine: The life, death, and legacy of one laptop per child*. Mit Press, 2019.
- [15] Nazanin Andalibi and Justin Buss. “The human in emotion recognition on social media: Attitudes, outcomes, risks”. In: *Proceedings of the 2020 CHI conference on human factors in computing systems*. 2020, pp. 1–16.
- [16] Payal Arora. “The leisure divide: Can the ‘Third World’ come out to play?” In: *Information Development* 28.2 (2012), pp. 93–101.

- [17] Isaac Arturo Ortega Alvarado. “Relationality in design: What can be understood?” In: (2024).
- [18] Stefan Baack. “A critical analysis of the largest source for generative ai training data: Common crawl”. In: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 2199–2208.
- [19] Stefan Baack. “Training Data for the Price of a Sandwich: Common Crawl’s Impact on Generative AI.(Feb. 2024)”. In: *Mozilla Insights* (2024).
- [20] Chelsea Barabas et al. “Studying up: reorienting the study of algorithmic fairness around issues of power”. In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020, pp. 167–176.
- [21] Jeffrey Bardzell and Shaowen Bardzell. “What is” critical” about critical design?” In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2013, pp. 3297–3306.
- [22] Eric PS Baumer and Jed R Brubaker. “Post-userism”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, pp. 6291–6303.
- [23] Eric PS Baumer, Rui Sun, and Peter Schaedler. “Departing and returning: Sense of agency as an organizing concept for understanding social media non/use transitions”. In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), pp. 1–19.
- [24] Andrew M Bean et al. “Measuring what Matters: Construct Validity in Large Language Model Benchmarks”. In: *arXiv preprint arXiv:2511.04703* (2025).
- [25] Howard S Becker. *Art worlds: updated and expanded*. Univ of California Press, 2023.
- [26] Emily M Bender et al. “On the dangers of stochastic parrots: Can language models be too big?” In: *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 2021, pp. 610–623.
- [27] Ruha Benjamin. “Race after technology: Abolitionist tools for the new jim code”. In: *Social forces* (2019).
- [28] Walter Benjamin. “The work of art in the age of mechanical reproduction”. In: *Illuminations*. Random House Digital, Inc., 1986, pp. 217–252.
- [29] Cynthia L Bennett, Erin Brady, and Stacy M Branham. “Interdependence as a frame for assistive technology research and design”. In: *Proceedings of the 20th international acm sigaccess conference on computers and accessibility*. 2018, pp. 161–173.
- [30] Cynthia L Bennett et al. “Painting with Cameras and Drawing with Text: AI Use in Accessible Creativity”. In: *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*. 2024, pp. 1–19.
- [31] John Berger. *Ways of seeing*. Penguin uK, 2008.
- [32] ZOË BERNARD. *Inside the Gay Tech Mafia*. Feb. 2026. URL: <https://www.wired.com/story/inside-the-gay-tech-mafia/>.
- [33] Federico Bianchi et al. “Easily accessible text-to-image generation amplifies demographic stereotypes at large scale”. In: *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 2023, pp. 1493–1504.
- [34] Heidi Biggs, Alexa Marcotte, and Shaowen Bardzell. “TikTok as a stage: Performing rural# farmqueer Utopias on TikTok”. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2023, pp. 946–956.

- [35] Wiebe E Bijker. *Of bicycles, bakelites, and bulbs: Toward a theory of sociotechnical change*. MIT press, 1997.
- [36] Chantelle Billson. *Gay billionaire Sam Altman accuses Democrats of ‘losing the plot’: ‘I am politically homeless’*. Jan. 2025. URL: <https://www.thepinknews.com/2025/07/09/gay-billionaire-sam-altman-accuses-democrats-of-losing-the-plot-i-am-politically-homeless/>.
- [37] Julia Binswanger. A.I. “completes” Keith Haring’s intentionally unfinished painting. Jan. 2024. URL: <https://www.smithsonianmag.com/smart-news/keith-haring-painting-artificial-intelligence-180983563/>.
- [38] Abeba Birhane et al. “Into the laion’s den: Investigating hate in multimodal datasets”. In: *Advances in Neural Information Processing Systems* 36 (2023), pp. 21268–21284.
- [39] Lindsay Blackwell et al. “LGBT parents and social media: Advocacy, privacy, and disclosure during shifting social movements”. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2016, pp. 610–622.
- [40] Su Lin Blodgett et al. “Stereotyping Norwegian salmon: An inventory of pitfalls in fairness benchmark datasets”. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 2021, pp. 1004–1015.
- [41] Margaret A Boden and Ernest A Edmonds. “What is generative art?” In: *Digital Creativity* 20.1-2 (2009), pp. 21–46.
- [42] Kirsten Boehner and Carl DiSalvo. “Data, design and civics: An exploratory study of civic tech”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 2970–2981.
- [43] Kirsten Boehner et al. “How HCI interprets the probes”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2007, pp. 1077–1086.
- [44] Julie Bort. *Coinbase CEO explains why he fired engineers who didn’t try ai immediately*. Aug. 2025. URL: <https://techcrunch.com/2025/08/22/coinbase-ceo-explains-why-he-fired-engineers-who-didnt-try-ai-immediately/>.
- [45] Pierre Bourdieu. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press, 1984.
- [46] Geoffrey C Bowker and Susan Leigh Star. *Sorting things out: Classification and its consequences*. MIT press, 2000.
- [47] Wayne Brekhus. “A sociology of the unmarked: Redirecting our focus”. In: *Sociological theory* 16.1 (1998), pp. 34–51.
- [48] Lauren E Bridges. “Digital failure: Unbecoming the “good” data subject through entropic, fugitive, and queer data”. In: *Big Data & Society* 8.1 (2021), p. 2053951720977882.
- [49] Amy S Bruckman. *Should You Believe Wikipedia?: Online Communities and the Construction of Knowledge*. Cambridge University Press, 2022.
- [50] Thema Bryant-Davis et al. “The trauma lens of police violence against racial and ethnic minorities”. In: *Journal of Social Issues* 73.4 (2017), pp. 852–871.
- [51] Joy Buolamwini and Timnit Gebru. “Gender shades: Intersectional accuracy disparities in commercial gender classification”. In: *Conference on fairness, accountability and transparency*. PMLR. 2018, pp. 77–91.

- [52] Ambar Butchart. “The artificial divide between fine art and textiles is a gendered issue”. In: *Frieze* 14 (2018).
- [53] Judith Butler. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge, 1990.
- [54] Judith Butler. “Performative acts and gender constitution: An essay in phenomenology and feminist theory”. In: *Theatre journal* 40.4 (1988), pp. 519–531.
- [55] Kendra Calhoun and Alexia Fawcett. ““They Edited Out her Nip Nops”: Linguistic Innovation as Textual Censorship Avoidance on TikTok”. In: *Language@ Internet* 21 (2023), pp. 1–30.
- [56] William A Camfield. “Marcel Duchamp’s fountain: Its history and aesthetics in the context of 1917”. In: *Marcel Duchamp: Artist of the century* (1990).
- [57] Baptiste Caramiaux and Sarah Fdili Alaoui. ““Explorers of Unknown Planets” Practices and Politics of Artificial Intelligence in Visual Arts”. In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW2 (2022), pp. 1–24.
- [58] Stuart K Card, Allen Newell, and Thomas P Moran. *The Psychology of Human-Computer Interaction*. 1983.
- [59] Ana Carmo. *Ai and anonymity fuel surge in digital violence against women — UN news*. Nov. 2025. URL: <https://news.un.org/en/story/2025/11/1166411>.
- [60] Matthew Carrasco and Andruud Kerne. “Queer visibility: Supporting LGBT+ selective visibility on social media”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–12.
- [61] Robert L Caserio. “The antisocial thesis in queer theory”. In: *pmla* 121.3 (2006), pp. 819–821.
- [62] Thibault Castells et al. “Ld-pruner: Efficient pruning of latent diffusion models using task-agnostic insights”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024, pp. 821–830.
- [63] Stevie Chancellor et al. “#thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities”. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 2016, pp. 1201–1213.
- [64] Emily Chang. *Brotopia: Breaking up the boys’ club of Silicon Valley*. Penguin, 2019.
- [65] Minsuk Chang et al. “The prompt artists”. In: *Proceedings of the 15th Conference on Creativity and Cognition*. 2023, pp. 75–87.
- [66] Isabelle Chapman, Casey Tolan, and Allison Morrow. *More than a dozen states have passed new laws that led to restrictions on pornography. now, the Supreme Court will weigh in*. Jan. 2025. URL: <https://www.cnn.com/2025/01/11/politics/invs-porn-age-verification-laws-supreme-court/index.html>.
- [67] Bill Chappell. *Trump has sued universities for billions. Here’s what the strategy tells us*. Jan. 2026. URL: <https://www.npr.org/2026/01/29/nx-s1-5559293/trump-settlements-colleges-universities>.
- [68] Caiwei Chen. *Can we fix ai’s evaluation crisis?* June 2025. URL: <https://www.technologyreview.com/2025/06/24/1119187/fix-ai-evaluation-crisis/>.

