ORIGINAL RESEARCH

How AI hype impacts the LGBTQ + community

Dawn McAra-Hunter[1,2]

Abstract

Hype around Artificial Intelligence (AI) has been a feature of this technology since its inception. However, the most recent wave of AI hype has been leveraged to encourage adoption of AI technologies that cause issues for marginalised communities. Hype is also a means to obfuscate real issues of bias, harm, and exploitation felt most sharply by marginalised communities when AI is implemented. This therefore raises the question of power imbalances as a feature of AI technologies as we currently know them. This paper will study the relationship of AI hype and marginalised communities, with particular emphasis on the LGBTQ + community, and look at the way that AI impacts on this community. This paper will pose two key questions: does hype affect marginalised communities, particularly hype around new technologies such as AI; and what impact does the LGBTQ + community experience as a result of hype. This paper will then move on to discuss areas that provide a focus for discourse of AI hype and the impact on the LGBTQ + community: policy and decision-making, the maintenance of the cisgender heteronormative (cishet) baseline, the ubiquity of a mythology of AI, and the role of market expansion.

Keywords Artificial intelligence · LGBTQ+ · Queer theory · Hype · Marginalisation

1 Introduction

In recent years, Artificial Intelligence (AI) has experienced a notable resurgence in both academic and commercial circles. This renewed interest has positioned AI as an integral element in the landscape of digital transformation, with the potential to revolutionise industries moving forward. However, the capabilities of AI are frequently overstated [1], and communications and general discourse around AI are often plagued by hype [2].

AI hype has been a concern of scholars, activities, and practitioners since the inception of artificial intelligence, as the exaggerated promises and inflated expectations of AI technologies can perpetuate harmful stereotypes and exclusionary practices against marginalised communities [3]. However, the advent of machine learning technologies and the advancements in generative AI have fuelled a new wave of hype around AI [4] and with that, increased risks of harm for marginalised groups.

There are many ways in which AI hype can be harmful to marginalised communities. This is particularly prevalent in exaggerating the capabilities of AI systems, misleadingly presenting them as highly accurate or infallible solutions to social problems, or downplaying their potential biases. Marginalised communities may then have negative experiences with AI through the reinforcement of stereotypes and biases [5], or from the harms that can be perpetrated by marginalised people not being part of the data that AI systems are trained on [6].

Hype is also an influencing factor in policy and decision-makers' understanding of AI capabilities, and subsequent AI implementations. The role of hype in policymakers' and key decision-makers' understandings of the capabilities, risks, and limitations of artificial intelligence has been addressed by key figures in the AI ecosystem, such as Zachary Lipton who stated in 2018 "policymakers don't read the scientific literature but they do read the clickbait that goes around" [2]. According to Lipton (2018), the media industry is partly responsible for this issue as it fails to effectively discern between genuine advancements in the field and promotional material [2].

Hype is also a means to obfuscate real issues of bias, harm, and exploitation felt most sharply by marginalised communities when AI is implemented. This therefore raises

✉ Dawn McAra-Hunter
dawn.hunter@sbe.uws.ac.uk

1 School of Business and Creative Industries, University of the West of Scotland, Paisley, UK

2 Diverse AI, Edinburgh, UK

the question of power imbalances as a feature of AI technologies as we currently know them. This paper will study the relationship of AI hype and marginalised communities, particularly the LGBTQ+ community, and the role which marketing communications plays in enhancing this hype and its impacts on the LGBTQ+ community.

Section 2 of this paper will discuss marginalisation and its origins. Section 3 will examine some of the key theoretical underpinnings of this paper. Section 4 looks specifically at power and AI, and some of the real-world impacts experienced by marginalised communities. Section 5 examines how hype impacts on the LGBTQ+ communities experience of AI. Section 6 then moves on to look exclusively at the Queer experience of AI. The paper then moves on to discuss these elements in Section 7, before offering recommendations for future research in Section 8.

This paper will pose two key questions: does hype affect marginalised communities, particularly hype around new technologies such as AI; and what impacts does the LGBTQ+ community experience as a result of hype. This paper will then move on to discuss areas that provide a focus for discourse of AI hype and the impact on the LGBTQ+ community: policy and decision-making, the maintenance of the cisgender heteronormative (cishet) baseline, the ubiquity of a mythology of AI, and the role of market expansion.

2 Marginalisation and its origins

2.1 Marginalisation and marginalised communities

Marginalised communities as a concept refer to groups of people who experience social, economic, and/or political disadvantages or exclusion due to factors such as their race, ethnicity, gender identity, sexual orientation, disability, or socioeconomic status. More expansive definitions extend to marginalised communities as those in communities considered to be outwith mainstream society. The concept of marginality can be traced back to Robert Park of the Chicago School of Sociology, who defined it as the position of individuals or groups in society that is characterised by a lack of power, and limited access to resources [7].

While social marginalisation can be experienced on an individual level, such as in the cases of single parenthood [8], unhoused people [9], or barriers faced by disabled people in the workforce [10], marginalisation occurs on a larger societal scale, where entire communities are at risk of marginalisation on account of systemic discrimination and prejudice based on their identities [10, 11].

Contributing factors to marginalisation are numerous; however, certain key elements arguably contribute to marginalisation on a global scale, such as gender [12], race and ethnicity [13], globalisation [14], and socioeconomic inequality [15,

16]. These factors intersect and interact, giving rise to intricate systems of marginalisation that impact individuals within the LGBTQ+ community. The intersectionality of different marginalising factors is also a key contributor of further marginalisation [17].

2.2 Marginalisation and the LGBTQ+ community

The LGBTQ+ community refers to individuals who identify as lesbian, gay, bisexual, transgender, queer, and other sexual and gender minorities. People within the LGBTQ+ community often face systemic discrimination and prejudice based on their sexual orientation and gender identity, leading to various socioeconomic and healthcare disparities compared to the general population [18]. People within the LGBTQ+ community may also experience marginalisation from other groups within the LGBTQ+ community. For example, trans-individuals or people of colour within the LGBTQ+ community may experience transphobia or racism from other group members [19]. It is therefore important to consider the ways in which the LGBTQ+ community is marginalised not only as a whole, but also how specific subgroups within the LGBTQ+ community may face intersecting forms of marginalisation.

A major concern regarding the expanded use of AI is its potential to unintentionally reinforce stereotypes around gender, which can hinder progress toward gender equality [20]. As AI algorithms primarily learn from vast amounts of data, the biases locked within these systems can be perpetuated and reinforced through unmonitored implementation. However, research specifically focusing on the LGBTQ experience as a result of marketing driven early adoption is limited.

The LGBTQ+ community has encountered numerous challenges when it comes to the integration of their queer identities with artificial intelligence. These issues stem from the fact that AI models often learn from data that reflect social biases, leading to instances of discrimination against transgender individuals on dating websites and misgendering by generative AI systems. Furthermore, algorithmic bias within healthcare systems perpetuates negative impacts on the LGBTQ+ community, undermining progress made in addressing bias for other marginalised groups. Over the past 2 decades, as AI technology has advanced, its interactions with the LGBTQ+ community have exhibited various harmful or unfavourable aspects.

3 Theoretical background and considerations

3.1 Queer theory

Much of the marginalisation of the LGBTQ+ community comes from the acceptance of heterosexuality, and the

characteristics and values that entail, as the dominant paradigm for understanding gender and sexuality. Queer theory challenges the dominant paradigm of using heterosexuality as the standard for comparison [21]. Coined by Teresa de Lauretis in 1991 [22], there are at least three interconnected concepts in queer theory [23]:

> Disregarding heterosexuality as the default state for sexuality.
> Challenging the notion that lesbian and gay studies comprise a singular area of study.
> Acknowledgement and subsequent focus on the interconnection of racism and sexism.

Queer theory offers the potential to encompass these various critiques and enable a reevaluation of sexuality and gender beyond traditional norms. In the context of AI hype and the experience of marginalised communities, queer and feminist theoretical perspectives can offer valuable insights into the experiences of marginalised communities and the potential biases embedded within AI.

However, it is important to note that Queer theory did not emerge in isolation; it has its roots in critical and feminist theories. These theories aimed to deconstruct dominant narratives and examine power imbalances within society. One influential figure in this field is Michel Foucault, whose analysis of power dynamics and exploration of the history of sexuality remain highly relevant [24, 25]. In examining the relationship between key concepts like heteronormativity and cisnormativity in Queer theory and mechanisms employed by AI systems, we can draw upon Foucault's insights regarding taxonomies, classifications, and power structures that contribute to the oppression faced by LGBTQ+ communities.

According to Foucault, subjectivity is not an isolated concept in philosophy; rather, it is shaped and influenced by knowledge and power. In his later work, he delved into examining the interplay between knowledge and power.

His work represents a significant departure from previous conceptions of power and cannot be easily assimilated into existing frameworks. Power is depicted as diffuse rather than concentrated, embodied, and enacted instead of possessed, discursive rather than solely coercive, and forming agents rather than being wielded by them [26].

3.2 The cishet baseline

The key concept in queer theory is that of *heteronormativity*: "the institutions, structures of understanding, and practical orientations that make heterosexuality seem not only coherent—that is, organised as a sexuality—but also privileged" [27].

Heteronormativity refers to the assumption that heterosexuality is the norm and all other forms of sexuality are deviations or abnormal. This assumption permeates societal institutions, cultural norms, and individual beliefs, creating a hierarchical system that marginalises and oppresses non-heterosexual identities.

Similarly, *cisnormativity* is the assumption that cisgenderism is the norm, and that everyone is (or ought to be) cisgender [28]. Originated in 2009 as the "the expectation that all people are cissexual" [29], it has been noted that "Cisnormative assumptions are so prevalent that they are difficult at first to even recognize", and "cisnormativity shapes social activity such as child rearing, the policies and practices of individuals and institutions, and the organisation of the broader social world" [29].

This "cishet baseline" has numerous implications for the LGBTQ+ community, both in terms of othering the lived experiences of LGBTQ+ people, and in terms of actively discriminating against the LGBTQ+ community or specific member groups within the community.

In Western Europe and North America, there has been a notable increase in the normalising of non-heterosexuality within institutional arrangements. Same-sex marriage and adoption have become legally recognised in countries, such as Canada, the USA, UK, and many parts of Western Europe. Furthermore, equalities' legislation prohibits employers from discriminating against individuals on the basis of their sexual orientation or gender identity to a certain extent (although this is narrowly defined in UK legislation). While these advancements may appear positive at first glance, it is important to recognise that rights enjoyed by queer individuals across these regions still face ongoing challenges and potential threats similar to those observed with abortion rights.

The recent repeal of Roe v Wade in the United States has created opportunities to challenge rulings on marriage equality and laws concerning private sexual activities. This shift in legal precedent is reflected in the introduction of bills like the "Don't Say Gay" bill, reminiscent of UK's s.28, which prohibits any discussion or promotion of gender and sexuality (currently enacted only in Florida but with plans for implementation in 11 other states). In December 2023, the UK government Department for Education released its guidance for a "parent first approach" to transgender and non-binary children in school [30], and guidance which has been widely criticised as transphobic and in violation of children's privacy [31]. Furthermore, the Eastern European queer community faces increasing discrimination and social isolation from the anti-LGBTQ+ political ideologies currently asserting dominance in nations, such as Poland and Hungary [32, 33], in stark juxtaposition to the more permissive political landscape of the 1960s and 70s and the social values of many citizens [34].

4 Power and AI

4.1 Defining power

Power itself can be understood either as structural or poststructural. Structural power refers to systemic bodies of power which are encoded in social and institutional structures, shaping relationships, norms, and values within society. However, defining structural power has been a subject of debate and varies among different scholars [35–37]. Poststructural power is concerned on how power functions through discourse and language, shaping subjectivities, and identities [38].