- [69] Junsong Chen et al. “Pixart- $\sigma$ : Weak-to-strong training of diffusion transformer for 4k text-to-image generation”. In: *European Conference on Computer Vision*. Springer. 2024, pp. 74–91.
- [70] Sophia Chen. “The lost data: how AI systems censor LGBTQ+ content in the name of safety”. In: *Nature computational science* 4.9 (2024), pp. 629–632.
- [71] Ted Chiang. *Why A.I. isn’t going to make art*. Aug. 2024. URL: <https://www.newyorker.com/culture/the-weekend-essay/why-ai-isnt-going-to-make-art>.
- [72] Hyunsung Cho et al. “Auptimize: Optimal Placement of Spatial Audio Cues for Extended Reality”. In: *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 2024, pp. 1–14.
- [73] YouJin Choi et al. “Exploring the Potential of Generative AI in Song-Signing”. In: *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2024, pp. 816–820.
- [74] Monojit Choudhury. “Generative AI has a language problem”. In: *Nature human behaviour* 7.11 (2023), pp. 1802–1803.
- [75] Kenneth Church and Patrick Hanks. “Word association norms, mutual information, and lexicography”. In: *Computational linguistics* 16.1 (1990), pp. 22–29.
- [76] Mark Coeckelbergh. “Technofascism: AI, Big Tech, and the new authoritarianism”. In: *AI & SOCIETY* (2026), pp. 1–14.
- [77] Susie Coen and David Millward. *Elon Musk reveals Twitter takeover driven by ‘woke mind virus’ that infected his trans daughter*. Sept. 2023. URL: <https://www.telegraph.co.uk/world-news/2023/09/01/elon-musk-reveals-twitter-takeover-hatred-woke-mind-virus/>.
- [78] Tanya Combrinck. *The rise of cara: The anti-AI Social Media Platform for artists*. Oct. 2024. URL: <https://www.creativeblog.com/art/digital-art/the-rise-of-cara-the-anti-ai-social-media-platform-for-artists>.
- [79] Juliet Corbin and Anselm Strauss. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications, 2014.
- [80] Sasha Costanza-Chock. *Design justice: Community-led practices to build the worlds we need*. The MIT Press, 2020.
- [81] Jake Coyle. “In Hollywood writers’ battle against AI, humans win (for now)”. In: *AP News*. Accessed January (2023).
- [82] Alex Cranz. *Meta Cheated on AI Benchmarks and It’s a Glimpse Into a New Golden Age*. Apr. 2025. URL: <https://gizmodo.com/meta-cheated-on-ai-benchmarks-and-its-a-glimpse-into-a-new-golden-age-2000586433>.
- [83] Ana-Maria Cretu et al. “Evaluating Concept Filtering Defenses against Child Sexual Abuse Material Generation by Text-to-Image Models”. In: *arXiv preprint arXiv:2512.05707* (2025).
- [84] Ana-Maria Cretu et al. “Evaluating Concept Filtering Defenses against Child Sexual Abuse Material Generation by Text-to-Image Models”. In: *arXiv preprint arXiv:2512.05707* (2025).

- [85] Will Crichton, Maneesh Agrawala, and Pat Hanrahan. “The role of working memory in program tracing”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–13.
- [86] Jay Cunningham et al. “On the grounds of solutionism: Ontologies of blackness and HCI”. In: *ACM Transactions on Computer-Human Interaction* 30.2 (2023), pp. 1–17.
- [87] Jay L. Cunningham et al. “The Cost of Culture: An Analysis of Cash App and the Financial Inclusion of Black American Communities”. In: *Proceedings of the 2022 ACM Designing Interactive Systems Conference*. DIS ’22. Virtual Event, Australia: Association for Computing Machinery, 2022, pp. 612–628. ISBN: 9781450393584. DOI: [10.1145/3532106.3533569](https://doi.org/10.1145/3532106.3533569). URL: <https://doi.org/10.1145/3532106.3533569>.
- [88] Catherine D’ignazio and Lauren F Klein. *Data feminism*. MIT press, 2020.
- [89] Ulrika Dahl. “Not gay as in happy, but queer as in fuck you”. In: *lambda nordica* 19.3-4 (2014), pp. 143–168.
- [90] Doraid Dalalah and Osama MA Dalalah. “The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT”. In: *The International Journal of Management Education* 21.2 (2023), p. 100822.
- [91] DALL-E 2. Jan. 2025. URL: <https://labs.openai.com>.
- [92] Ritendra Datta et al. “Studying aesthetics in photographic images using a computational approach”. In: *European conference on computer vision*. Springer. 2006, pp. 288–301.
- [93] Aida Davani et al. “Disentangling Perceptions of Offensiveness: Cultural and Moral Correlates”. In: *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 2007–2021.
- [94] Alicia DeVrio et al. “Toward User-Driven Algorithm Auditing: Investigating users’ strategies for uncovering harmful algorithmic behavior”. In: *CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–19.
- [95] Wenxin Ding et al. “Understanding implosion in text-to-image generative models”. In: *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*. 2024, pp. 1211–1225.
- [96] Cory Doctorow. *Enshittification: Why everything suddenly got worse and what to do about it*. Verso Books, 2025.
- [97] Cory Doctorow. *Information doesn’t want to be free: laws for the internet age*. McSweeney’s, 2014.
- [98] Jesse Dodge et al. “Documenting large webtext corpora: A case study on the colossal clean crawled corpus”. In: *arXiv preprint arXiv:2104.08758* (2021).
- [99] Jesse Dodge et al. “Measuring the carbon intensity of ai in cloud instances”. In: *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency*. 2022, pp. 1877–1894.
- [100] Lynn Dombrowski, Ellie Harmon, and Sarah Fox. “Social justice-oriented interaction design: Outlining key design strategies and commitments”. In: *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 2016, pp. 656–671.
- [101] Paul Dourish. “Re-Space-Ing Place: ”Place” and ”Space” Ten Years On”. In: *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*. CSCW ’06. Banff, Alberta, Canada: Association for Computing Machinery, 2006,

- pp. 299–308. ISBN: 1595932496. DOI: [10.1145/1180875.1180921](https://doi.org/10.1145/1180875.1180921). URL: <https://doi.org/10.1145/1180875.1180921>.
- [102] Paul Dourish. “Reading and interpreting ethnography”. In: *Ways of Knowing in HCI*. Springer, 2014, pp. 1–23.
  - [103] Lisa Duggan. “The new homonormativity: The sexual politics of neoliberalism”. In: *Materializing democracy: Toward a revitalized cultural politics* 10 (2002), pp. 175–194.
  - [104] Michaelanne Dye. “Un Grano de Arena: Infrastructural Care, Social Media Platforms, and the Venezuelan Humanitarian Crisis”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW3 (2021), pp. 1–28.
  - [105] Brianna Dym, Namita Pasupuleti, and Casey Fiesler. “Building a pillowfort: Political tensions in platform design and policy”. In: *Proceedings of the ACM on Human-Computer Interaction* 6.GROUP (2022), pp. 1–23.
  - [106] Brianna Dym et al. “” Coming Out Okay” Community Narratives for LGBTQ Identity Recovery Work”. In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), pp. 1–28.
  - [107] Lee Edelman. *No future: Queer theory and the death drive*. Duke University Press, 2020.
  - [108] Brian Eno. *A year with swollen appendices: Brian Eno’s Diary*. Faber & Faber, 2020.
  - [109] Ziv Epstein, Antonio Alonso Arechar, and David Rand. “What label should be applied to content produced by generative AI?” In: (2023).
  - [110] Zezhong Fan et al. “Prompt optimizer of text-to-image diffusion models for abstract concept understanding”. In: *Companion Proceedings of the ACM Web Conference 2024*. 2024, pp. 1530–1537.
  - [111] Michael Feffer et al. “Red-Teaming for generative AI: Silver bullet or security theater?” In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. Vol. 7. 2024, pp. 421–437.
  - [112] Jessica L Feuston and Anne Marie Piper. “Everyday experiences: small stories and mental illness on Instagram”. In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 2019, pp. 1–14.
  - [113] Casey Fiesler and Amy S Bruckman. “Creativity, copyright, and close-knit communities: a case study of social norm formation and enforcement”. In: *Proceedings of the ACM on Human-Computer Interaction* 3.GROUP (2019), pp. 1–24.
  - [114] Casey Fiesler and Amy S Bruckman. “Remixers’ understandings of fair use online”. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 2014, pp. 1023–1032.
  - [115] Casey Fiesler and Brianna Dym. “Moving across lands: Online platform migration in fandom communities”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW1 (2020), pp. 1–25.
  - [116] Casey Fiesler, Jessica L Feuston, and Amy S Bruckman. “Understanding copyright law in online creative communities”. In: *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 2015, pp. 116–129.
  - [117] Casey Fiesler, Shannon Morrison, and Amy S Bruckman. “An archive of their own: a case study of feminist HCI and values in design”. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2016, pp. 2574–2585.

- [118] Ekaterina Filikhina and Lennart Elsaß. *Robert Kneschke v. Laion: Judgment of 10 december 2025 (ref.: 5 U 104/24)*. Dec. 2025. URL: <https://www.technologysleageledge.com/2025/12/robert-kneschke-v-laion-judgment-of-10-december-2025-ref-5-u-104-24/>.
- [119] Bi Qi Fong and John See. “Branddiffusion: Multimodal personalized marketing visual content generation”. In: *Proceedings of the 2nd International Workshop on Multimedia Content Generation and Evaluation: New Methods and Practice*. 2024, pp. 72–77.
- [120] Diana Forsythe. *Studying those who study us: An anthropologist in the world of artificial intelligence*. Stanford University Press, 2001.
- [121] Michel Foucault. *The history of sexuality: An introduction*. Vintage, 1990.
- [122] Sarah E Fox and Samantha Shorey. *How augmentation-washing hides labor automation*. 2025.
- [123] Nancy Fraser. “Politics, culture, and the public sphere: Toward a postmodern conception”. In: *Social postmodernism: Beyond identity politics* 291 (1995), p. 295.
- [124] Nancy Fraser. “Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy”. In: *Social Text* 25/26 (1990). Publisher: Duke University Press, pp. 56–80. ISSN: 0164-2472. DOI: [10.2307/466240](https://doi.org/10.2307/466240). URL: <http://www.jstor.org/stable/466240> (visited on 07/02/2021).
- [125] Christopher Frauenberger. “Entanglement HCI the next wave?” In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 27.1 (2019), pp. 1–27.
- [126] Roger Fry. “An essay in aesthetics”. In: *Vision and Design*. Chatto & Windus, 1920, pp. 11–26.
- [127] Kexue Fu et al. “” Being Eroded, Piece by Piece”: Enhancing Engagement and Storytelling in Cultural Heritage Dissemination by Exhibiting GenAI Co-Creation Artifacts”. In: *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2024, pp. 2833–2850.
- [128] Xing Fu and Navid Razmjooy. “Preserving and enhancing cultural heritage through art design using feature pyramid network optimized by modified builder optimization algorithm”. In: *Scientific Reports* 15.1 (2025), p. 42603.
- [129] Anna Furman. *Merriam-Webster’s Word of the Year for 2025 is AI “slop”*. Dec. 2025. URL: <https://www.pbs.org/newshour/nation/merriam-websters-word-of-the-year-for-2025-is-ais-slop>.
- [130] Xiang Gao and Jiaying Liu. “Fbsdif: Plug-and-play frequency band substitution of diffusion features for highly controllable text-driven image translation”. In: *Proceedings of the 32nd ACM International Conference on Multimedia*. 2024, pp. 4101–4109.
- [131] Christine Geeng, Jevan Hutson, and Franziska Roesner. “Usable sexurity: Studying {People’s} concerns and strategies when sexting”. In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 2020, pp. 127–144.
- [132] Clifford Geertz. *The interpretation of cultures*. Vol. 5019. Basic books, 1973.
- [133] R Stuart Geiger and David Ribes. “Trace ethnography: Following coordination through documentary practices”. In: *2011 44th Hawaii international conference on system sciences*. IEEE. 2011, pp. 1–10.

- [134] Marc Gerbracht et al. “‘We are no Luddites!’-CSCW, Co-Determination and Digital Transformation in Germany”. In: *Proceedings of the ACM on Human-Computer Interaction* 8.CSCW1 (2024), pp. 1–23.
- [135] Sourojit Ghosh and Aylin Caliskan. “‘Person’== Light-skinned, Western Man, and Sexualization of Women of Color: Stereotypes in Stable Diffusion”. In: *arXiv preprint arXiv:2310.19981* (2023).
- [136] Sourojit Ghosh, Nina Lutz, and Aylin Caliskan. “‘I Don’t See Myself Represented Here at All’: User Experiences of Stable Diffusion Outputs Containing Representational Harms across Gender Identities and Nationalities”. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. Vol. 7. 2024, pp. 463–475.
- [137] Cassidy Gibson et al. “Analyzing the {AI} nudification application ecosystem”. In: *34th USENIX Security Symposium (USENIX Security 25)*. 2025, pp. 1–20.
- [138] Tarleton Gillespie. “Content moderation, AI, and the question of scale”. In: *Big Data & Society* 7.2 (2020), p. 2053951720943234.
- [139] Tarleton Gillespie. “Do not recommend? Reduction as a form of content moderation”. In: *Social Media+ Society* 8.3 (2022), p. 20563051221117552.
- [140] Tarleton Gillespie. “Generative AI and the politics of visibility”. In: *Big Data & Society* 11.2 (2024), p. 20539517241252131.
- [141] Isaiah Glick. “Consciousness and commodity: towards a critical theory of genre in American popular music”. In: *Culture, Theory and Critique* (2024), pp. 1–19.
- [142] Trystan S Goetze. “AI Art is Theft: Labour, Extraction, and Exploitation: Or, On the Dangers of Stochastic Pollocks”. In: *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 186–196.
- [143] Emma Goldberg. *Living the Slop Life*. May 2025. URL: <https://www.nytimes.com/2025/05/19/style/ai-slop-slop-bowls-shein-slop-hauls.html>.
- [144] Naomi G. Goldberg et al. “Police and the Criminalization of LGBT People”. In: *The Cambridge Handbook of Policing in the United States* (2019). URL: <https://api.semanticscholar.org/CorpusID:202311107>.
- [145] Samuel Goree, Jackson Domingo, and David Crandall. “Human-Centered Evaluation of Aesthetic Quality Assessment Models Using a Smartphone Camera Application”. In: *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 2025, pp. 3265–3275.
- [146] Joan Greenbaum. “Back to Labor: Returning to labor process discussions in the study of work”. In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work*. 1996, pp. 229–237.
- [147] Inderpal Grewal and Caren Kaplan. “Global identities: Theorizing transnational studies of sexuality”. In: *GLQ: A journal of lesbian and gay studies* 7.4 (2001), pp. 663–679.
- [148] John Gruen. *Keith Haring: The Authorized Biography*. Simon and Schuster, 1992.
- [149] Jiayi Guo et al. “Smooth diffusion: Crafting smooth latent spaces in diffusion models”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024, pp. 7548–7558.

- [150] Sireesh Gururaja et al. “To build our future, we must know our past: Contextualizing paradigm shifts in natural language processing”. In: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 2023, pp. 13310–13325.
- [151] Kevin Guyan. *Rainbow trap: Queer lives, classifications and the dangers of inclusion*. Bloomsbury Publishing, 2025.
- [152] Gina Gwenffrewi. “JK Rowling and the echo chamber of secrets”. In: *Transgender Studies Quarterly* 9.3 (2022), pp. 507–516.
- [153] Hana Habib et al. “‘’ It’s a scavenger hunt’: Usability of Websites’ Opt-Out and Data Deletion Choices”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–12.
- [154] Oliver L Haimson. *Trans Technologies*. MIT Press, 2025.
- [155] Oliver L Haimson and Anna Lauren Hoffmann. “Constructing and enforcing” authentic” identity online: Facebook, real names, and non-normative identities”. In: *First Monday* (2016).
- [156] Oliver L Haimson et al. “AI Attitudes Among Marginalized Populations in the US: Nonbinary, Transgender, and Disabled Individuals Report More Negative AI Attitudes”. In: *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 2025, pp. 1224–1237.
- [157] Oliver L Haimson et al. “Tumblr was a trans technology: the meaning, importance, history, and future of trans technologies”. In: *Feminist Media Studies* 21.3 (2021), pp. 345–361.
- [158] Jack Halberstam. “The queer art of failure”. In: *The queer art of failure*. Duke University Press, 2011.
- [159] Jack Halberstam. “What’s that smell? Queer temporalities and subcultural lives”. In: *International journal of cultural studies* 6.3 (2003), pp. 313–333.
- [160] Brett A Halperin and Daniela K Rosner. “‘AI is Soulless’: Hollywood Film Workers Strike and Emerging Perceptions of Generative Cinema”. In: *ACM Transactions on Computer-Human Interaction* (2025).
- [161] Jean Hardy and Silvia Lindtner. “Constructing a desiring user: Discourse, rurality, and design in location-based social networks”. In: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2017, pp. 13–25.
- [162] Elizabeth Harney and Ruth B Phillips. *Mapping modernisms: Art, indigeneity, colonialism*. Duke University Press, 2018.
- [163] Christina N Harrington, Shamika Klassen, and Yolanda A Rankin. “‘All that You Touch, You Change’: Expanding the Canon of Speculative Design Towards Black Futuring”. In: *CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–10.
- [164] Elizabeth A. Harris. *Here are the most targeted books of 2023*. Apr. 2024. URL: <https://www.nytimes.com/2024/04/08/books/banned-books-2023.html>.
- [165] Steve Harrison, Deborah Tatar, and Phoebe Sengers. “The three paradigms of HCI”. In: *Alt. Chi. Session at the SIGCHI Conference on human factors in computing systems San Jose, California, USA*. 2007, pp. 1–18.
- [166] Nabil Hassein. “Against black inclusion in facial recognition”. In: *Digital Talking Drum* 15 (2017).