Power is according to Weedon (1987)

a dynamic of control and lack of control between discourses and the subjects, constituted by discourses, who are their agents. Power is exercised within discourses in the ways in which they constitute and govern individual subjects [39].

Furthermore, systems of power are not solely determined by individual actions but rather by the existence of power itself. Power permeates society and is present in countless everyday situations that involve various issues. The combined impact of these situations leads to the establishment of specific power structures [40]. Additionally, individuals themselves are shaped by both external and internal constraints imposed by these power structures. External controls restrict certain identities, particularly through labelling numerous bodily desires as unacceptable.

Power is not limited to specific individuals or rigid structures, but is present in every aspect of society without any singular source or fixed form [41]. Every society has its system of truth, known as the "general politics" of truth [42]. This refers to the specific kinds of discourse that are accepted and treated as true within a particular society. It also includes the mechanisms and institutions that enable individuals to distinguish between true and false statements, as well as how these statements are validated.

4.2 Real-world implications for marginalised communities

The hopeful optimism that AI will assist in overcoming biases in human decision-making has been challenged by instances of bias and unfairness against marginalised communities [43]. As asserted by Abeba Birhane (2022):

"Let's ditch the common narrative that AI is a tool that promotes and enhances human 'prosperity' (whatever that means) & start with the assumption that AI is a tool that exacerbates inequality & injustice & harms the most marginalised unless people actively make sure it doesn't." [44]

Women and people of colour (and particularly women of colour) experience real-world implications of inherent power imbalances encoded in AI systems trained on biased data and created in biased and unbalanced conditions. For example, facial recognition technology has demonstrated racial and gender bias in data sets, leading to the misclassification of women and people of colour [45–47]. Disparities in facial recognition classification accuracy are also greater between light skinned people and dark skinned people, with inaccuracies rising sharply for dark skinned people [45]. The existence of biases in areas like law enforcement, where facial recognition technology is employed for identification purposes, can result in significant negative outcomes with potentially devastating consequences [48–50].

Structural inequalities in access to healthcare are often replicated in AI-enabled healthcare systems [51–53]. These systems have been found to exhibit biases in diagnosis and treatment recommendations, leading to disparities in healthcare outcomes for marginalised communities.

The technocentric discourse which centres fast and ubiquitous implementation of new technologies yields numerous demonstrations of systemic power leveraged against marginalised communities [54]. This can be seen across key aspects of social administration, as demonstrated in Table 1.

5 How hype impacts the LGBTQ + AI experience

5.1 Definitions and understanding of hype

Hype as a concept has a variety of different definitions, ranging from deception or fraud, to excitability [71]. However, for the purposes of this paper, the definition of hype that will be utilised is the use of media, marketing, and promotional channels to elicit interest in a product or service [71]. Where this intersects with new technologies, such as AI, this is often achieved on the basis of overinflated claims of capabilities, although this is not always the case [72].

Hype is often a catalyst in the implementation and adoption of emergent technologies [73]. This has been seen in previously emergent technologies of the modern era, such as the Internet [74–76], Big data [77–79], the Internet of Things [80–82], and Blockchain [83–85].

Perhaps, the most prevalent model of emergent technology hype is the Gartner Hype Cycle, which provides a visual representation of the hype surrounding various technologies over time [86]. The Gartner Hype Cycle is a graphical representation that illustrates the evolution of technologies, their adoption rates, and social impact over time. It consists of five

Table 1 Technology implementation impact on marginalised communities across social categories

	Policing and justice	Border control	Healthcare	Social welfare	Financial services	Employment
Technology	Facial recognition	Facial recognition and deception detection	Patient prioritisation algorithm	Welfare benefits fraud detection algorithm	Credit scoring algorithm	Data classification schemes
Impact	Misclassification of women and people of colour [45, 55]	Misclassification of statements as deception by marginalised people [56, 57]	Prioritisation of white people for additional healthcare services [58, 59]	Heightened surveillance and high level of false positives for lower socio-economic communities [60, 61]	Use of statistical models deemed to violate anti-discrimination laws [62]	Misclassification of protected characteristics in interviewing, hiring and promotion decision-making [63]
Technology	Predictive policing	Biometrics	Melanoma detection algorithm	Digitalisation of community care operations	Credit scoring algorithm	Video interview systems
Impact	Overrepresentation of lower socioeconomic areas in data [64, 65]	Reproduction of historic bias across different systems. [66]	Underrepresentation of people of colour decreasing accuracy [67]	Decreased autonomy and decision-making power for female workers and patients [68]	Classification bias of people of colour leads to discriminatory impacts across financial services [69]	Centring of dominant characteristics as normative negatively impacting disabled people [70]

distinct phases that reflect various stages in the technology adoption cycle [87].

The concept of the Gartner Hype Cycle has been the subject of numerous criticisms [88–90], mainly due to its subjective nature and lack of scientific rigour. A few technologies have been shown to travel through an identifiable hype cycle, and the model has been described as more of a conceptual framework rather than a precise predictive tool [91].

However, despite these criticisms, the hype cycle framework has been widely used to understand the adoption and maturity of emerging technologies, including in the field of artificial intelligence.

As a specific marketing approach, hype involves utilising exaggerated measures of publicity to generate excitement and anticipation for a product or service [92]. This modern practice is closely linked with social media marketing, particularly through influencer and viral marketing [93]. Consumers targeted by hype marketing may engage in hype-generating activities around the product in question as a demonstrator of conspicuous consumption and as a means to signal their affiliation with a particular brand or trend, and the personal characteristics or social capital that this signifies [93].

Hype in itself is not necessarily a bad thing. As noted by Milne (2020):

"Hype, like any tool, isn't inherently good or bad. It can be the tool with which we gather communities around positive change, and it can be the tool that misleads to satisfy the ill-conceived wants of a few immoral actors. Sometimes people don't even know they are propagating it. But when hype starts to grow unchecked, it doesn't really matter who started it or why; what matters is that it is spotted before any damage happens" [94].

5.2 The evolution of AI hype

Throughout the history of artificial intelligence, there has been a persistent pattern of inflated expectations and grandiose claims regarding its transformative potential. The field has consistently experienced periods of hype, with exaggerated promises surrounding AI's capabilities.

Artificial Intelligence encompasses various technologies, such as machine learning, natural language processing, natural language generation, deep learning, and neural networks. While the term "Artificial Intelligence" serves as a broad concept that simplifies complex technological processes for lay audiences, computer scientists and developers are starting to view it as a marketing hype term [95–97]. This is due to its continuous use to mask the true capabilities of different technologies behind an illusion of a singular magical technology.

AI and Ethics

Springer

Hype has been a significant aspect of artificial intelligence research and development since the 1950s. While there have been notable advancements in the field of artificial intelligence in recent years, much of this progress can be attributed to the availability of Big Data and increased computing power rather than substantial strides in what is commonly understood as "intelligence" by the general public [98–100].

The effect of such hype has led to a perception from some commentators that "*AI-powered' is tech's meaningless equivalent of 'all natural'*" [101].

According to a report by Slate, an analysis of press releases and technology articles dating back to the 1990s reveals a recurring pattern: predictions about technological advancements, especially those related to artificial intelligence, consistently project developments that are 5–10 years away [102].

The report compiled a list of 81 such predictions to illustrate this common cliché. These inflated expectations and overpromises contribute to the hype surrounding AI, creating a sense of anticipation and excitement among both industry professionals and the general public.

5.3 Elements of AI hype

There are several elements of AI as a concept that are simultaneously hyped up while driving further hype towards AI. This can be seen in anthropomorphisation and perceived objectivity of AI systems.

In the case of anthropomorphism, it has been observed that to facilitate customer–robot interactions, humanlike service robots may be preferred to increase customers' perceptions of social presence [103]. Equally, this can be seen in the human mimicry of chatbots, which can often convince users that their interactions have been with another human actor [104]. There is a growing consensus, noted by Novak and Hoffman (2019) that anthropomorphism is an important tool in understanding how customers experience interactions with inanimate objects [105]. According to Epley et al., this perception results from "the attribution of human characteristics or traits to nonhuman agents" [106].

Anthropomorphism has been found to increase product and brand liking [107], although whether anthropomorphism in service robots enhances customers' experiences is unclear. It has been argued that humanlike qualities "incorporates the underlying principles and expectations people use in social settings in order to fine-tune the social robot's interaction with humans" [107]. However, there is also the argument that anthropomorphism is less positive: "consumers will experience discomfort—specifically, feelings of eeriness and a threat to their human identity" [107]. This is also known as the "uncanny valley" effect.

However, Troshani et al. have purported that enhancing the humanness of an AI application is likely to amplify the human user's perception of goodness of an AI application, and consequently the extent to which it can be trusted. The further posit that humanity in AI applications in service can improve trust of consumers in these applications which can, in turn, facilitate relationship building between consumers and service providers [108].

Hype also drives and reinforces the idea of perceived objectivity that underpins AI technologies. Datasets on which AI systems are trained, and subsequently analysed, often reflect inequities that occur in the world at large [109, 110]. However, the highly technical nature of data-driven AI systems often provides a rhetoric of objectivity which veils the complicated and much more fallible systems underneath [111].

These "appeals to objectivity" are embedded in technological discourses and practices [112]. This notion of objectivity increases the difficulty with which to challenge this fundamentally misleading dichotomy and to demand accountability [113].

When defining Big Data, on which artificially intelligent systems are developed, Boyd & Crawford (2012) define it as a cultural, technological, and scholarly phenomenon that intersects with technology, analysis, and mythology [114]. The concept of *mythology* offers a foundation for appeals to objectivity in perpetuating the belief that data offers a "higher form of intelligence and knowledge that can generate insights that were previously impossible with the aura of truth, objectivity and accuracy" [114]. These claims to objectivity in Big Data, the information on which artificial intelligence is fed, are fundamentally misleading.

It is argued by Gitelman (2013) that an interpretative process of the imagination is shaped by the norms and standards for every discipline and disciplinary institution and their own perception or "imagination" of data [115]. Boyd & Crawford note thusly "As computational scientists have started engaging in acts of social science, there is a tendency to claim their work as the business of facts and not interpretation. A model may be mathematically sound, an experiment may seem valid, but as soon as a researcher seeks to understand what it means, the process of interpretation has begun. This is not to say that all interpretations are created equal, but rather that not all numbers are neutral" [114].

Closely related to perceived objectivity is the fallacy of inscrutability [116]. This fallacy of inscrutability is a category error: when critics argue that the actions of a system cannot be comprehended, they are attributing values to mechanical technologies rather than to the humans who created and implemented them [116]. The fallacy of inscrutability is highlighted as one of 18 key issues with AI journalism which contributes to AI hype set out by Kapoor and Narayanan [117]. This can be seen in media which claims

that it is impossible to understand how models work, and as such they cannot be used in a non-discriminatory way.

The remainder of the 18 common pitfalls most often seen in AI journalism, include utilising flawed human-AI comparisons, hyperbolic, incorrect, or non-falsifiable claims about AI, uncritically platforming those with self-interest, and failure to address limitations [117]. These common issues with media representation of AI contribute to the perpetuation of unrealistic expectations and the culture of hype surrounding AI technologies.

This is compounded by the role of technology developers and other private interests in driving AI hype. Technology firms have a strong motivation to keep information concealed. Some may aim to protect the confidentiality of their intellectual property, while others seek to capitalise on the allure of "AI" without truly engaging in AI itself [118]. Many software developments might employ quite ordinary statistical methods that do not reflect true artificial intelligence. Consequently, it is not advantageous for a company to disclose how basic their technology actually is.