- [167] Peter S Hawkins. “Naming names: The art of memory and the NAMES project AIDS quilt”. In: *Critical inquiry* 19.4 (1993), pp. 752–779.
- [168] Huiguo He et al. “Learning profitable NFT image diffusions via multiple visual-policy guided reinforcement learning”. In: *Proceedings of the 31st ACM International Conference on Multimedia*. 2023, pp. 6831–6840.
- [169] Jiahui He et al. “Flocking to mastodon: Tracking the great twitter migration”. In: *Proceedings of the 2023 ACM on Internet Measurement Conference*. 2023, pp. 111–123.
- [170] Melissa Hellmann. *A professor had a 2.4m grant to study Black maternal health. Then Trump was elected*. Aug. 2025. URL: <https://www.theguardian.com/world/2025/aug/02/black-maternal-health-infant-trump-cuts>.
- [171] Drew Hemment et al. “AI in the public eye: Investigating public AI literacy through AI art”. In: *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 2023, pp. 931–942.
- [172] Benjamin Henne, Christian Szongott, and Matthew Smith. “SnapMe if you can: Privacy threats of other peoples’ geo-tagged media and what we can do about it”. In: *Proceedings of the sixth ACM conference on Security and privacy in wireless and mobile networks*. 2013, pp. 95–106.
- [173] Alex Hern. *Institute bans use of Playboy Test Image in engineering journals*. Mar. 2024. URL: <https://www.theguardian.com/technology/2024/mar/31/tech-publisher-bans-playboy-centrefold-test-image-from-its-journals>.
- [174] Martin Heusel et al. “Gans trained by a two time-scale update rule converge to a local nash equilibrium”. In: *Advances in neural information processing systems* 30 (2017).
- [175] Daniel Hickey et al. “X under Musk’s leadership: Substantial hate and no reduction in inauthentic activity”. In: *PloS one* 20.2 (2025), e0313293.
- [176] Kashmir Hill. “Microsoft plans to eliminate face analysis tools in push for ‘responsible ai’”. In: *The New York Times* (2022).
- [177] Lauren Hodges. *Former top general calls military’s removal of trans troops a costly mistake*. Feb. 2026. URL: <https://www.npr.org/2026/02/21/nx-s1-5720178/trump-military-ban-rights-transgender-troops>.
- [178] Anna Lauren Hoffmann. “Terms of inclusion: Data, discourse, violence”. In: *New Media & Society* 23.12 (2021), pp. 3539–3556.
- [179] Rachel Hong et al. “Who’s in and who’s out? A case study of multimodal CLIP-filtering in DataComp”. In: *Proceedings of the 4th ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*. 2024, pp. 1–17.
- [180] Lara Houston et al. “Values in repair”. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2016, pp. 1403–1414.
- [181] Dirk Hovy, Federico Bianchi, and Tommaso Fornaciari. ““you sound just like your father” commercial machine translation systems include stylistic biases”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020, pp. 1686–1690.
- [182] Daniel C Howe and Helen Nissenbaum. “Engineering privacy and protest: A case study of adnauseam.” In: *IWPE@ SP*. 2017, pp. 57–64.

- [183] Yipo Huang et al. “Aesbench: An expert benchmark for multimodal large language models on image aesthetics perception”. In: *arXiv preprint arXiv:2401.08276* (2024).
- [184] Yipo Huang et al. “AesExpert: Towards Multi-modality Foundation Model for Image Aesthetics Perception”. In: *Proceedings of the 32nd ACM International Conference on Multimedia*. MM ’24. Melbourne VIC, Australia: Association for Computing Machinery, 2024, pp. 5911–5920. ISBN: 9798400706868. DOI: [10.1145/3664647.3680649](https://doi.org/10.1145/3664647.3680649). URL: <https://doi.org/10.1145/3664647.3680649>.
- [185] Hilary Hutchinson et al. “Technology probes: inspiring design for and with families”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, pp. 17–24.
- [186] Sami P Itävuori. “The Computational Approach To Aesthetics: Value Alignment And The Political Economy Of Attention In Museums And Text-To-Image Generator Stable Diffusion”. In: *A Peer-Reviewed Journal About: Everything Is A Matter Of Distance* 14.1 (2025), pp. 82–100.
- [187] Steven J Jackson and Laewoo Kang. “Breakdown, obsolescence and reuse: HCI and the art of repair”. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2014, pp. 449–458.
- [188] Kyle Jahner. *Google hit with copyright class action over Imagen AI model*. Apr. 2024. URL: <https://news.bloomberglaw.com/ip-law/google-hit-with-copyright-class-action-over-imagen-ai-model>.
- [189] Shagun Jhaver et al. “Evaluating the effectiveness of deplatforming as a moderation strategy on Twitter”. In: *Proceedings of the ACM on human-computer interaction* 5.CSCW2 (2021), pp. 1–30.
- [190] Shagun Jhaver et al. “Online harassment and content moderation: The case of blocklists”. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 25.2 (2018), pp. 1–33.
- [191] Harry H Jiang, Jordan Taylor, and William Agnew. “How Professional Visual Artists are Negotiating Generative AI in the Workplace”. In: *arXiv preprint arXiv:2603.04537* (2026).
- [192] Harry H Jiang et al. “AI Art and its Impact on Artists”. In: *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 2023, pp. 363–374.
- [193] Weiwei Jiang and Louisa Ha. “Smartphones or computers for online sex education? A contraception information seeking model for Chinese college students”. In: *Sex Education* 20.4 (2020), pp. 457–476.
- [194] Jaime Jiménez and Jari Arkko. “AI, Robots. txt”. In: (2024).
- [195] CT Jones. *Fan fiction is about community. could ai ruin that?* May 2025. URL: <https://www.rollingstone.com/culture/culture-features/ai-fan-fiction-community-1235333636/>.
- [196] Nagai Kafū, Kyoko Selden, and Alisa Freedman. “Ukiyo-e landscapes and edo scenic places (1914)”. In: *Review of Japanese Culture and Society* 24 (2012), pp. 210–232.
- [197] Ruogu Kang et al. “{“My” data just goes {Everywhere:”} user mental models of the internet and implications for privacy and security”. In: *Eleventh symposium on usable privacy and security (SOUPS 2015)*. 2015, pp. 39–52.
- [198] E Ann Kaplan. *Looking for the other: Feminism, film and the imperial gaze*. Routledge, 2012.

- [199] Anna Kawakami et al. “Studying Up Public Sector AI: How Networks of Power Relations Shape Agency Decisions Around AI Design and Use”. In: *Proceedings of the ACM on Human-Computer Interaction* 8.CSCW2 (2024), pp. 1–24.
- [200] Reishiro Kawakami and Sukrit Venkatagiri. “The Impact of Generative AI on Artists”. In: *Proceedings of the 16th Conference on Creativity & Cognition*. 2024, pp. 79–82.
- [201] Jofish Kaye et al. “HCI and AI in Industry: Current and Future”. In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–5.
- [202] Os Keyes, Josephine Hoy, and Margaret Drouhard. “Human-computer insurrection: Notes on an anarchist HCI”. In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 2019, pp. 1–13.
- [203] Pranav Khadpe et al. “Hug Reports: Supporting Expression of Appreciation between Users and Contributors of Open Source Software Packages”. In: *Proceedings of the ACM on Human-Computer Interaction* 9.2 (2025), pp. 1–32.
- [204] Aditya Khosla et al. “Understanding and predicting image memorability at a large scale”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 2390–2398.
- [205] Will Knight. *Elon Musk Has Fired Twitter’s ‘Ethical AI’ Team*. Nov. 2022. URL: <https://www.wired.com/story/twitter-ethical-ai-team/>.
- [206] Alistair Knott et al. “Generative AI models should include detection mechanisms as a condition for public release”. In: *Ethics and Information Technology* 25.4 (2023), p. 55.
- [207] Thomas S Kuhn. *The structure of scientific revolutions*. Vol. 111. Chicago University of Chicago Press, 1970.
- [208] Anastasia Kuzminykh and Jessica R Cauchard. “Be mine: Contextualization of ownership research in HCI”. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–9.
- [209] Huiqian Lai. ““Please, don’t kill the only model that still feels human”: Understanding the# Keep4o Backlash”. In: *arXiv preprint arXiv:2602.00773* (2026).
- [210] Jackie Landess. “Politics, Religion, and Sex: Social, Legal, and Medical Equality for LGBTQI Americans”. In: *AMA Journal of Ethics* 12.8 (2010), pp. 606–607.
- [211] Bruno Latour and Steve Woolgar. *Laboratory life: The construction of scientific facts*. Princeton University Press, 2013.
- [212] Matt Laviertes. *Meta’s new hate speech rules allow users to call LGBTQ people mentally ill*. Jan. 2025. URL: <https://www.nbcnews.com/tech/social-media/meta-new-hate-speech-rules-allow-users-call-lgbtq-people-mentally-ill-rcna186700>.
- [213] Kung Jin Lee et al. “The show must go on: A conceptual model of conducting synchronous participatory design with children online”. In: *Proceedings of the 2021 CHI conference on human factors in computing systems*. 2021, pp. 1–16.
- [214] Lawrence Lessig. *Code and Other Laws of Cyberspace*. USA: Basic Books, Inc., 1999. ISBN: 046503912X.
- [215] Lawrence Lessig. *Remix: Making art and commerce thrive in the hybrid economy*. Bloomsbury Academic, 2008.
- [216] Hanlin Li et al. “How do people change their technology use in protest? understanding”. In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), pp. 1–22.

- [217] Hanlin Li et al. “How do people change their technology use in protest?: Understanding “Protest Users””. In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), pp. 1–22.
- [218] Jingyi Li. “Toward Appropriating Tools for Queer Use”. In: *Proceedings of the Halfway to the Future Symposium*. 2024, pp. 1–4.
- [219] Kun Li et al. “AesBiasBench: Evaluating Bias and Alignment in Multimodal Language Models for Personalized Image Aesthetic Assessment”. In: *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. 2025, pp. 7618–7631.
- [220] Zejian Li et al. “Laion-sg: An enhanced large-scale dataset for training complex image-text models with structural annotations”. In: *arXiv preprint arXiv:2412.08580* (2024).
- [221] Zhimin Li et al. “Hunyuan-dit: A powerful multi-resolution diffusion transformer with fine-grained chinese understanding”. In: *arXiv preprint arXiv:2405.08748* (2024).
- [222] Ann Light. “HCI as heterodoxy: Technologies of identity and the queering of interaction with computers”. In: *Interacting with computers* 23.5 (2011), pp. 430–438.
- [223] Cindy Lin and Silvia Margot Lindtner. “Techniques of use: Confronting value systems of productivity, progress, and usefulness in computing and design”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–16.
- [224] Travis Lloyd et al. “AI Rules? Characterizing Reddit Community Policies Towards AI-Generated Content”. In: *arXiv preprint arXiv:2410.11698* (2024).
- [225] Jasmine Lu and Pedro Lopes. “Unmaking electronic waste”. In: *ACM Transactions on Computer-Human Interaction* 31.6 (2024), pp. 1–30.
- [226] Xin Lu et al. “Rapid: Rating pictorial aesthetics using deep learning”. In: *Proceedings of the 22nd ACM international conference on Multimedia*. 2014, pp. 457–466.
- [227] Simian Luo et al. “Latent consistency models: Synthesizing high-resolution images with few-step inference”. In: *arXiv preprint arXiv:2310.04378* (2023).
- [228] Renkai Ma and Yubo Kou. ““I’m not sure what difference is between their content and mine, other than the person itself” A Study of Fairness Perception of Content Moderation on YouTube”. In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW2 (2022), pp. 1–28.
- [229] Zilin Ma et al. “Evaluating the Experience of LGBTQ+ People Using Large Language Model Based Chatbots for Mental Health Support”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–15.
- [230] Cameron MacDonald. “The Alien Resonances and Queer Obscurities of Hyperpop’s 100 geecs”. In: *Journal of Popular Music Studies* 36.2 (2024), pp. 76–98.
- [231] Kelly Avery Mack et al. ““They only care to show us the wheelchair”: Disability representation in text-to-image AI models”. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–23.
- [232] Hughie Mackay and Gareth Gillespie. “Extending the social shaping of technology approach: ideology and appropriation”. In: *Social studies of science* 22.4 (1992), pp. 685–716.
- [233] I Scott MacKenzie. “Fitts’ law as a research and design tool in human-computer interaction”. In: *Human-computer interaction* 7.1 (1992), pp. 91–139.
- [234] Stephen Maddison. “Fags, hags and queer sisters”. In: *Gender Dissent and Heterosocial Bonds in Gay Culture* (2000).

- [235] Clare Malone. *Inside Bari Weiss’s Hostile Takeover of CBS News*. Jan. 2026. URL: <https://www.newyorker.com/magazine/2026/01/26/inside-bari-weiss-hostile-takeover-of-cbs-news>.
- [236] Nahema Marchal et al. “Generative AI Misuse: A Taxonomy of Tactics and Insights from Real-World Data”. In: *arXiv preprint arXiv:2406.13843* (2024).
- [237] George E Marcus. “Ethnography in/of the world system: The emergence of multi-sited ethnography”. In: *Annual review of anthropology* 24.1 (1995), pp. 95–117.
- [238] Vidushi Marda and Shivangi Narayan. “Data in New Delhi’s predictive policing system”. In: *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 2020, pp. 317–324.
- [239] Annette Markham. “The limits of the imaginary: Challenges to intervening in future speculations of memory, data, and algorithms”. In: *New media & society* 23.2 (2021), pp. 382–405.
- [240] Colin Martin. “Art and activism in New York”. In: *The Lancet HIV* 6.10 (2019), e653–e654.
- [241] Alice Marwick et al. “Child-sacrificing drag queens: Historical antecedents in disinformative narratives supporting the Drag Queen Story Hour moral panic”. In: *Women’s Studies in Communication* 47.4 (2024), pp. 459–479.
- [242] Shannon Mattern. “Maintenance and care”. In: *Places Journal* (2018).
- [243] Shannon Mattern. “Step by Step: Thinking through and beyond the repair manual”. In: *Places Journal* (2024).
- [244] Samuel Mayworm, Kendra Albert, and Oliver L Haimson. “Misgendered During Moderation: How Transgender Bodies Make Visible Cisnormative Content Moderation Policies and Enforcement in a Meta Oversight Board Case”. In: *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 301–312.
- [245] Liv McMahan and Laura Cress. *X could face UK ban over deepfakes, minister says*. Jan. 2026. URL: <https://www.bbc.com/news/articles/c99kn52nx9do>.
- [246] Brian Merchant. *Ai is already taking jobs in the video game industry*. July 2024. URL: <https://www.wired.com/story/ai-is-already-taking-jobs-in-the-video-game-industry/>.
- [247] Brian Merchant. *Column: What Stephen King - and nearly everyone else - gets wrong about ai and the Luddites*. Aug. 2023. URL: <https://www.latimes.com/business/technology/story/2023-08-31/column-stephen-king-i-love-you-but-youre-wrong-about-the-luddites-and-technological-progress>.
- [248] Brian Merchant. *What’s really going on with AI and jobs?* Nov. 2025. URL: <https://www.bloodinthemachine.com/p/what-the-hell-is-going-on-with-ai>.
- [249] Milagros Miceli et al. “Documenting computer vision datasets: An invitation to reflexive data practices”. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 2021, pp. 161–172.
- [250] Michael G Milgroom. “The Germ Theory Paradigm”. In: *Biology of Infectious Disease: From Molecules to Ecosystems*. Springer, 2023, pp. 9–22.

- [251] Nusrat Jahan Mim et al. “In-Between Visuals and Visible: The Impacts of Text-to-Image Generative AI Tools on Digital Image-making Practices in the Global South”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–18.
- [252] Jaron Mink et al. “It’s Trying Too Hard To Look Real: Deepfake Moderation Mistakes and Identity-Based Bias”. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–20.
- [253] Piotr Mirowski et al. “A Robot Walks into a Bar: Can Language Models Serve as Creativity Support Tools for Comedy? An Evaluation of LLMs’ Humour Alignment with Comedians”. In: *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 1622–1636.
- [254] Youssef Mohamed et al. “Artelingo: A million emotion annotations of wikiart with emphasis on diversity over language and culture”. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. 2022, pp. 8770–8785.
- [255] Alexander Mordvintsev, Christopher Olah, and Mike Tyka. “Inceptionism: Going deeper into neural networks”. In: *Google research blog* 20.14 (2015), p. 5.
- [256] Bill Morgan. *Howl on Trial: the battle for free expression*. City Lights Books, 2021.
- [257] Meredith Ringel Morris. “HCI for AGI”. In: *Interactions* 32.2 (2025), pp. 26–32.
- [258] Meredith Ringel Morris. “Prompting Considered Harmful”. In: *Communications of the ACM* (2024).
- [259] Meredith Ringel Morris et al. “Is Human-AI Interaction CSCW?” In: *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing*. 2024, pp. 95–97.
- [260] Patt Morrison. ‘Degenerate’ or ‘woke,’ Paris museum exhibit shows what happens to art in the crosshairs of politics. May 2025. URL: <https://www.latimes.com/entertainment-arts/story/2025-05-07/degenerate-or-woke-paris-museum-exhibit-evokes-what-happens-to-art-in-the-crosshairs-of-politics>.
- [261] Laura Mulvey. “Visual Pleasure and Narrative Cinema”. In: *Screen* 16.3 (Oct. 1975), pp. 6–18. ISSN: 0036-9543. DOI: [10.1093/screen/16.3.6](https://doi.org/10.1093/screen/16.3.6). URL: <https://doi.org/10.1093/screen/16.3.6> (visited on 12/31/2025).
- [262] Dylan Mulvin. *Proxies: The cultural work of standing in*. MIT Press, 2021.
- [263] José Esteban Muñoz. “Cruising utopia”. In: *Cruising Utopia, 10th Anniversary Edition*. New York University Press, 2019.
- [264] José Esteban Muñoz. “Ephemera as evidence: Introductory notes to queer acts”. In: (1996).
- [265] David C Munson. “A note on Lena”. In: *IEEE Transactions on Image Processing* 5.1 (1996), pp. 3–3.
- [266] Naila Murray, Luca Marchesotti, and Florent Perronnin. “AVA: A large-scale database for aesthetic visual analysis”. In: *2012 IEEE conference on computer vision and pattern recognition*. IEEE. 2012, pp. 2408–2415.
- [267] Laura Nader. “Up the anthropologist: Perspectives gained from studying up.” In: (1972).
- [268] Marissa Newman and Aggi Cantrill. *The Future of AI Relies on a High School Teacher’s Free Database*. Apr. 2023. URL: <https://www.bloomberg.com/news/features/>

[2023-04-24/a-high-school-teacher-s-free-image-database-powers-ai-unicorns.](#)