The role of tech companies is further compounded by the media. A key pitfall of AI journalism according to Kapoor and Narayanan includes the platforming of self-interested parties without critique. This can be seen by the media treating company spokespeople or sources as though they are neutral sources, repeating PR terms rather than describing how an AI tool works [117]. This uncritical platforming allows corporate interests to control the narrative and perpetuate the hype surrounding their AI technologies, without providing a balanced and factual representation of their limitations or potential risks [117]. This lack of critical analysis in AI journalism contributes to the formation of false impressions and unrealistic expectations about AI capabilities.

All of this hype has real-world consequences which both directly and indirectly harm marginalised communities such as the LGBTQ+ community.

Hype drives the early adoption of new tech, even when there is little evidence to support its effectiveness or usefulness [119], and when there are potentially negative impacts for marginalised communities [120]. The rush to adopt AI technologies without fully understanding their capabilities and limitations can lead to the creation and perpetuation of biased and discriminatory algorithms [121].

The fear of being left behind and the rush to early adoption also drives the traffic in "fake" AI [122]. For example, in 2019, venture capital firm MMC found that out of 2,830 startups classified as AI companies, only 1580 actually met the criteria [123].

Related to this is also the way in which hype impacts AI implementation decisions [124]. Where executions of AI have resulted in adverse social impacts, implementation decision-makers ultimately perceive the technology to be impartial, and results generated found to be fair and correct even in the presence of biased or poorly structured data [124].

There is also the notion that the failure of previous hyped technologies such as blockchain or cryptocurrency has led to a desire for AI to be the "golden solution" that will solve all economic, social, and political woes, encouraging an attitude of "all bets are on AI" [125]. This plays into a discourse of inevitability, whereby AI implementation is presented as a necessity. In a horizon scan of discourses on artificial intelligence and development in education, Nemorin (2021) states:

"This view of AI in education rests on the assumption that no space in the human body is sacred enough to be protected from the creep of AI's attention. This social imaginary suggests that every aspect of bare life is and should be thrown open for measurement and behavioural management..." [126].

6 The queer AI hype experience

6.1 The LGBTQ+ historical perspective: engagement with technology and media

The LGBTQ+ community has had a complex relationship with AI since its inception, as seen in the transgender politics of Alan Turing's original Turing Test [127]. The presence of other LGBTQ+ individuals among prominent AI pioneers further emphasises this intersection between AI and the LGBTQ+ community on a human level. However, it is also important to consider the ethical implications and potential biases that can arise when integrating AI algorithms into societal frameworks. Christopher Strachey, the creator of C programming language and considered one of the pioneers in computer-generated art, faced personal challenges regarding his sexuality while working within the restrictive environment of British academia during the 1960s [128]. Peter Landin, a prominent figure in computer science who recognised the mathematical expression behind programming languages, later expressed regret for his involvement with this field due to its growing utilisation in state surveillance activities [129].

The LGBTQ+ community faces various forms of harm in the digital realm, even outside the scope of AI technologies. In the United Kingdom, legislation such as The Protection of Freedoms Act 2012 [130] and the Policing and Crime Act 2017 (known as the 'Turing Law') [131] aimed to address this issue by allowing gay men with historic cautions or convictions for certain offences to have them disregarded or pardoned. However, due to inadequate consideration for a wide range of different historical crimes stored digitally, many individuals ended up having these convictions included in Disclosure and Barring Service checks alongside serious

sex offences, terror acts, and crimes like murder. As a result, their careers suffered significant damage [132]. In 2023, the programme was revised to cover a broader range of crimes eligible for pardon. For the first time, this now includes pardons for women convicted of any past same-sex activity offences that have since been repealed or abolished [133].

6.2 The modern experience: a history of the present

With the advancement of AI technology, its impact on the LGBTQ + community has evolved. Unfortunately, this evolution has often resulted in harmful effects. A review of some key harmful impacts of the implementation of AI systems on the LGBTQ + community can be seen across a variety of social factors in Table 2.

One notable instance occurred in 2017 when a Stanford study claimed that facial images could be used by artificial intelligence to determine sexual orientation [134]. This claim was swiftly criticised by advocacy groups as "junk science" [135] and sparked debates regarding privacy, ethical boundaries, and potential misuse of AI technology. Metcalf (2017) argues that part of the issue with this paper lies at the intersection of research ethics and research hype [136]. Where academic ethics boards come up against data science research, they are often ill equipped to deal with the outcome, leaving people at the risk of harm in the rush to be the next "scientific gaydar" claim [136]. The questionable and centuries old search for positive physical identifiers of sexuality is in itself a product of perceived objectivity hype (albeit for science more generally rather than AI specifically) [137].

Automated gender recognition refers to a specific application of facial recognition technology that uses AI algorithms to identify the gender of individuals based on photographs and videos. However, it is important to acknowledge that AGR models hold outdated and potentially harmful assumptions about gender presentation, particularly for transgender and non-binary communities [138]. Studies indicate that AGR technologies reinforce the existing biases against marginalised groups, including trans-, non-binary, gender nonconforming individuals [139] as well as people with darker skin tones who belong to racial minority groups [140]. These examples highlight the discriminatory outcomes and representational harms that can arise from data-intensive practices and AI systems of surveillance and social sorting [141]. The development and proliferation of AGR technologies can be seen to be a direct reaction to hype. Similarly to the aforementioned appeals to scientific objectivity found in "scientific gaydar" [136, 137], the variations of means by which to classify gender are numerous [142]. AGR systems also fall prey to the fallacy of inscrutability [116], the commercial computer services almost always being proprietary "black box" systems [142].

Utilisation of AI content moderation has led to the LGBTQ+ community experiencing several issues. Social media platforms like YouTube and TikTok have come under scrutiny for their alleged discriminatory practices towards the LGBTQ+ community. In 2019, YouTube faced a class action lawsuit that accused its content moderation algorithm of falsely identifying videos with keywords related to "lesbian," "transgender," and "gay" as adult content and restricting access to them through the use of "restricted mode" [143]. Similarly, TikTok has been accused of engaging in various anti-LGBTQ+ activities through its algorithms, including limiting exposure to LGBTQ+ hashtags and suppressing disabled, plus-sized, and LGBTQ+ creators' content [144, 145]. Despite frequent and regular reports about the failures of automated content moderation, whether it "is neither reliable nor effective" [146], "it might not work" [147], or it is still reliant on human labour to the extent that a former content moderator filed suit against TikTok alleging failure to provide adequate safeguards for moderator mental health after she developed PTSD [148], AI continues to be hyped as a panacea for content moderation by industry players such as OpenAI [149] and the UK Government [150].

This is in spite of evidence that even highly efficient moderation systems could exacerbate, instead of improve, numerous current content policy issues on platforms by potentially further enhancing opacity, introducing complexity to existing problems related to fairness and justice in large-scale sociotechnical systems; and concealing the inherently political nature of speech decisions being made at scale [151].

These examples represent just a small fraction of the reported cases of algorithmic bias targeting the LGBTQ+ community on various social media platforms. It is evident that marginalised groups face stricter content moderation regulations and are subject to disproportionate account suspensions, especially when their content challenges the dominant group [152].

The LGBTQ+ community has also faced the algorithmic promotion of content that is harmful and discriminatory. TikTok, for instance, has been accused of actively promoting homophobic and anti-LGBTQ+ content to its users [153]. The cyclical nature of misinformation and the viral spread of content on social media platforms, combined with intentional promotion to boost user engagement, have played a significant role in amplifying narratives such as the "groomer" accusation against LGBTQ+ individuals. This type of narrative has garnered substantial attention on social media platforms, with the top 500 influential tweets containing hateful 'grooming' allegations being viewed over 72 million times [154].

Table 2 A review of AI implementation impacts on LGBTQ+ community

	Impact			
Safety and privacy	Logistical regression model claims a high reported accuracy of self-reported sexual orientation, leading to claims of AI "gaydar" [134]	Automated gender recognition exacerbates existing biases and opens trans- and non-binary people up to harm [136, 138]	Using video gender recognition systems on online video content, researchers attempted to ascertain hormone replacement treatment in trans content creators [155, 156]	Online content algorithms promote the proliferation of harmful content about LGBTQ people, particularly in amplification of the "groomer" narrative [153, 154]
Policing, security, and border control	Deception detection software incapable of detecting trauma, putting women and LGBTQ people at greater risk [56, 57]	Border facial recognition software programmed to recognise gender binary placing trans-people at greater risk [157]	Use of Google Maps and geolocation data has been used to create maps of trans groups and trans healthcare providers as targets of harm [158]	Egyptian police monitor social media, dating and geolocation apps to track and arrest LGBTQ people [159]
Healthcare	Due to a lack of sexual orientation and gender identity data in medical records, LGBTQ+ people are underrepresented in health data, impacting on patient outcomes [160]	Currently implemented data sets in AI healthcare are limited by binary gender fields and lack of gender disaggregation, making bias testing difficult [160]	Binary gender classification in healthcare data causes harm for the transgender, non-binary and intersex communities, exacerbating health disparities [161]	Using machine learning models to approximate gender identity and sexuality from health records introduces risks of forced outing of patients [162]
Cultural	Online socialising technologies reinforce gender binary causing marginalisation and barriers to services for LGBTQ people [163]	Machine learning model is used to predict sexual orientation based on Facebook profiles, perpetuating stereotypes about how LGBTQ+ people express themselves [164]	Automated content moderation systems censor LGBTQ content at a disproportionate level to non-LGBTQ content [143–145]	Social networking app aimed at women and girls utilises "biometric gender verification software" to exclude trans women from using the site [165]
Employment	Utilisation of Microsoft Real ID facial recognition system by Uber resulted in a registered driver, a transgender woman, being locked out of the system [166]	The use of facial recognition systems in recruitment applications are inherently biased putting LGBTQ+ job applicants at risk of unintentional discrimination [167]		

7 Discussion

7.1 Introduction

This paper seeks to investigate the ways in which AI hype impacts the LGBTQ+ community. As a means of exploring this proposition, hype as a concept both in AI and in marketing communications more generally was examined, along with looking at the various ways in which AI itself can cause harm for marginalised communities, in particular the LGBTQ+ community. This paper also looked into the ways in which marketing and communication strategies for AI contribute to and mirror systemic power dynamics, specifically concerning the LGBTQ+ community. The findings of this paper suggest that AI hype can indeed mirror and perpetuate the existing power structures related to LGBTQ+ identities.

The LGBTQ+ community experience of AI is fundamentally shaped by the biases inherent in AI systems, the subsequent preconceived means of interacting with LGBTQ+ individuals that AI systems can deliver. AI technologies are then able to manifest a variety of different harms on the LGBTQ+ community. The lack of diverse representation and inclusivity in developing and applying AI technologies further perpetuates these biases.

Problematic issues that arise from the intersection of AI and the LGBTQ+ community include the use of controversial and discriminatory AI technologies like facial recognition, deception detection, and predictive policing. Not only are these technologies shown to cause harm to the LGBTQ+ community as a whole, but those most at risk from the sharpest harms are the most disenfranchised of the community, namely trans-people, refugees and asylum seekers, and people of colour. It remains imperative to approach AI and the people impacted by its implementation through an intersectional lens.

This paper has also examined the way in which AI hype drives early adoption of AI technologies, often without sufficient consideration of their potential impact on marginalised communities.

This is particularly the case for AI systems that purport to predict social outcomes, such as crime prediction, child protection, and welfare benefits administration. The context in which these technologies are implemented can have significant consequences for the LGBTQ+ community. Given that many of these technologies are implemented without sufficient transparency, oversight, or accountability mechanisms, the potential risks to marginalised communities are amplified. The consequences of the AI hype cycle on the LGBTQ+ community therefore cannot be overlooked.

The findings of this paper point to a variety of different factors at play in driving hype and the adoption of AI technologies towards an ultimately harmful experience for the LGBTQ+ community. This discussion will now move to consider these factors and their implications in more detail, before making recommendations as to how to address these issues and create a better experience for LGBTQ+ people in which to interact with AI.

7.2 How AI hype actively harms the LGBTQ+ community

7.2.1 Influence on policy and decision-making

Artificial intelligence hype, particularly claims pertaining to perceived objectivity, infallibility, and techno-solutionism, is pervasive in many areas of decision-making. In particular, AI hype exerts an increasingly significant influence on key decision-makers and AI implementers, subsequently increasing the impact on marginalised communities. On the subject of navigating claims of artificial superintelligence, The Carnegie Endowment for International Peace commented that "leaders are far less equipped to evaluate claims made in a media and investment environment that incentivizes hype over level-headed assessment" [168].

The Observatory of Public Sector Innovation noted the impact that AI would have on government policy making and made recommendations as to policy frameworks to overcome AI hype in government decision-making [169]. However, policy decisions to implement AI systems with questionable and arguably overstated claims have seen AI implemented in European Union border control policy [170], the United Kingdom Defence Strategy [171], Dutch welfare benefits policy [172], as well as across a host of other public sector implementations [173]. Sadly, but somewhat unsurprisingly, the key drivers of AI implementation in the aforementioned policies are centred around claims of efficiency and objectivity, be that of tracing fraud [172], in warfare [171], or border security [170]. It is also with a somewhat bitter irony that protection of vulnerable people from harm is mentioned frequently as a benefit of AI implementation in the aforementioned policies, although without the assistance of any clarification as to how or why this might be achieved.

One does not need to look far to find key elements of hype, such as techno-solutionism and simplistic appeals to infallibility, in public policies. The introduction of the UK Government's Defence Artificial Intelligence Strategy, states boldly:

> "We also recognise that the use of AI in many contexts, and especially by the military, raises profound issues. We take these very seriously – but think for a moment about the number of AI-enabled devices you have at home and ask yourself whether we shouldn't

make use of the same technology to defend ourselves and our values" [171].

Overstatement of the effects of AI systems may also be seen as a means to secure funding or set funding agendas, in the public sector, in research, and in commercial enterprises [174]. Indeed, this is not only seen as a particular phase in the Gartner Hype Cycle as depicted in Appendix 1, but is also a defining factor in the concept of "AI Summers" and "AI Winters" whereby government funding of AI projects along with commercial investment go through periods of boom and bust, predicated largely on overstatement of AI capabilities thus leading to a reduced capacity for exploration and innovation in the field [175].

AI hype also drives the pervasive notion that it operates outwith human weaknesses such as bias [176]. This claim has actively impacted on policy decisions that affect marginalised people at great concentrations.

This claim is often used to promote and implement AI systems in various social contexts, such as recruitment, criminal justice, child protection, and welfare administration. However, these AI tools for social outcomes are frequently applied in situations where marginalised individuals and communities are excluded from the decision-making process yet are most likely to be affected by them [177].

AI has been used to predict outcomes in the fields where marginalised communities are at the greatest likelihood of intersecting with these technologies, in areas such as crime prevention [178], domestic violence [179], child welfare [180], and welfare benefit administration [172]. In 2022, British innovation agency Nesta noted that a key issue with utilising AI to predict social outcomes was a lack of robustness in AI models, leaving them generally poor at generalising outwith the narrow confines of which they had been trained, and favouring targeted interventions that could achieve unfair outcomes [181].

In the recruitment and human resources sectors, AI is often touted to business leaders as an effective way to reduce bias in the hiring process [182–184]. This is a popular misconception around AI that lingers in decision-making circles to this day despite evidence that AI at best reproduces race and gender bias in a similar way that humans do [185] largely on account of limited and biased datasets [186]. Despite these concerns, AI continues to be adopted in the recruitment sector, with recruiters focused largely on the efficiencies that AI brings over concerns around bias [187].

The impact that AI hype has on policy and decision making powers has harmed, and continues to harm, marginalised people, of which the LGBTQ+ community is a particularly vulnerable constituent.

7.2.2 Obfuscation and diverted priorities

Artificial intelligence hype also actively harms marginalised communities, particularly the LGBTQ+ community, through obfuscating or otherwise distracting from real issues of bias, harm, and exploitation felt most sharply by marginalised communities when AI is implemented. The persistent focus on superintelligence, or the existential threat of AI [188] effectively diverts public attention (and capital investment) away from real and routine matters of discrimination in housing [189], healthcare [12, 43, 51–53, 58], security, policing and criminal justice [48–50, 64], and the spread of hate speech and misinformation in non-English languages [190], all of which drastically impact on marginalised communities and where the LGBTQ+ community experiences some of the sharper harms.

In the instance of a social media platform such as TikTok, where the LGBTQ+ community is impacted by discriminatory content moderation policies as well as the increasing promotion of anti-LGBTQ+ content, the hype around the use of the platform itself (or the subsequent US ban), serves to overshadow the obscurity around the use of AI on the platform itself. In a case study examining the use of AI in TikTok and Facebook, Grandinetti (2021) states "the discursive promotional strategies of TikTok represent a hype cycle that obfuscates as much as it clarifies" [191].

7.2.3 Supporting the cishet baseline as the dominant discourse

Heteronormativity and cisnormativity continue to prevail in environments where the rights of queer individuals are undermined or taken away, as well as in spaces where their very existence is challenged by those who hold power. The politicisation and marginalisation of queer communities further reinforce cishet ideologies and uphold the belief that heterosexuality and cisgenderism are the norm, while any other sexual or gender identity is viewed as a deviation influenced by external factors.

This "cishet baseline" is arguably the main driver for the exclusion of the LGBTQ+ community in AI systems and their underlying data sets. This is particularly the case where models are trained on historical data, such as health records, but we also see its effects in other areas, such as social media algorithms and targeted advertising. These AI models are often built on data sets that reflect and reinforce heteronormative and cisnormative biases and assumptions, leading to inaccurate or discriminatory outcomes for the LGBTQ+ community.

To combat this issue, it is crucial to actively question and dismantle heteronormative and cisnormative prejudices in AI systems. It is vital to diversify data sources by integrating LGBTQ+ experiences to guarantee equitable and inclusive

AI algorithms. This involves incorporating diverse perspectives and experiences, specifically those of LGBTQ+ individuals, into data collection processes for the development of more accurate and inclusive AI systems.

7.2.4 Market expansion above human impact

Exaggerated claims have been a common occurrence in the field of artificial intelligence research and development. In the 1950s, Alan Turing envisioned a future where computers would reach such an advanced level of intelligence that distinguishing between human interaction and AI would become indistinguishable. While recent advancements like the GPT-4 algorithm demonstrate progress towards this goal, Turing initially anticipated achieving this technological breakthrough by the end of the twentieth century. In 1970, Life Magazine quoted multiple computer scientists who predicted that within 3–15 years, we would have machines with comparable general intelligence to humans. But with the advent of generative AI, we are seeing a rush to market for AI that values getting ahead of the AI trend above the real consequences that marginalised people face from its implementation.

Despite the regular highlighting of the potential and realised issues with AI for marginalised communities, the demands of the market, of commerce and of innovation mean that AI implementation continues barely abated. This has been seen over the past decade of increased AI implementation, but has seen an upswing with the easy accessibility of generative AI tools. This is despite warnings that generative AI is perpetuating harmful gender stereotypes [192], data colonialism and exacerbation of poverty in the Global South [193], and monopolistic business practices [194] to name just a few. Whether these warnings will come to pass is of little consequence. The hype, the rush to market and the fear of missing out or being left behind is pervasive. As discussed earlier, decisions to implement AI in situations where limitations of the data or biases in the algorithmic outcomes are a known fact are often made on the basis of cost cutting, efficiency, and competitive advantage. While this rush to market continues, or indeed intensifies, we will continue to see AI systems implemented that harm the LGBTQ+ community.

7.2.5 A mythology of AI and AI hype

The burgeoning ubiquity and discourse of inevitability around AI, coupled with the omnipresent nature of AI hype, arguably gives way to a pervasive mythology of AI. Technological myths are made by their ability to enter the collective imagination [195], which AI can be seen to have achieved within its socio-technological system [196].

This mythology of AI perpetuates unrealistic expectations and exaggerations about the capabilities and potential impact of AI technologies, encouraged by a variety of actors and influences, becoming pervasive throughout politics, society, and culture [195].

The mythology of AI is distributed in the form of AI hype, from a combination of factors, including media sensationalism, marketing strategies by tech companies, and the desire to attract investment and gain competitive advantage [197].

The ideology at the core of the AI myth has been suggested as "a political and social ideology rather than as a basket of algorithms. The core of the ideology is that a suite of technologies, designed by a small technical elite, can and should become autonomous from and eventually replace, rather than complement, not just individual humans but much of humanity" [198].

Operating under the auspices of such a mythology can drive implementations of AI technology that may not adequately consider the specific needs, experiences, and challenges faced by marginalised communities, or where it is not the activating reason for implementation, can ensure that the media coverage, public perceptions, corporate and financial interests, and general environments of heightened hype, will contribute to the mythologisation of AI technology.

Technology myths have a direct impact on the capabilities of policy-makers to make decisions, whereby industry generated and media hyped technology myths ultimately degrade the quality of decision-making [199]. Existing in a mythology of AI, driven by AI hype, where AI is ubiquitous, inevitable, inscrutable, and infallible ultimately frames all AI implementations and outcomes within the context of the AI myth and AI hype.

8 Areas for further research

8.1 AI, regimes of truth, and power dynamics

The AI ecosystem can be seen as a prevailing dominant discourse that functions as a regime of truth. This regime is established by those who create artificial intelligence and influence decision-makers within organisations, ultimately supporting the existing power structure. Hype is also a key feature in the construction of a regime of truth.

Individuals who are not fluent in the dominant discourse often face marginalisation, ridicule, or exclusion. However, it is worth considering that AI has only been widely recognised for less than half a century. This raises the question of whether there is potential to alter the prevailing discourse surrounding AI. Can regimes of truth be reconstructed and modified by challenging perspectives that may seem unconventional or irrational? Such an inquiry would highlight the

inherent contradictions between knowledge about AI, how this knowledge is acquired, and the decision-making processes within the field.

8.2 Accountability

The domains of ethics, technology, and society have become increasingly cognisant of the direct repercussions that AI technologies wield on their end users, as well as how the creation and utilisation of AI systems can reinforce the existing power structures and systemic biases against affected populations. As a result, there has been a widespread demand to establish mechanisms for AI accountability in a manner that is both transparent and effective. Framing the question around who or what is accountable for AI hype, AI implementation, and the impacts of these implementations is an area of investigation which would add to these discussions.

8.3 Influencing factors in AI design, investment, and implementation

In the course of developing this paper, a significant number of cases of AI design, implementation, and socio-political decision-making have been analysed, examined, and, in many cases, critiqued. However the influencing factors in the decision-making processes as pertains to AI, particularly as experienced directly by the decision-makers themselves, is rarely to be found in the literature. This paper lays the foundations for further exploration and novel data collection from decision-makers themselves to explore the influencing factors that drive implementations of AI, steers AI design and development, and underpins AI policy.

9 Conclusion

This paper aimed to address two key questions: does hype affect marginalised communities, particularly hype around new technologies such as AI; and how do AI marketing and communication strategies that leverage hype reflect systemic power dynamics, particularly as they pertain to the LGBTQ+ community.