- [269] Safiya Umoja Noble. “Algorithms of oppression: How search engines reinforce racism”. In: *Algorithms of oppression*. New York university press, 2018.
- [270] Chon Noriega. “”Something’s Missing Here!”: Homosexuality and Film Reviews during the Production Code Era, 1934-1962”. In: *Cinema Journal* 30.1 (1990), pp. 20–41.
- [271] Fayika Farhat Nova et al. “” Facebook Promotes More Harassment” Social Media Ecosystem, Skill and Marginalized Hijra Identity in Bangladesh”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW1 (2021), pp. 1–35.
- [272] Claudio Novelli et al. “Generative AI in EU law: liability, privacy, intellectual property, and cybersecurity”. In: *arXiv preprint arXiv:2401.07348* (2024).
- [273] Martha C Nussbaum. “The professor of parody.” In: *New Republic* 220.8 (1999), pp. 37–45.
- [274] William Odom. “Understanding long-term interactions with a slow technology: an investigation of experiences with FutureMe”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 575–584.
- [275] William Odom et al. “Attending to slowness and temporality with olly and slow game: A design inquiry into supporting longer-term relations with everyday computational objects”. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 2018, pp. 1–13.
- [276] William T Odom et al. “Designing for slowness, anticipation and re-visitation: a long term field study of the photobox”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014, pp. 1961–1970.
- [277] A.W. Ohlheiser. *A poster’s guide to WHO’s selling your data to train AI*. Feb. 2024. URL: <https://www.vox.com/technology/24086039/reddit-tumblr-wordpress-whos-selling-your-data-to-train-ai>.
- [278] Will Orr and Edward B Kang. “AI as a sport: On the competitive epistemologies of benchmarking”. In: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 1875–1884.
- [279] Antti Oulasvirta and Pertti Saariluoma. “Long-term working memory and interrupting messages in human–computer interaction”. In: *Behaviour & Information Technology* 23.1 (2004), pp. 53–64.
- [280] Susanna Paasonen and Jenny Sundén. “Objectionable nipples: Puritan data politics and sexual agency in social media”. In: *Queer data*. Routledge (2024).
- [281] Nitish Pahwa. *DeviantArt’s downfall is devastating, depressing, and dumb*. May 2024. URL: <https://slate.com/technology/2024/05/deviantart-what-happened-ai-decline-lawsuit-stability.html>.
- [282] Rock Yuren Pang et al. “Understanding the LLM-ification of CHI: Unpacking the Impact of LLMs at CHI through a Systematic Literature Review”. In: *arXiv preprint arXiv:2501.12557* (2025).
- [283] Samir Passi and Solon Barocas. “Problem formulation and fairness”. In: *Proceedings of the conference on fairness, accountability, and transparency*. 2019, pp. 39–48.

- [284] Blakeley H Payne et al. “How to Ethically Engage Fat People in HCI Research”. In: *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 2023, pp. 117–121.
- [285] Lucy Pei and Bonnie Nardi. “We did it right, but it was still wrong: Toward assets-based design”. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, pp. 1–11.
- [286] Antonio Pequeño. *Tiktok Ban Timeline: Tiktok returns to U.S. app stores-for now*. Feb. 2025. URL: <https://www.forbes.com/sites/antoniopequenoiv/2025/02/14/tiktok-ban-timeline-tiktok-returns-to-us-app-stores-for-now/>.
- [287] Trevor J Pinch and Wiebe E Bijker. “The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other”. In: *Social studies of science* 14.3 (1984), pp. 399–441.
- [288] Anthony T Pinter, Morgan Klaus Scheuerman, and Jed R Brubaker. “Entering doors, evading traps: Benefits and risks of visibility during transgender coming outs”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW3 (2021), pp. 1–27.
- [289] David Porter. “Chinoiserie and the Aesthetics of Illegitimacy”. In: *Studies in Eighteenth-Century Culture* 28.1 (1999), pp. 27–54.
- [290] Matthew Potolsky. *Mimesis*. Routledge, 2006.
- [291] Whitney Pow. “A trans historiography of glitches and errors”. In: *Feminist media histories* 7.1 (2021), pp. 197–230.
- [292] John David Pressman, Katherine Crowson, and Simulacra Captions Contributors. *Simulacra Aesthetic Captions*. Version 1.0. Stability AI, 2022. URL: <https://github.com/JD-P/simulacra-aesthetic-captions>.
- [293] Sally Price. *Primitive art in civilized places*. University of Chicago Press, 2001.
- [294] Sally Price. *Primitive art in civilized places*. University of Chicago Press, 2001.
- [295] Mahima Pushkarna, Andrew Zaldivar, and Oddur Kjartansson. “Data cards: Purposeful and transparent dataset documentation for responsible ai”. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 2022, pp. 1776–1826.
- [296] Rida Qadri, Piotr Mirowski, and Remi Denton. “AI and Non-Western Art Worlds: Reimagining Critical AI Futures through Artistic Inquiry and Situated Dialogue”. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 2025, pp. 1–17.
- [297] Li Qiwei et al. “The sociotechnical stack: Opportunities for social computing research in non-consensual intimate media”. In: *Proceedings of the ACM on Human-Computer Interaction* 8.CSCW2 (2024), pp. 1–21.
- [298] Organizers Of Queerinaï et al. “Queer in AI: A case study in community-led participatory AI”. In: *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 2023, pp. 1882–1895.
- [299] Alec Radford et al. “Learning transferable visual models from natural language supervision”. In: *International conference on machine learning*. PmLR. 2021, pp. 8748–8763.
- [300] Courtney C Radsch et al. “1 year later, how has the White House AI Executive Order delivered on its promises?” In: (2024).

- [301] Inioluwa Deborah Raji et al. “AI and the everything in the whole wide world benchmark”. In: *arXiv preprint arXiv:2111.15366* (2021).
- [302] Aditya Ramesh et al. “Hierarchical text-conditional image generation with clip latents”. In: *arXiv preprint arXiv:2204.06125* 1.2 (2022), p. 3.
- [303] Javier Rando et al. “Red-teaming the stable diffusion safety filter”. In: *arXiv preprint arXiv:2210.04610* (2022).
- [304] Alex Reisner. “Revealed: The authors whose pirated books are powering generative AI”. In: *The Atlantic* 19 (2023).
- [305] Alex Reisner. “The Company Quietly Funneling Paywalled Articles to AI Developers”. In: *Atlantic* (2025).
- [306] Alex Reisner. “These 183,000 books are fueling the biggest fight in publishing and tech”. In: *Atlantic* (2023).
- [307] Alex Reisner and Annie Gilbertson. *Apple, Nvidia, anthropic used thousands of swiped YouTube videos to train AI*. July 2024. URL: <https://www.wired.com/story/youtube-training-data-apple-nvidia-anthropic/>.
- [308] Gisela Reyes-Cruz et al. “Resisting AI Solutionism: Where Do We Go From Here?”. In: *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 2025, pp. 1–6.
- [309] Samantha Riedel. “Lewdness” Arrests Up at NYC Penn Station Bathrooms, Police Likely Using Sniffies. Sept. 2025. URL: <https://www.them.us/story/nyc-penn-station-sniffies-arrests>.
- [310] Alexandra Teixeira Riggs et al. “Designing an archive of feelings: Queering tangible interaction with button portraits”. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–17.
- [311] Sarah Riley. “Overriding (in) justice: pretrial risk assessment administration on the front-lines”. In: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 480–488.
- [312] Adi Robertson. *Artists’ lawsuit against stability AI and Midjourney gets more punch*. Aug. 2024. URL: <https://www.theverge.com/2024/8/13/24219520/stability-midjourney-artist-lawsuit-copyright-trademark-claims-approved>.
- [313] Reece Rogers. *Here’s how Generative AI depicts queer people*. Apr. 2024. URL: <https://www.wired.com/story/artificial-intelligence-lgbtq-representation-openai-sora/>.
- [314] Yvonne Rogers. *HCI theory: classical, modern, and contemporary*. Vol. 14. Morgan & Claypool Publishers, 2012.
- [315] Amber Roldan. *LGBTQ-friendly resources, policies recede on Georgia’s college campuses as DEI rollbacks continue*. Aug. 2025. URL: <https://georgiarecorder.com/2025/08/18/lgbtq-friendly-resources-policies-recede-on-georgias-college-campuses-as-dei-rollbacks-continue/>.
- [316] Robin Rombach et al. “High-resolution image synthesis with latent diffusion models”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, pp. 10684–10695.