This paper explored the connection between AI hype and its impact on the LGBTQ+ community, as well as the influence of marketing in amplifying this hype. The findings of this paper suggest that AI hype can reflect and perpetuate the existing power dynamics surrounding LGBTQ+ identities, leading to a reinforcement of heteronormative and cisnormative ideologies, and subsequently compounding the marginalisation of queer communities.

The LGBTQ+ community is fundamentally impacted by the biases and preconceptions ingrained in AI systems and algorithms. These biases can appear through misgendering or inaccurately depicting LGBTQ+ individuals in tailored ads, omitting LGBTQ+ subjects from AI-generated content, or reinforcing stereotypes about the LGBTQ+ community. Moreover, the absence of varied representation and inclusiveness in the creation and application of AI technologies contributes to perpetuating these biases. For example, AI development in societies with a history of discrimination may reinforce and worsen these biases and oppressions [202].

The findings of this paper point to several driving factors at play in encouraging hype and the adoption of AI technologies towards an ultimately harmful experience for the LGBTQ+ community. This can be predominantly seen in the impact that AI hype has on the decision-making powers with respect to AI implementation, and the fundamental societal state of heteronormativity underpinning issues around data sets, model development, and general diversity and inclusion in the AI ecosystem. Overall, this paper sheds new light on the ways in which the LGBTQ+ community is impacted by AI implementation across a number of different areas, and the ways in which hype drives the adoption of and implementation of these technologies irrespective of the harms felt by the LGBTQ+ community.

Appendix 1: Gartner hype cycle stages

Gartner hype cycle stages		
#	Phase	Description
1	Technology trigger	Potential breakthrough in technology. Stories and media attention highlighting initial proof-of-concept generate significant publicity. At this stage, there may not be any practical products available, and the commercial feasibility remains uncertain

Gartner hype cycle stages		
#	Phase	Description
2	Peak of inflated expectations	Early publicity generates a number of success stories and use cases, although this is often accompanied by many failures. At this stage, some companies are likely to become early adopters of the technology although many do not
3	Trough of disillusionment	Initial interest tapers off as early optimism fails to deliver on promised benefits. Technology producers may begin to fail and drop out of the market. Investment is predicated on existing providers improving their technology offerings
4	Slope of enlightenment	A more balanced and realistic understanding of the technology begins to emerge. Instances of how the technology may be of benefit begin to strengthen. Second- and third-generation products are released as technology creators develop further and the market responds with increased demand

Gartner hype cycle stages		
#	Phase	Description
5	Plateau of productivity	The technology in question begins to find mainstream acceptance. Adoption increases and technology is picked up beyond the early adopter group and into more cautious demographics. Assessment criteria for ongoing product viability are better defined and market applicability and relevance increases

Declarations

References

1. Gathering Strength, Gathering Storms. Retrieved from https://ai100.stanford.edu/sites/g/files/sbiybj18871/files/media/file/AI100Report_MT_10.pdf (2021)
2. Artificial intelligence is often overhyped—and here's why that's dangerous. MIT Technology Review. Retrieved from https://www.technologyreview.com/2018/09/13/240156/artificial-intelligence-is-often-overhypedand-heres-why-thats-dangerous/ (2018)
3. Udupa, S., Maronikolakis, A., Wisiorek, A.: Ethical scaling for content moderation: extreme speech and the (in)significance of artificial intelligence. Big Data Soci. **10**(1), (2023). https://doi.org/10.1177/20539517231172424
4. Yang, J., Zhang, B.: Artificial intelligence in intelligent tutoring robots: a systematic review and design guidelines. Appl. Sci. **9**, 2078 (2019). https://doi.org/10.3390/app9102078
5. Krakowski, A., Greenwald, E., Hurt, T., Nonnecke, B., Cannady, M.: Authentic integration of ethics and AI through sociotechnical, problem-based learning. Proc. AAAI Conf. Artif. Intell. **36**(11), 12774–12782 (2022). https://doi.org/10.1609/aaai.v36i11.21556

6. Baird, A., Schuller, B.: Considerations for a more ethical approach to data in AI: on data representation and infrastructure. Front. Big Data. **3**, 25 (2020). https://doi.org/10.3389/fdata.2020.00025

7. Park, R.E.: Human migration and the marginal man. Am. J. Sociol. Sociol. **33**(6), 881–893 (1928). https://doi.org/10.1086/214592

8. Lessa, I.: Discursive struggles within social welfare: restaging teen motherhood. Br. J. Soc. Work. **36**(2), 283–298 (2006). https://doi.org/10.1093/bjsw/bch256

9. Walsh, T.: A right to inclusion? Homelessness, human rights and social exclusion. Aust. J. Human Rights **12**(1), 185–204 (2006). https://doi.org/10.1080/1323238X.2006.11910818

10. Young, I.M.: Five faces of oppression. In: Adams, M. (ed.) Readings for diversity and social justice, pp. 35–49. Routledge, New York (2000)

11. Yee, J.: Critical anti-racism praxis: The concept of whiteness implicated. In: Hick, S., Fook, J., Pozzuto, R. (eds.) Social work, a critical turn, pp. 87–104. Thompson, Toronto (2005)

12. Tricco, A.C., Nincic, V., Darvesh, N., Rios, P., Khan, P.A., Ghassemi, M.M., MacDonald, H., Yazdi, F., Lai, Y., Warren, R., Austin, A., Cleary, O., Baxter, N.N., Burns, K.E.A., Coyle, D., Curran, J.A., Graham, I.D., Hawker, G., Légaré, F., Straus, S.E.: Global evidence of gender equity in academic health research: a scoping review. BMJ Open (2023). https://doi.org/10.1136/bmjopen-2022-067771

13. Stone, P.: Ghettoized and marginalized: the coverage of racial and ethnic groups in introductory sociology texts. Teach. Sociol. Sociol. (1996). https://doi.org/10.2307/1318872

14. Diamond, P.: How globalisation is changing patterns of marginalisation and inclusion in the UK. Joseph Rowntree Foundation, New York (2010)

15. Epstein, D., Quinn, K.: Markers of online privacy marginalization: Empirical examination of socioeconomic disparities in social media privacy attitudes, literacy, and behavior. Soc. Media Soc. **6**(2), 2056305120916853 (2020). https://doi.org/10.1177/2056305120916853

16. Benner, A.D., Wang, Y.: Demographic marginalization, social integration, and adolescents' educational success. J. Youth Adolesc. Adolesc. **43**, 1611–1627 (2014). https://doi.org/10.1007/s10964-014-0151-6

17. Crenshaw, K.: Mapping the margins: intersectionality, identity politics, and violence against women of color. Stanf. Law Rev. **43**(6), 1241–1299 (1991). https://doi.org/10.2307/1229039

18. Garg, I., Hanif, H., Javed, N., Abbas, R., Mirza, S., Javaid, M.A., Sheikh, A.B.: COVID-19 vaccine hesitancy in the LGBTQ+ population: a systematic review. Infect. Dis. Rep. **13**(4), 872–887 (2021). https://doi.org/10.3390/idr13040079

19. Miller, M. K.: Discrimination and Barriers to Well-Being: The State of the LGBTQI+ Community in 2022. Center for American Progress. https://www.americanprogress.org/article/discrimination-and-barriers-to-well-being-the-state-of-the-lgbtqi-community-in-2022 (2023)

20. Lau, V.W., Scott, V.L., Warren, M.A., Bligh, M.C.: Moving from problems to solutions: a review of gender equality interventions at work using an ecological systems approach. J. Organ. Behav. Behav. **44**(2), 399–419 (2023). https://doi.org/10.31234/osf.io/cy63n

21. Cerezo, A., Cummings, M., Holmes, M., Williams, C.: Identity as resistance: identity formation at the intersection of race, gender identity, and sexual orientation. Psychol. Women Q. **44**(1), 67–83 (2020). https://doi.org/10.1177/0361684319875977

22. De Lauretis, T.: Queer theory: Lesbian and gay sexualities an introduction. Differences **3**(2), iii–xviii (1991). https://doi.org/10.1215/10407391-3-2-iii

23. Halperin, D.M.: The normalization of queer theory. In: Queer theory and communication, pp. 339–343. Routledge (2014). https://doi.org/10.1300/J082v45n02_17

24. Dore, I.: Foucault on power. UMKC Law Rev. **78**(3), 737–748 (2009)

25. Iacono, G.: Epistemic injustice: towards uncovering knowledge of bisexual realities in social work research. Adv. Soc. Work **18**(2), 563–582 (2017). https://doi.org/10.18060/21427

26. Gaventa, J.: Power after Lukes: An Overview of Theories of Power since Lukes and Their Application to Development. Brighton: Participation Group, Institute of Development Studies. https://www.powercube.net/wp-content/uploads/2009/11/power_after_lukes.pdf (2003)

27. Berlant, L., Warner, M.: Sex in public. Crit. Inq. **24**(2), 547–566 (1998). https://doi.org/10.1086/448884

28. Israel, B., Gavriel, A.Y.: Cisnormativity. In: Abbie, G., Genny, B. (eds.) The SAGE Encyclopedia of Trans Studies, pp. 121–125. SAGE Publishing (2021)

29. Bauer, G.R., Hammond, R., Travers, R., Kaay, M., Hohenadel, K.M., Boyce, M.: 'I don't think this is theoretical; this is our lives': how erasure impacts health care for transgender people. J. Assoc. Nurses AIDS Care **20**(5), 348–361 (2009). https://doi.org/10.1016/j.jana.2009.07.004

30. Department for Education. Parent first approach at the core of new guidance on gender questioning children. GOV.UK. https://www.gov.uk/government/news/parent-first-approach-at-the-core-of-new-guidance-on-gender-questioning-children (2023)

31. Booth, R.: Schools in England 'face legal risks if they follow new transgender guidance.' *The Guardian*. https://www.theguardian.com/society/2023/dec/20/schools-in-uk-face-legal-risks-if-they-follow-new-transgender-guidance (2023)

32. Mehmet D, Mehmet F.: Poland Anti-LGBTI hate timeline | ILGA-Europe. *ILGA-Europe | Safety, equality and freedom for LGBTI people in Europe & Central Asia*. https://www.ilga-europe.org/report/poland-anti-lgbti-hate-timeline/ (2023)

33. Gall, L.: Discriminatory bill harms trans women in Hungary. *Human Rights Watch*. https://www.hrw.org/news/2023/07/20/discriminatory-bill-harms-trans-women-hungary (2023)

34. Kottosova, I.: LGBTQ rights Hungary Eastern Europe. CNN. https://edition.cnn.com/2021/07/01/europe/lgbtq-rights-hungary-eastern-europe-intl-cmd/index.html (2021)

35. Gill, S.R., Law, D.: Global hegemony and the structural power of capital. Int. Stud. Quart. **33**(4), 475–499 (1989). https://doi.org/10.2307/2600523

36. Kähkönen, A.K., Virolainen, V.M.: Sources of structural power in the context of value nets. J. Purch. Supply Manag. Purch. Supply Manag. **17**(2), 109–120 (2011). https://doi.org/10.1016/j.pursup.2011.01.001

37. Kitchen, N., Cox, M.: Power, structural power, and American decline. Camb. Rev. Int. Aff.. Rev. Int. Aff. **32**(6), 734–752 (2019). https://doi.org/10.1080/09557571.2019.1606158

38. Khan, T.H., MacEachen, E.: Foucauldian discourse analysis: moving beyond a social constructionist analytic. Int J Qual Methods (2021). https://doi.org/10.1177/16094069211018009

39. Weedon, C.: Feminist Practice and Post-Structuralist Theory. Basil Blackwell, Oxford (1987)

40. Bevir, M.: Foucault, power, and institutions. Polit. Stud. **47**, 345–359 (1999). https://doi.org/10.1111/1467-9248.00204

41. Foucault, M.: The Will to Knowledge, History of Sexuality. Penguin, London (1998)

42. Foucault, M., Rabinow, P.: The Essential works of Michel Foucault, 1954–1984, vol. 1. New Press (1997)

43. Burger, M.: The risk to population health equity posed by automated decision systems: a narrative review. arXiv preprint arXiv:2001.06615. https://doi.org/10.48550/arXiv.2001.06615 (2020)

44. Birhane, A. [@Abebab]: Let's ditch the common narrative that "AI is a tool that promotes and enhances human [...]". [Tweet]. Retrieved from https://twitter.com/Abebab/status/1526530264717004802 (2022).

45. Buolamwini, J., Gebru, T.: Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Proceedings of the Conference on Fairness, Accountability, and Transparency, pp. 77–91. PMLR (2018)

46. Leslie, D.: Understanding bias in facial recognition technologies. arXiv preprint arXiv:2010.07023. https://doi.org/10.48550/arXiv.2010.07023 (2020)

47. Krishnapriya, K.S., Albiero, V., Vangara, K., King, M.C., Bowyer, K.W.: Issues related to face recognition accuracy varying based on race and skin tone. IEEE Trans. Technol. Soc. 1(1), 8–20 (2020). https://doi.org/10.1109/TTS.2020.2974996

48. Civil rights concerns regarding law enforcement use of face recognition technology. New America. https://www.newamerica.org/oti/briefs/civil-rights-concerns-regarding-law-enforcement-use-of-face-recognition-technology/ (2021)

49. Gentzel, M.: Biased face recognition technology used by government: a problem for liberal democracy. Philos. Technol. 34(4), 1639–1663 (2021). https://doi.org/10.1007/s13347-021-00478-z

50. Schuetz, P.N.: Fly in the face of bias: algorithmic bias in law enforcement's facial recognition technology and the need for an adaptive legal framework. Law Inequal. 39, 221 (2021). https://doi.org/10.24926/25730037.391

51. Takshi, S.: Unexpected Inequality: Disparate-Impact From Artificial Intelligence in Healthcare Decisions. J Law Health. Retrieved from https://pubmed.ncbi.nlm.nih.gov/34185974/ (2021)

52. Koski, E., Scheufele, E., Karunakaram, H., Foreman, M.A., Felix, W., Dankwa-Mullan, I.: Understanding Disparities in Healthcare: Implications for health systems and AI applications, pp. 375–387. In Springer eBooks (2022)

53. Preventing bias and inequities in AI-Enabled health tools. Margolis Center for Health Policy. https://healthpolicy.duke.edu/publications/preventing-bias-and-inequities-ai-enabled-health-tools (2022)

54. Peña Gangadharan, S., Niklas, J.: Decentering technology in discourse on discrimination. Inf. Commun. Soc.Commun. Soc. 22(7), 882–899 (2019). https://doi.org/10.1080/1369118X.2019.1593484

55. Lynch, J.: Face off: law enforcement use of face recognition technology. SSRN (2020). https://doi.org/10.2139/ssrn.3909038

56. Hall, L., Clapton, W.: Programming the machine: gender, race, sexuality, AI, and the construction of credibility and deceit at the border. Internet Policy Rev. (2021). https://doi.org/10.14763/2021.4.1601

57. Hawthorne, C.: Dangerous Networks: Internet Regulations as Racial Border Control in Italy. In: Vertesi, J., Ribes, D. (eds.) DigitalSTS: A Field Guide for Science & Technology Studies. Princeton University Press (2019)

58. Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S.: Dissecting racial bias in an algorithm used to manage the health of populations. Science 366(6464), 447–453 (2019). https://doi.org/10.1126/science.aax2342

59. Obermeyer, Z., Nissan, R., Stern, M., Eaneff, S., Bembeneck, E.J., Mullainathan, S.: Algorithmic bias playbook. Center for Applied AI at Chicago Booth, pp. 7–8 (2021). https://www.chicagobooth.edu/-/media/project/chicago-booth/centers/caai/docs/algorithmic-bias-playbook-june-2021.pdf

60. Charette, R. N.: Michigan's MIDAS unemployment system: algorithm alchemy created lead, not gold. IEEE Spectrum. Retrieved from https://spectrum.ieee.org/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold (2021)

61. Elyounes, D.A.: "Computer Says No!": The impact of automation on the discretionary power of public officers. Vanderbilt J. Entertain. Technol. Law 23, 451 (2021)

62. AlgorithmWatch.: Finnish Credit Score Ruling raises Questions about Discrimination and how to avoid it - AlgorithmWatch. Retrieved from https://algorithmwatch.org/en/finnish-credit-score-ruling-raises-questions-about-discrimination-and-how-to-avoid-it/ (2018)

63. Kim, P.T.: Data-driven discrimination at work. William Mary Law Rev. 58(3), 857 (2017)

64. Alikhademi, K., Drobina, E., Prioleau, D., et al.: A review of predictive policing from the perspective of fairness. Artif. Intell. Law. Intell. Law 30, 1–17 (2022). https://doi.org/10.1007/s10506-021-09286-4

65. European Agency for Fundamental Rights (FRA): Bias in Algorithms: Artificial intelligence and discrimination. Publications Office of the European Union (2022)

66. Valdivia, A., Serrajòrdia, J.C., Swianiewicz, A.: There is an elephant in the room: towards a critique on the use of fairness in biometrics. AI Ethics 3, 1407–1422 (2023). https://doi.org/10.1007/s43681-022-00249-2

67. Phillips, M., Marsden, H., Jaffe, W., et al.: Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. JAMA Netw. OpenNetw. Open 2(10), e1913436 (2019). https://doi.org/10.1001/jamanetworkopen.2019.13436

68. Frennert, S.: Gender blindness: On health and welfare technology, AI and gender equality in community care. Nurs. Inq.. Inq. 28, e12419 (2021). https://doi.org/10.1111/nin.12419

69. Rice, L., Swesnik, D.: Discriminatory effects of credit scoring on communities of color. Suffolk Univ. Law Rev. 46, 935–936 (2013)

70. Whittaker, M., Alper, M., Bennett, C.L., Hendren, S., Kaziunas, L., Mills, M., West, S.M.: Disability, bias, and AI. AI Now Inst. 8, 15–17 (2019)

71. HYPE definition and meaning | Collins English Dictionary. In Collins Dictionaries. https://www.collinsdictionary.com/dictionary/english/hype (2024)

72. Siegel, E.: The AI hype cycle is distracting companies. Harvard Business Review. https://hbr.org/2023/06/the-ai-hype-cycle-is-distracting-companies (2023)

73. Flew, T.: New media: an introduction. Oxford University Press, Cham (2008)

74. Cukier, W., Ryan, P. M., & Fornssler, B. The Rhetoric of the "Information Highway" in the Media 1992–2008.

75. Pärna, K.: Believe in the net: the construction of the sacred in utopian tales of the internet. Implic. Relig. (2006). https://doi.org/10.1558/imre2006.v9i2.180

76. Howcroft, D.: The hyperbolic age of information: an empirical study of Internet usage. Inf. Commun. Soc. 2(3), 277–299 (1999). https://doi.org/10.1080/136911899359592

77. Khan, M., Wu, X., Xu, X., Dou, W.: Big data challenges and opportunities in the hype of Industry 40. In: 2017 IEEE International Conference on Communications (ICC), pp. 1–6. IEEE (2017)

78. Kitchin, R.: Big data—Hype or revolution. The SAGE handbook of social media research methods, pp 27–39. (2017).

79. Fox, S., Do, T.: Getting real about big data: applying critical realism to analyse Big Data hype. Int. J. Manag. Proj. Bus. 6(4), 739–760 (2013). https://doi.org/10.1108/IJMPB-08-2012-0049

80. Rayes, A., Salam, S.: Internet of things from hype to reality. Springer (2017)

81. Pal, A.: Internet of things: making the hype a reality. IT Prof. 17(3), 2–4 (2015). https://doi.org/10.1109/MITP.2015.36

82. Singh, A.K., Firoz, N., Tripathi, A., Singh, K.K., Choudhary, P., Vashist, P.C.: Internet of things: From hype to reality. In: An

industrial IoT Approach for Pharmaceutical Industry Growth, pp. 191–230. Academic Press (2020)

83. Michelman, P.: Seeing beyond the blockchain hype. MIT Sloan Manag. Rev. **58**(4), 17 (2017)

84. Notheisen, B., Hawlitschek, F., & Weinhardt, C.: Breaking down the blockchain hype–towards a blockchain market engineering approach. https://core.ac.uk/download/pdf/301372336.pdf (2017)

85. Perera, S., Nanayakkara, S., Rodrigo, M.N.N., Senaratne, S., Weinand, R.: Blockchain technology: Is it hype or real in the construction industry? J. Ind. Inf. Integr.Integr. **17**, 100125 (2020). https://doi.org/10.1016/j.jii.2020.100125

86. Hype Cycle for Emerging Technologies. Gartner. Retrieved from https://www.gartner.com/en/documents/4597499 (2023)

87. Understanding Gartner's Hype Cycles. Gatner. Retrieved from https://www.gartner.com/en/documents/3887767 (2018)

88. Dedehayir, O., Steinert, M.: The hype cycle model: a review and future directions. Technol. Forecast. Soc. Chang. **108**, 28–41 (2016). https://doi.org/10.1016/j.techfore.2016.04.005

89. Steinert, M., Leifer, L.: Scrutinizing Gartner's hype cycle approach. In: Picmet (2010) Technology management for global economic growth, pp. 1–13. IEEE (2010)

90. Khodayari, M., Aslani, A.: Analysis of the energy storage technology using hype cycle approach. Sustain. Energy Technol. Assess. **25**, 60–74 (2018). https://doi.org/10.1016/j.seta.2017.12.002

91. Stilgoe, J., Stilgoe, J.: In Dreams Begins Responsibility. Who's Driving Innovation? New Technologies and the Collaborative State, 39–54. https://doi.org/10.1007/978-3-030-32320-2 (2020)

92. Kumar, D.: The A—Z Of What Is Hype Marketing - Hype Marketing - Medium. Medium. https://medium.com/hype-marketing/the-a-z-of-what-is-hype-marketing-ead337fd798d (2021)

93. Huang, C.C., Lin, T.C., Lin, K.J.: Factors affecting pass-along email intentions (PAEIs): Integrating the social capital and social cognition theories. Electron. Commer. Res. Appl.Commer. Res. Appl. **8**(3), 160–169 (2009). https://doi.org/10.1016/j.elerap.2008.11.001

94. Milne, G.: Smoke & mirrors: How hype obscures the future and how to see past it. Robinson (2020)

95. Singh, S.: *Is artificial intelligence just a stupid marketing term?* https://www.computan.com/blog/is-artificial-intelligence-just-a-stupid-marketing-term. (2022)

96. Scott, S.: AI or BS? How to tell if a marketing tool really uses artificial intelligence. *The Drum.* https://www.thedrum.com/opinion/2023/03/30/ai-or-bs-how-tell-if-marketing-tool-really-uses-artificial-intelligence (2023)

97. Smith, G.: *The word "AI" has become a marketing ploy*. Mind Matters. https://mindmatters.ai/2021/07/the-word-ai-has-become-a-marketing-ploy/ (2023)

98. Surya, L.: An exploratory study of AI and Big Data, and it's future in the United States. Int. J. Creat. Res. Thoughts (IJCRT), ISSN, 2320–2882. https://doi.org/10.1729/Journal.25788 (2015)

99. Ali, M., & Abdel-Haq, M. K.: Bibliographical analysis of artificial intelligence learning in Higher Education: is the role of the human educator and educated a thing of the past?. In Fostering Communication and Learning With Underutilized Technologies in Higher Education. IGI Global, pp. 36–52, (2021)