- [317] Paul Röttger et al. “Two contrasting data annotation paradigms for subjective NLP tasks”. In: *Proceedings of the 2022 conference of the North American chapter of the association for computational linguistics: human language technologies*. 2022, pp. 175–190.
- [318] Paul Röttger et al. “Xstest: A test suite for identifying exaggerated safety behaviours in large language models”. In: *arXiv preprint arXiv:2308.01263* (2023).
- [319] Vito Russo. *The Celluloid Closet: Homosexuality in the Movies*. HarperCollins, 1987.
- [320] Samar Sabie et al. “Unmaking as Emancipation: Lessons and Reflections from Ludism”. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–15.
- [321] Edward W Said. *Orientalism*. Vintage, 1979.
- [322] Babak Saleh and Ahmed Elgammal. “Large-scale classification of fine-art paintings: Learning the right metric on the right feature”. In: *arXiv preprint arXiv:1505.00855* (2015).
- [323] Princess Sampson, Ro Encarnacion, and Danaë Metaxa. “Representation, Self-Determination, and Refusal: Queer People’s Experiences with Targeted Advertising”. In: *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 2023, pp. 1711–1722.
- [324] Téó Sanchez. “Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs?” In: *Proceedings of the 15th Conference on Creativity and Cognition*. 2023, pp. 43–61.
- [325] Maarten Sap et al. “Annotators with attitudes: How annotator beliefs and identities bias toxic language detection”. In: *Proceedings of the 2022 conference of the north american chapter of the association for computational linguistics: Human language technologies*. 2022, pp. 5884–5906.
- [326] Maarten Sap et al. “The risk of racial bias in hate speech detection”. In: *Proceedings of the 57th annual meeting of the association for computational linguistics*. 2019, pp. 1668–1678.
- [327] Christine Satchell and Paul Dourish. “Beyond the user: use and non-use in HCI”. In: *Proceedings of the 21st annual conference of the Australian computer-human interaction special interest group: Design: Open 24/7*. 2009, pp. 9–16.
- [328] Mark Savage. *Nick Cave says Chatgpt’s ai attempt to write Nick Cave Lyrics “sucks”*. Jan. 2023. URL: <https://www.bbc.com/news/entertainment-arts-64302944>.
- [329] Morgan Klaus Scheuerman. “In the walled garden: Challenges and opportunities for research on the practices of the AI tech industry”. In: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*. 2024, pp. 456–466.
- [330] Morgan Klaus Scheuerman, Jacob M Paul, and Jed R Brubaker. “How computers see gender: An evaluation of gender classification in commercial facial analysis services”. In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), pp. 1–33.
- [331] Morgan Klaus Scheuerman et al. “From human to data to dataset: Mapping the traceability of human subjects in computer vision datasets”. In: *Proceedings of the ACM on Human-Computer Interaction* 7.CSCW1 (2023), pp. 1–33.

- [332] Christoph Schuhmann. *LAION-AESTHETICS*. Aug. 2022. URL: <https://laion.ai/blog/laion-aesthetics>.
- [333] Christoph Schuhmann et al. “Laion-5b: An open large-scale dataset for training next generation image-text models”. In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 25278–25294.
- [334] Nick Seaver. “Algorithms as culture: Some tactics for the ethnography of algorithmic systems”. In: *Big data & society* 4.2 (2017), p. 2053951717738104.
- [335] Nick Seaver. “Computing taste: Algorithms and the makers of music recommendation”. In: *Computing Taste*. University of Chicago Press, 2022.
- [336] Nick Seaver. “Computing taste: Algorithms and the makers of music recommendation”. In: *Computing Taste*. University of Chicago Press, 2022.
- [337] Eve Kosofsky Sedgwick. “Queer performativity: Henry James’s *The art of the novel*”. In: *GLQ-NEW YORK*- 1 (1993), pp. 1–16.
- [338] Eve Kosofsky Sedgwick. *Tendencies*. Duke University Press, 1993.
- [339] Bryan C Semaan, Lauren M Britton, and Bryan Dosono. “Transition Resilience with ICTs: ‘Identity Awareness’ in Veteran Re-Integration”. In: *Proceedings of the 2016 CHI Conference on human factors in computing systems*. 2016, pp. 2882–2894.
- [340] Phoebe Sengers and Bill Gaver. “Staying open to interpretation: engaging multiple meanings in design and evaluation”. In: *Proceedings of the 6th conference on Designing Interactive systems*. 2006, pp. 99–108.
- [341] Shawn Shan et al. “Glaze: Protecting artists from style mimicry by {Text-to-Image} models”. In: *32nd USENIX Security Symposium (USENIX Security 23)*. 2023, pp. 2187–2204.
- [342] Shawn Shan et al. “Nightshade: Prompt-Specific Poisoning Attacks on Text-to-Image Generative Models”. In: *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE Computer Society. 2024, pp. 212–212.
- [343] Eva Sharma et al. “Analyzing ideological discourse on social media: A case study of the abortion debate”. In: *Proceedings of the 2017 International Conference of The Computational Social Science Society of the Americas*. 2017, pp. 1–8.
- [344] Vishal Sharma, Neha Kumar, and Bonnie Nardi. “Post-growth human–computer interaction”. In: *ACM Transactions on Computer-Human Interaction* 31.1 (2023), pp. 1–37.
- [345] Tania Sharmin and Sanyat Sattar. “Gender politics in the projection of “Disney” villains”. In: *Journal of Literature and Art Studies* 8.1 (2018), pp. 53–57.
- [346] Matthew Sheffield. *User ban controversy reveals Bluesky’s decentralized aspiration isn’t reality*. 2025. URL: <https://plus.flux.community/p/banning-controversy-reveals-blueskys>.
- [347] Renee Shelby, Shalaleh Rismani, and Negar Rostamzadeh. “Generative AI in creative practice: ML-artist folk theories of T2I use, harm, and harm-reduction”. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–17.
- [348] Renee Shelby et al. “Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction”. In: *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 2023, pp. 723–741.

- [349] Hong Shen et al. “Everyday algorithm auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2 (2021), pp. 1–29.
- [350] Amanda Silberling. *Fan fiction writers are trolling AIS with Omegaverse stories*. June 2023. URL: <https://techcrunch.com/2023/06/13/fan-fiction-writers-are-trolling-ais-with-omegaverse-stories/>.
- [351] Amanda Silberling. *Fan fiction writers are trolling AIS with Omegaverse stories*. June 2023. URL: <https://techcrunch.com/2023/06/13/fan-fiction-writers-are-trolling-ais-with-omegaverse-stories/>.
- [352] Ellen Simpson and Bryan Semaan. “For You, or For” You”? Everyday LGBTQ+ Encounters with TikTok”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW3 (2021), pp. 1–34.
- [353] Christian Sivertsen et al. “Machine Learning Processes as Sources of Ambiguity: Insights from AI Art”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–14.
- [354] Shanté Paradigm Smalls and Elliott H Powell. *Introduction: An impossibility: Black queer and trans\* aesthetics*. 2019.
- [355] Sebastian Smee. *The nazis’ war on ‘degenerate art’ was modern art’s greatest crisis*. Mar. 2025. URL: <https://www.washingtonpost.com/entertainment/art/2025/03/14/degenerate-art-musee-picasso-nazi-culture/>.
- [356] Aisha Sobey. “The thinness of GenAI: body size in relation to the construction of the normate through GenAI image models”. In: *AI and Ethics* (2025), pp. 1–16.
- [357] Robert Soden et al. “Time for Historicism in CSCW: An Invitation”. In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW2 (Oct. 2021). DOI: 10.1145/3479603. URL: <https://doi.org/10.1145/3479603>.
- [358] Johan Söderberg. “Copyleft vs. copyright: A Marxist critique”. In: *First Monday* (2002).
- [359] Susan Sontag. *Notes on camp*. Penguin UK, 2018.
- [360] Susan Sontag. *On Photography*. Picador, 1977.
- [361] Taylor Sorensen et al. “A roadmap to pluralistic alignment”. In: *arXiv preprint arXiv:2402.05070* (2024).
- [362] Taylor Sorensen et al. “Value kaleidoscope: Engaging ai with pluralistic human values, rights, and duties”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 38. 18. 2024, pp. 19937–19947.
- [363] Dean Spade. *Mutual aid: Building solidarity during this crisis (and the next)*. Verso Books, 2020.
- [364] Katta Spiel. “” Why are they all obsessed with Gender?”—(Non) binary Navigations through Technological Infrastructures”. In: *Designing Interactive Systems Conference 2021*. 2021, pp. 478–494.
- [365] Nick Srnicek. *Platform capitalism*. John Wiley & Sons, 2017.
- [366] Verge Staff. *The hunt for the next Twitter: all the news about alternative social media platforms*. 2026. URL: <https://www.theverge.com/23686584/twitter-alternative-social-media-platforms-mastodon-bluesky-activitypub-protocol>.