100. Duan, Y., Edwards, J.S., Dwivedi, Y.K.: Artificial intelligence for decision making in the era of big data–evolution, challenges and research agenda. Int. J. Inf. Manage. **48**, 63–71 (2019). https://doi.org/10.1016/j.ijinfomgt.2019.01.021

101. Coldewey, D.: *AI-powered' is tech's meaningless equivalent of 'all natural*. Tech Crunch. https://techcrunch.com/2017/01/10/ai-powered-is-techs-meaningless-equivalent-of-all-natural (2017)

102. Ballenger, G. In technology, everything is always "Five to 10 years away." Slate Magazine. https://www.slate.com/articles/technology/future_tense/2017/09/in_technology_everything_is_always_five_to_10_years_away.html (2017)

103. Niemelä, M., Heikkilä, P., and Lammi, H.: A social service robot in a shopping mall: Expectations of the management, retailers and consumers. In: HRI '17: ACM/IEEE International conference on human-robot interaction, Association for Computing Machinery, New York, NY. https://doi.org/10.1145/3029798.3038301 (2017)

104. Wünderlich, N.V., and Paluch, S.: A nice and friendly chat with a bot. 38th International conference on information systems, Association for Information Systems, pp. 1–11. (2017)

105. Novak, T.P., Hoffman, D.L.: Relationship journeys in the internet of things: A new framework for understanding interactions between consumers and smart objects. J. Acad. Mark. Sci. **47**(2), 216–237 (2019). https://doi.org/10.1007/s11747-018-0608-3

106. Epley, N., Waytz, A., Akalis, S., Cacioppo, J.T.: When we need a human: Motivational determinants of anthropomorphism. Soc. Cogn. **26**(2), 143–155 (2008). https://doi.org/10.1521/soco.2008.26.2.143

107. Blut, M., Wang, C., Wünderlich, N.V., et al.: Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI. J. Acad. Mark. Sci. **49**, 632–658 (2021). https://doi.org/10.1007/s11747-020-00762-y

108. Troshani, I., Hill, S.R., Sherman, C., Arthur, D.: Do we trust in AI? Role of anthropomorphism and intelligence. J Computer Inf. Syst. **61**(5), 481–491 (2020). https://doi.org/10.1080/08874417.2020.1788473

109. Smith, B.C.: The promise of artificial intelligence: reckoning and judgement. Mit Press (2019)

110. Richardson, B., Prioleau, D., Alikhademi, K., Gilbert, J.E.: Public accountability: Understanding sentiments towards artificial intelligence across dispositional identities. In: 2020 IEEE International Symposium on Technology and Society (ISTAS), pp. 489–496. IEEE (2020)

111. Waseem, Z., Lulz, S., Bingel, J., Augenstein, I.: Disembodied machine learning: On the illusion of objectivity in nlp. arXiv preprint arXiv:2101.11974. https://doi.org/10.48550/arXiv.2101.11974 (2021)

112. Amironesei, R., Denton, E., Hanna, A., Nicole, H., Smart, A.: The case for interpretive techniques in machine learning in Fake AI, Frederike Kaltheuner, Meatspace Press (2021). (https://fakeaibook.com/Chapter-5-The-case-for-interpretive-techniques-in-machine-learning (2021)

113. Benjamin, R.: Race After Technology: Abolitionist Tools for the New Jim Code. Polity Press, Medford, MA (2019)

114. Boyd, D., Crawford, K.: Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Inf. Commun. Soc.Commun. Soc. **15**(5), 662–679 (2012). https://doi.org/10.1080/1369118X.2012.678878

115. Gitelman, L. (ed.): Raw data is an oxymoron. MIT press (2013)

116. Kroll, J.A.: The fallacy of inscrutability. Phil. Trans. R. Soc. A **376**(2133), 20180084 (2018). https://doi.org/10.1098/rsta.2018.0084

117. Kapoor, S. and Narayanan, A.: A checklist of eighteen pitfalls in AI journalism. https://www.cs.princeton.edu/~sayashk/ai-hype/ai-reporting-pitfalls.pdf?ref=hackernoon.com (2022)

118. Ray, T.: Why is AI reporting so bad? *ZDNET*. https://www.zdnet.com/article/why-is-a-i-reporting-so-bad (2020)

119. Hedman, J., Gimpel, G.: The adoption of hyped technologies: a qualitative study. Inf. Technol. Manag. **11**, 161–175 (2010). https://doi.org/10.1007/s10799-010-0075-0

120. Reventlow, N.J.: *Why tech needs to focus on the needs of marginalized groups*. World Economic Forum. https://www.wefor

um.org/agenda/2021/07/tech-focus-needs-marginalized-groups/ (2022)

121. Williams, D.P.: *Belief, Values, Bias, and Agency: Development of and Entanglement with "Artificial Intelligence"*. PhD Diss. Virginia Polytechnic Institute and State University. https://vtechworks.lib.vt.edu/server/api/core/bitstreams/51b1e215-744c-45ae-817f-f39bfd677be0/content (2022)

122. Kaltheuner, F.: Fake AI. Meatspace Press (2021)

123. Dickson, B.: Hype is killing AI – here's how we can stop it. *TNW | Syndication*. https://thenextweb.com/news/hype-is-killing-ai-heres-how-can-we-can-stop-it?utm_campaign=profeed&utm_medium=feed&utm_source=social (2019)

124. Slota, S.C., Fleischmann, K.R., Greenberg, S.R., Verma, N., Cummings, B., Li, L., Shenefiel, C.: Good systems, bad data?: Interpretations of AI hype and failures. Proc. Assoc. Inf. Sci. Technol. (2020). https://doi.org/10.1002/pra2.275

125. Mallazzo, M.: The BS-Industrial Complex of Phony A.I. *GEN*. https://web.archive.org/web/20190625190051/https://gen.medium.com/the-bs-industrial-complex-of-phony-a-i-44bf1c0c60f8 (2019)

126. Nemorin, S., Vlachidis, A., Ayerakwa, H.M., Andriotis, P.: AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. Learn. Media Technol. **48**(1), 38–51 (2022). https://doi.org/10.1080/17439884.2022.2095568

127. Publius X. A.: "Turing Tests, Catfishing, and Other Technologies of Transgender Rage." In: "Trans Fandom," edited by Jennifer Duggan and Angie Fazekas, special issue, Transformative Works and Cultures, no. 39. https://doi.org/10.3983/twc.2023.2269 (2023)

128. Gaboury, J.: "Queer Affects at the Origins of Computation," JCMS 61, no. 4 (Summer 2022): 169–174. https://doi.org/10.1353/cj.2022.0053 (2022)

129. Bornat, R.: Peter Landin obituary, The Guardian, https://www.theguardian.com/technology/2009/sep/22/peter-landin-obituary Accessed 1 Sept 2023 (2009)

130. Protection of Freedoms Act 2012, https://www.legislation.gov.uk/ukpga/2012/9/pdfs/ukpga_20120009_en.pdf

131. Policing and Crime Act 2017 https://www.legislation.gov.uk/ukpga/2017/3/pdfs/ukpga_20170003_en.pdf

132. Lynch, C. Gay men with historic convictions failed by government scheme, The Justice Gap, https://www.thejusticegap.com/gay-men-with-historic-convictions-failed-by-government-scheme/ Accessed 1 Sept 2023 (2019)

133. Syal, R.: Home Office expands scheme to pardon those criminalised for gay activity. *The Guardian*. https://www.theguardian.com/world/2023/jun/13/home-office-expands-scheme-to-pardon-those-criminalised-for-gay-activity (2023)

134. Wang, Y., Kosinski, M.: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. J. Pers. Soc. Psychol. **114**(2), 246–257 (2018). https://doi.org/10.1037/pspa0000098

135. Levin, S.: LGBT groups denounce 'dangerous' AI that uses your face to guess sexuality, The Guardian, https://www.theguardian.com/world/2017/sep/08/ai-gay-gaydar-algorithm-facial-recognition-criticism-stanford Accessed 1 Sept 2023 (2017)

136. Metcalf, J.: "The study has been approved by the IRB": Gayface AI, research hype and the pervasive data ethics gap. *Medium*. https://medium.com/pervade-team/the-study-has-been-approved-by-the-irb-gayface-ai-research-hype-and-the-pervasive-data-ethics-ed76171b882c (2017)

137. Gelman, A., Mattson, G., Simpson, D.: Gaydar and the fallacy of decontextualized measurement. Sociol. Sci. **5**, 270–280 (2018)

138. Os Keyes.: The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition. Proc. ACM Hum.-Comput.

Interact. 2, CSCW, Article 88 (November 2018), 22 pages. https://doi.org/10.1145/3274357 (2018)

139. Quinan, C.L., Hunt, M.: Biometric bordering and automatic gender recognition: challenging binary gender norms in everyday biometric technologies. Commun. Cult. Crit. **15**(2), 211–226 (2022). https://doi.org/10.1093/ccc/tcac013

140. Sharkey, N.: The impact of gender and race bias in AI, Humanitarian Law & Policy, https://blogs.icrc.org/law-and-policy/2018/08/28/impact-gender-race-bias-ai/#:~:text=The%20impact%20of%20gender%20and%20race%20bias%20in,gender%20and%20race%20discrimination%20throughout%20our%20global%20community (2018)

141. Chi, N., Lurie, E., & Mulligan, D. K.: Reconfiguring diversity and inclusion for AI ethics. In: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. pp. 447–457. https://doi.org/10.48550/arXiv.2105.02407 (2021)

142. Scheuerman, M.K., Paul, J., Brubaker, J.R.: How computers see gender. Proc. ACM Human-computer Interaction **3**(CSCW), 1–33 (2019). https://doi.org/10.1145/3359246

143. Lang, N.: This Lawsuit Alleging YouTube Discriminates Against LGBTQ+ Users was Just Tossed Out, Them, https://www.them.us/story/lawsuit-alleging-youtube-discriminates-against-lgbtq-users-tossed-out (2021)

144. Fox, C.: Tiktok admits restricting some LGBT hashtags, BBC Technology, https://www.bbc.co.uk/news/technology-54102575 (2020)

145. Botella, E. Tiktok admits is suppressed videos by disabled, queer and fat creators, Slate. https://slate.com/technology/2019/12/tiktok-disabled-users-videos-suppressed.html (2019)

146. Bassett, C.: *Will AI take over content moderation?* Mind Matters. https://mindmatters.ai/2022/01/will-ai-take-over-content-moderation/ (2023)

147. Kersley, A.: *The one problem with AI content moderation? It doesn't work*. ComputerWeekly.com. https://www.computerweekly.com/feature/The-one-problem-with-AI-content-moderation-It-doesnt-work (2023)

148. Vincent, J.: TikTok sued by former content moderator for allegedly failing to protect her mental health. *The Verge*. https://www.theverge.com/2021/12/24/22852817/tiktok-content-moderation-lawsuit-candie-frazier (2021)

149. Weng, L., Gaul, V., Vallone, A.: Using GPT-4 for content moderation. *OpenAI*. https://openai.com/blog/using-gpt-4-for-content-moderation (2023)

150. Chin-Rothmann, C., Rajic, T., & Brown, E.: A New Chapter in Content Moderation: Unpacking the UK Online Safety Bill. *Centre for Strategic International Studies*. https://www.csis.org/analysis/new-chapter-content-moderation-unpacking-uk-online-safety-bill (2023)

151. Gorwa, R., Binns, R., Katzenbach, C.: Algorithmic content moderation: Technical and political challenges in the automation of platform governance. Big Data Soc. (2020). https://doi.org/10.1177/2053951719897945

152. Haimson, O.L., Delmonaco, D., Nie, P., Wegner, A.: Disproportionate removals and differing content moderation experiences for conservative, transgender, and black social media users: marginalisation and moderation gray areas. Proc. ACM Human-Computer Interaction **5**(CSCW2), 466 (2021). https://doi.org/10.1145/3479610

153. Little, O.: Tiktok's recommendation algorithm is promoting homophobia and anti-trans violence. Media Matters for America. Retrieved from https://www.mediamatters.org/tiktok/tiktoks-recommendation-algorithm-promoting-homophobia-and-anti-trans-violence (2021)

154. Center for Countering Digital Hate.: Digital Hate: Social Media's Role in Amplifying Dangerous Lies About LGBTQ+ People. https://counterhate.com/research/digital-hate-lgbtq/ (2023)

155. Mahalingam, G., Ricanek, K.: Is the eye region more reliable than the face? A preliminary study of face-based recognition on a transgender dataset. In: 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–7. IEEE (2013)

156. Ovalle, A., Liang, D., & Boyd, A. Should they? Mobile Biometrics and Technopolicy Meet Queer Community Considerations. In Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '23) (Article 32, pp. 1–10). Association for Computing Machinery. https://doi.org/10.1145/3617694.3623255 (2023)

157. Scheuerman, M. K., Branham, S. M., Hamidi, F. Safe Spaces and safe places: unpacking technology-mediated experiences of safety and harm with transgender people. https://doi.org/10.1145/3274424 (2018)

158. Factora, J. (2022). TERFs Are Using Google Maps to Track and Target Trans Healthcare Providers. Them. Retrieved December 5, 2022, from https://www.them.us/story/terfs-google-maps-hospitals-community-centers in Giantini, G. (2023). The sophistry of the neutral tool: Weaponizing artificial intelligence and big data into threats toward social exclusion. AI Ethics, 3, 1049–1061. https://doi.org/10.1007/s43681-023-00311-7

159. Culzac, N. (2014). Egypt's police 'using social media and apps like Grindr to trap gay people'. The Independent. Retrieved from https://www.independent.co.uk/news/world/africa/egypts-police-using-social-media-and-apps-grindr-trap-gay-people-9738515.html

160. Norori, N., Hu, Q., Aellen, F.M., Faraci, F.D., Tzovara, A.: Addressing bias in big data and AI for health care: A call for open science. Patterns (2021). https://doi.org/10.1016/j.patter.2021.100347

161. Barbee, H., Deal, C., Gonzales, G.: Anti-Transgender Legislation—A Public Health Concern for Transgender Youth. JAMA Pediatr. 176(2), 125–126 (2022). https://doi.org/10.1001/jamapediatrics.2021.4483

162. Nenad Tomasev, Kevin R. McKee, Jackie Kay, and Shakir Mohamed. (2021). Fairness for Unobserved Characteristics: Insights from Technological Impacts on Queer Communities. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21), May 19–21, 2021, Virtual Event, USA. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3461702.3462540

163. Donnelly, N., Stapleton, L.: Digital enterprise technologies: Do enterprise control and automation technologies reinforce gender biases and marginalisation? IFAC-PapersOnLine 54(13), 551–556 (2021). https://doi.org/10.1016/j.ifacol.2021.10.507

164. Bhattasali, N.X., & Maiti, E. (2015). Machine " Gaydar " : Using Facebook Profiles to Predict Sexual Orientation. https://cs229.stanford.edu/proj2015/019_report.pdf

165. Chudy, E. (2022). 'Just for females' social media app Giggle under fire for 'excluding' trans women. PinkNews. Retrieved from https://www.thepinknews.com/2022/01/24/giggle-app-trans-women-females-sall-grover/

166. Melendez, S. (2018). Uber driver troubles raise concerns about transgender face recognition. Fast Company & Inc. Retrieved from https://www.fastcompany.com/90216258/uber-facerecognition-tool-has-locked-out-some-transgender-drivers

167. Wilkinson, P.H.: The legal implications of sexual orientation-detecting facial recognition technology. Dukeminier Awards Best Sex. Orientat. Gender Identity Law Rev. 20, 301–342 (2021)

168. O'Shaughnessy, M.: How hype over AI superintelligence could lead policy astray. Carnegie Endowment for International Peace. (2023)

169. Santos, R.: Government and Artificial Intelligence: From hype to strategy. Observatory of Public Sector Innovation. Retrieved from https://oecd-opsi.org/blog/government-and-artificial-intelligence-from-hype-to-strategy/ (2019)

170. European Union.: Artificial intelligence and EU Border: Overview of applications and key issues. https://doi.org/10.2861/91831 (2021)

171. UK Government: Defence Artificial Intelligence Strategy. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1082416/Defence_Artificial_Intelligence_Strategy.pdf (2022)

172. Rachovitsa, A., Johann, N.: The human rights implications of the use of ai in the digital welfare state: lessons learned from the dutch SyRI case. Human Rights Law Rev. 22(2), ngac010 (2022). https://doi.org/10.1093/hrlr/ngac010

173. Misuraca, G., van Noordt, C., & Boukli, A.: The use of AI in public services: Results from a preliminary mapping across the EU. In: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance. pp. 90–99. https://doi.org/10.1145/3428502.3428513 (2020)

174. Galanos, V.: Expectations and expertise in artificial intelligence: Specialist views and historical perspectives on conceptualisation, promise, and funding https://doi.org/10.7488/era/3188 (2023).

175. Gonsalves, T.: The summers and winters of artificial intelligence. In IGI Global eBooks (2018). https://doi.org/10.4018/978-1-5225-2255-3.ch021

176. Kapania, S., Siy, O., Clapper, G., Meena, S. P. A., & Sambasivan, N.: Because AI is 100% right and safe": User Attitudes and Sources of AI Authority in India. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22) (Article 158). Association for Computing Machinery. https://doi.org/10.1145/3491102.3517533 (2022)

177. Kim, P.: AI and Inequality. The Cambridge Handbook on Artificial Intelligence & the Law. Washington University in St. Louis Legal Studies Research Paper, (21–09). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3938578 (2021)

178. Dakalbab, F., Abu Talib, M., Abu Waraga, O., Bou Nassif, A., Abbas, S., Nasir, Q.: Artificial intelligence & crime prediction: a systematic literature review. Soc. Sci. Human. Open 6(1), 100342 (2022). https://doi.org/10.1016/j.ssaho.2022.100342

179. Salehi, M., Ghahari, S., Hosseinzadeh, M., Ghalichi, L.: Domestic violence risk prediction in Iran using a machine learning approach by analyzing Persian textual content in social media. Heliyon 9(5), e15667 (2023). https://doi.org/10.1016/j.heliyon.2023.e15667

180. Kawakami, A., Sivaraman, V., Cheng, H.-F., Stapleton, L., Cheng, Y., Qing, D., Perer, A., Wu, Z. S., Zhu, H., & Holstein, K.: Improving Human-AI Partnerships in Child Welfare: Understanding Worker Practices, Challenges, and Desires for Algorithmic Decision Support. In CHI Conference on Human Factors in Computing Systems (CHI '22), April 30 – May 6, 2022. ACM. https://doi.org/10.1145/3491102.3517439 (2022).

181. Garcia, J. M., & Richardson, G.: Using artificial intelligence for social impact. Nesta. Retrieved from https://www.nesta.org.uk/blog/using-artificial-intelligence-for-social-impact/ (2022)

182. Polli, F. Using AI to Eliminate Bias from Hiring. Harvard Business Review. Retrieved from https://hbr.org/2019/10/using-ai-to-eliminate-bias-from-hiring (2023)

183. Van Loon, R.: How to use AI in hiring to eliminate bias. Simplilearn.com. Retrieved from https://www.simplilearn.com/how-to-use-ai-in-recruitment-to-eliminate-bias-article (2023)

184. Walters, L.: 10 AI-Powered Tools for reducing Bias in Recruitment. PharmiWeb.jobs. Retrieved from https://www.pharmiweb.jobs/article/10-ai-powered-tools-for-reducing-bias-in-recruitment (2023)

185. Drage, E., & Mackereth, K.: Does AI Debias Recruitment? Race, Gender, and AI's "Eradication of Difference. Philos. Technol. 35: 89. https://doi.org/10.1007/s13347-022-00543-1 (2022)

186. Chen, Z.: Ethics and discrimination in artificial intelligence-enabled recruitment practices. Humanit. Soc. Sci. Commun. **10**, 567 (2023). https://doi.org/10.1057/s41599-023-02079-x

187. Horodyski, P.: Recruiter's perception of artificial intelligence (AI)-based tools in recruitment. Computers Human Behav. Rep. **10**, 100298 (2023). https://doi.org/10.1016/j.chbr.2023.100298

188. Bender, E. M., & Hanna, A.: *AI Causes Real Harm. Let's Focus on That over the End-of-Humanity Hype*. Scientific American. https://www.scientificamerican.com/article/we-need-to-focus-on-ais-real-harms-not-imaginary-existential-risks/ (2023)

189. Johnson, K.: Algorithms allegedly penalized Black renters. The US government is watching. *WIRED*. https://www.wired.com/story/algorithms-allegedly-penalized-black-renters-the-us-government-is-watching/ (2023)

190. Dave, P.: ChatGPT is cutting Non-English languages out of the AI revolution. *WIRED*. https://www.wired.com/story/chatgpt-non-english-languages-ai-revolution/ (2023)

191. Grandinetti, J.: Examining embedded apparatuses of AI in Facebook and TikTok. AI Soc. **38**, 1273–1286 (2023). https://doi.org/10.1007/s00146-021-01270-5

192. Lamensch, M.: Generative AI tools are perpetuating harmful gender stereotypes. Centre for International Governance Innovation. https://www.cigionline.org/articles/generative-ai-tools-are-perpetuating-harmful-gender-stereotypes/ (2023b)

193. Arora, A., Barrett, M., Lee, E., Oborn, E., Prince, K.: Risk and the future of AI: Algorithmic bias, data colonialism, and marginalization. Inf. Organ. **33**(3), 100478 (2023). https://doi.org/10.1016/j.infoandorg.2023.100478

194. Anderson, J., & Rainie, L.: As AI Spreads, Experts Predict the Best and Worst Changes in Digital Life by 2035. https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2023/06/PI_2023.06.21_Best-Worst-Digital-Life_2035_FINAL.pdf (2023)

195. Ballatore, A. & Natale, S.: Technological failures, controversies and the myth of AI. In: Handbook of Critical Studies of Artificial Intelligence. Ed. Simon Lindgren. Cheltenham, UK: Edward Elgar, 2023, pp. 237–44. (2023)

196. Johnson, D.G., Verdicchio, M.: Reframing AI discourse. Mind. Mach. **27**(4), 575–590 (2017). https://doi.org/10.1007/s11023-017-9417-6

197. Lente, H V., Spitters, C., & Peine, A.: Comparing technological hype cycles: Towards a theory. https://doi.org/10.1016/j.techfore.2012.12.004 (2013)

198. Lanier, J.: AI Is An Ideology, Not A Technology. *Wired*. https://www.wired.com/story/opinion-ai-is-an-ideology-not-a-technology/ (2021)

199. Peeters, P., Higham, J., Kutzner, D., Cohen, S., Gössling, S.: Are technology myths stalling aviation climate policy? Transp. Res. Part D Transp. Environ. **44**, 30–42 (2016). https://doi.org/10.1016/j.trd.2016.02.004

200. Głowacka, D., Youngs, R., Pintea, A., Wołosik, E.: Digital technologies as a means of repression and social control. Policy Department for External Relations, Directorate General for External Policies of the Union. https://www.europarl.europa.eu/thinktank/en/document/EXPO_STU(2021)653636 (2021).