- [367] Logan Stapleton et al. “Seeing seeds beyond weeds: Green teaming generative ai for beneficial uses”. In: *arXiv preprint arXiv:2306.03097* (2023).
- [368] Kate Starbird, Ahmer Arif, and Tom Wilson. “Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations”. In: *Proceedings of the ACM on human-computer interaction* 3.CSCW (2019), pp. 1–26.
- [369] Ella Steen, Kathryn Yurechko, and Daniel Klug. “You can (not) say what you want: Using algospeak to contest and evade algorithmic content moderation on TikTok”. In: *Social Media+ Society* 9.3 (2023), p. 20563051231194586.
- [370] Lucy Suchman. “Do categories have politics? The language/action perspective reconsidered”. In: *Computer supported cooperative work (CSCW)* 2 (1993), pp. 177–190.
- [371] Lucy Suchman. “The uncontroversial ‘thingness’ of AI”. In: *Big Data & Society* 10.2 (2023), p. 20539517231206794.
- [372] Lucy A Suchman. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press, 1987.
- [373] Joice Tang et al. “Back to “Back to Labor”: Revisiting Political Economies of Computer-Supported Cooperative Work”. In: *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing*. 2023, pp. 522–526.
- [374] Brooke Tanner and Nicole Turner Lee. *Children’s online safety laws are failing LGBTQ+ youth*. July 2025. URL: <https://www.brookings.edu/articles/childrens-online-safety-laws-are-failing-lgbtq-youth/>.
- [375] Yan Tao et al. “Cultural bias and cultural alignment of large language models”. In: *PNAS nexus* 3.9 (2024), p. 346.
- [376] Amancay Tapia. *The 20 biggest museums in the world*. Sept. 2021. URL: <https://www.newsweek.com/biggest-museums-world-louvre-metropolitan-vatican-china-1629383>.
- [377] Jordan Taylor and Amy Bruckman. “Mitigating Epistemic Injustice: The Online Construction of a Bisexual Culture”. In: *ACM Transactions on Computer-Human Interaction* (2024).
- [378] Jordan Taylor et al. “Carefully Unmaking the “Marginalized User:” A Diffractive Analysis of a Gay Online Community”. In: *ACM Transactions on Computer-Human Interaction* (2024).
- [379] Jordan Taylor et al. “Cruising Queer HCI on the DL: A Literature Review of LGBTQ+ People in HCI”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–21.
- [380] Jordan Taylor et al. “Un-Straightening Generative AI: How Queer Artists Surface and Challenge the Normativity of Generative AI Models”. In: *arXiv preprint arXiv:2503.09805* (2025).
- [381] Dias Oliva Thiago, Antonialli Dennys Marcelo, and Alessandra Gomes. “Fighting hate speech, silencing drag queens? artificial intelligence in content moderation and risks to lgbtq voices online”. In: *Sexuality & culture* 25.2 (2021), pp. 700–732.
- [382] David Thiel. “Identifying and eliminating csam in generative ml training data and models”. In: *Stanford Internet Observatory, Cyber Policy Center, December 23* (2023), p. 3.
- [383] Nanna Thylstrup and Zeerak Talat. “Detecting ‘dirt’ and ‘toxicity’: Rethinking content moderation as pollution behaviour”. In: *Available at SSRN 3709719* (2020).

- [384] Andros Tjandra et al. “Meta audiobox aesthetics: Unified automatic quality assessment for speech, music, and sound”. In: *arXiv preprint arXiv:2502.05139* (2025).
- [385] Alexandra To et al. “Flourishing in the everyday: Moving beyond damage-centered design in HCI for BIPOC communities”. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2023, pp. 917–933.
- [386] Sam Tobin. *Getty Images largely loses landmark UK lawsuit over AI Image Generator*. Nov. 2025. URL: <https://www.reuters.com/sustainability/boards-policy-regulation/getty-images-largely-loses-landmark-uk-lawsuit-over-ai-image-generator-2025-11-04/>.
- [387] Anh-Ton Tran et al. “Making Trouble: Techniques for Queering Data and AI Systems”. In: *Companion Publication of the 2024 ACM Designing Interactive Systems Conference*. 2024, pp. 381–384.
- [388] Petru-Daniel Tudosiu et al. “Mulan: A multi layer annotated dataset for controllable text-to-image generation”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024, pp. 22413–22422.
- [389] Matthew Turk and George Robertson. “Perceptual user interfaces (introduction)”. In: *Communications of the ACM* 43.3 (2000), pp. 32–34.
- [390] Eddie Ungless, Björn Ross, and Anne Lauscher. “Stereotypes and Smut: The (Mis) representation of Non-cisgender Identities by Text-to-Image Models”. In: *Findings of the Association for Computational Linguistics: ACL 2023*. 2023, pp. 7919–7942.
- [391] *Usage Policy*. Jan. 2025. URL: <https://web.archive.org/web/20241211193440/https://openai.com/policies/usage-policies/>.
- [392] P Vaughan. “The New York Times’ AI copyright lawsuit shows that forgiveness might not be better than permission”. In: *The Conversation* (2024).
- [393] Janet Vertesi. *About the Opt Out Project*. Oct. 2022. URL: <https://www.optoutproject.net/about-the-opt-out-project/>.
- [394] Bruno Caldas Vianna. “Aesthetic biases and opacity tactics in the training of visual artificial intelligence models”. In: *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. Springer. 2025, pp. 278–293.
- [395] Nicholas Vincent et al. “Data leverage: A framework for empowering the public in its relationship with technology companies”. In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 2021, pp. 215–227.
- [396] Matthew Vines. *I’m Gay, Not Queer. It Matters*. June 2026. URL: <https://www.nytimes.com/2026/06/30/opinion/queer-gay-rights.html>.
- [397] Johanna Walker et al. “AI art and misinformation: approaches and strategies for media literacy and fact checking”. In: *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 2023, pp. 26–37.
- [398] Ding Wang et al. “A case for moving beyond “gold data” in AI safety evaluation”. In: *Proceedings of the workshop on Human-centered Evaluation and Auditing of Language Models (HEAL) at CHI*. Vol. 2024. 2024.
- [399] Jianyi Wang, Kelvin CK Chan, and Chen Change Loy. “Exploring clip for assessing the look and feel of images”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 37. 2. 2023, pp. 2555–2563.

- [400] Kelly Wang et al. “Counting Carrds: Investigating Personal Disclosure and Boundary Management in Transformative Fandom”. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–13.
- [401] Sitong Wang et al. “PodReels: Human-AI Co-Creation of Video Podcast Teasers”. In: *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2024, pp. 958–974.
- [402] Bryan Washington. *All The Ways To Be with Bryan Washington* Ocean Vuong. Dec. 2020. URL: <https://a24films.com/notes/2020/12/all-the-ways-to-be-with-bryan-washington-ocean-vuong>.
- [403] John Waters. “Shock Value: A Tasteful Book about Bad Taste”. In: *Delta* 11 (1981).
- [404] Sabine Weber et al. “Queer NLP: A Critical Survey on Literature Gaps, Biases and Trends”. In: *arXiv preprint arXiv:2602.16151* (2026).
- [405] Joel Wester et al. ““As an AI language model, I cannot”: Investigating LLM Denials of User Requests”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–14.
- [406] David Gray Widder, Meredith Whittaker, and Sarah Myers West. “Why ‘open’ AI systems are actually closed, and why this matters”. In: *Nature* 635.8040 (2024), pp. 827–833.
- [407] Chrisopher Wiggins. *Former DOGE staffer says ChatGPT helped feds cancel grants mentioning ‘LGBTQ+’*. Mar. 2026. URL: <https://www.advocate.com/politics/national/doge-chatgpt-dei-lgbtq-grants>.
- [408] Christopher Wiggins. *Meet the gay minds behind the RuPublicans*. May 2023. URL: <https://www.advocate.com/drag/rupublicans-creator-instagram#rebelltitem1>.
- [409] Jacob O Wobbrock, Brad A Myers, and John A Kembel. “EdgeWrite: a stylus-based text entry method designed for high accuracy and stability of motion”. In: *Proceedings of the 16th annual ACM symposium on User interface software and technology*. 2003, pp. 61–70.
- [410] Robert Wolfe and Aylin Caliskan. “American== white in multimodal language-and-image ai”. In: *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. 2022, pp. 800–812.
- [411] Sally Wyatt. “Non-users also matter: The construction of users and non-users of the Internet”. In: *How users matter: The co-construction of users and technology* (2003), pp. 67–79.
- [412] Sijia Xiao, Coye Cheshire, and Amy Bruckman. “Sensemaking and the Chemtrail Conspiracy on the Internet: Insights from Believers and Ex-believers”. In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2 (2021), pp. 1–28.
- [413] Angela Yang. *Pope Leo XIV says advancement of AI played a factor in his papal name selection*. May 2025. URL: <https://www.nbcnews.com/world/the-vatican/pope-leo-ai-played-factor-papal-name-rcna206126>.
- [414] Qian Yang et al. “The Future of HCI-Policy Collaboration”. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–15.

- [415] Akhila Yerukola et al. “Mind the Gesture: Evaluating AI Sensitivity to Culturally Offensive Non-Verbal Gestures”. In: *ACL*. 2025. URL: <https://arxiv.org/abs/2502.17710>.
- [416] Meg Young et al. “Participation versus scale: Tensions in the practical demands on participatory AI”. In: *First Monday* (2024).
- [417] Haozhuo Zhang et al. “Hand1000: Generating realistic hands from text with only 1,000 images”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 39. 9. 2025, pp. 9905–9913.
- [418] Haonan Zhong et al. “Copyright protection and accountability of generative ai: Attack, watermarking and attribution”. In: *Companion Proceedings of the ACM Web Conference 2023*. 2023, pp. 94–98.
- [419] Xuhui Zhou et al. “Cobra frames: Contextual reasoning about effects and harms of offensive statements”. In: *arXiv preprint arXiv:2306.01985* (2023).
- [420] Wenqi Zhu et al. “Challenges and Opportunities for the Design of Inclusive Social Media Experiences with LGBT+ Older Adults”. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 2025, pp. 1–16.
- [421] Ben Zimmer and Kelly E. Wright. *American Dialect Society Selects Slop as 2025 Word of the Year*. 2026. URL: <https://americandialect.org/wp-content/uploads/2026/01/2025-Word-of-the-Year-PRESS-RELEASE.pdf>.
- [422] Michael Zimmer. ““But the data is already public”: on the ethics of research in Facebook”. In: *Ethics and information technology* 12 (2010), pp. 313–325.
- [423] Shoshana Zuboff. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. 2018.
- [424] Douglas Zytke, Jane Im, and Jonathan Zong. “Consent: A research and design lens for human-computer interaction”. In: *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*. 2022, pp. 205–208.