JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 5, NO. 4, 2023, 1–30

MAPPING THE SOCIAL IMPLICATIONS OF PLATFORM ALGORITHMS FOR LGBTQ+ COMMUNITIES

David Myles[a], Stefanie Duguay[b] and Lucia Flores Echaiz[c]

ABSTRACT

LGBTQ+ communities were among the first to appropriate the Internet to experiment with their identities and socialize outside of mainstream society. Recently, those platforms have implemented algorithmic systems that curate, exploit, and predict user practices and identities. Yet, the social implications that platform algorithms raise for LGBTQ+ communities remain largely unexplored. At the intersection of media and communication studies, science and technology studies, as well as gender and sexuality studies, this paper maps the main issues that platform algorithms raise for LGBTQ+ users and analyzes their implications for social justice and equity. To do so, it identifies and discusses public controversies through a review and analysis of journalistic articles. Our analysis points to five important algorithmic issues that affect the lives of LGBTQ+ users in ways that require additional scrutiny from researchers, policymakers, and tech developers alike: the ability for sorting algorithms to identify, categorize, and predict the sexual orientation and/or gender identity of users; the role that recommendation algorithms play in mediating LGBTQ+ identities, kinship, and cultures; the development of automated anti-LGBTQ+ speech detection/filtering software and the collateral harm caused to LGBTQ+ users; the power struggles over the nature and effects of visibility afforded to LGBTQ+ issues/people online; and the overall enactment of cisheteronormative biases through platform affordances.

Keywords: algorithms; digital platforms; LGBTQ+ communities.

[a] Institut national de la recherche scientifique, Canada
[b] Concordia University, Canada
[c] Université du Québec à Montréal, Canada

1 INTRODUCTION

Lesbian, gay, bisexual, trans, and queer (LGBTQ+) communities were among the first to appropriate the Internet in the late 1990s to experiment with their identities and socialize outside of mainstream society (Campbell, 2005; Gray, 2009). Today, digital platforms are integrated in nearly all components of LGBTQ+ cultures and lives. Social media platforms are used by LGBTQ+ people to socialize (Duguay, 2019), especially among younger publics (Robards et al., 2018). Dating sites and apps, like Grindr and Her, are used to foster intimate and sexual relationships (Ferris & Duguay, 2020; Myles, 2020), whereas platforms like Facebook and Instagram play an integral part in LGBTQ+ activism and political organizing (Ayoub & Brzezińska, 2015; Myles & Lewis, 2019). The Internet has represented somewhat of a safe harbour for LGBTQ+ communities (Lucero, 2017), especially since many physical spaces for queer socializing, like bars and clubs, have closed (Renninger, 2018), been violently attacked (Ramirez et al., 2018), or gentrified to attract mainstream clienteles (Nash, 2013).

However, the claim that digital platforms directly empower LGBTQ+ citizens should not be made uncritically, as platforms can also play an important part in their oppression. Researchers are increasingly examining how digital platforms are implicated in forms of gender and sexual discrimination (Hanckel et al., 2019; Mainardi & Pavan, 2020). Studies examining the potential risks posed by digital platforms for LGBTQ+ communities have highlighted the propensity for their members to become the targets of online hate speech (Lingiardi et al., 2019), cyberbullying (Elipe et al., 2018), and harassment (Marciano & Antebi-Gruszka, 2020). These studies have concluded that LGBTQ+ individuals, like people of color (Daniels, 2013) and women (Mendes et al., 2018), are more likely to experience violence online than their cisheterosexual, white, and male counterparts (Abreu & Kenny, 2018).

Though undoubtedly valuable, these studies tend to focus on individual user practices and underplay how platform affordances and operating models may themselves reproduce social inequalities (Gillespie, 2010; Hoffmann, 2019). Beyond individual user practices, researchers must also consider how digital platforms increasingly rely on sophisticated algorithms that can affect the lives of LGBTQ+ users, namely, by automating sexual- and gender-based biases (Massanari, 2017). Indeed, digital platforms have recently implemented algorithmic systems that curate, exploit, and predict user practices and identities (Bucher, 2018). So far, researchers have criticized platform algorithms for introducing biases in automated decision-making that disproportionally affect women and people of color (Benjamin, 2019; Noble, 2018). This paper extends these critical reflections by exploring how platform algorithms can specifically

affect the lives of LGBTQ+ users. Through a review and analysis of journalistic articles (Chartier, 2003), it aims to identify significant issues that platform algorithms raise for LGBTQ+ communities and reflect on their implications, especially in terms of social justice and equity.

2 LITERATURE REVIEW

If the Internet was initially composed of countless independent sites, today, most users interact on a handful of privately owned platforms, what is sometimes referred to as the platformization of the Web (Helmond, 2015). Platform studies invite researchers to examine how digital platforms, like Facebook and Twitter, partly configure user activities by enacting a series of affordances, that is, the sociotechnical possibilities provided to, and the constraints imposed on, users (Bucher & Helmond, 2018). These affordances typically serve platforms' common imperatives of fostering user engagement online, breaking down user activities into data points, and curating them into large datasets to build predictive models and attract potential third-party companies (van Dijck, 2014). The datafication of user activities relies on sophisticated algorithmic systems that manage, organize, and exploit increasingly complex data infrastructures (Musiani, 2013). Algorithms now oversee nearly all activities that take place on digital platforms: they select and order the results of searches, they filter, recommend, or censor certain contents, they monitor user activities to predict their preferences, and they score, evaluate, and moderate user content or even users themselves, among other tasks (Latzer et al., 2016).

Social researchers are increasingly interested in algorithms for their propensity to reshape our private lives (De Filippi, 2016) and our possibilities for collective action (Milan, 2015), as well as in their capacity to exert power (Beer, 2009; Bucher, 2018). Studying algorithms amounts, then, to assessing how they are shaped by human meaning as well as how they shape human meaning in return (Seaver, 2017). For example, digital platforms can reproduce patterns of exclusion by imposing users with preconfigured sociotechnical categories (e.g., gender, racial, and sexual identities) that help these platforms produce data that is easier for algorithms to recognize and exploit (Gillespie, 2017). Platforms also rely on algorithms to automatically determine who or what is deemed important, legitimate, valuable, or socially acceptable online (Crawford & Gillespie, 2016). Inquiries into the political nature of algorithms have largely sought to challenge their presumed neutrality (Cardon, 2015; Crawford, 2016a). Their objective is not to assess whether algorithms can be wrong in a technical sense as much as it is to understand how they can do wrong in an ethical or political sense (Gillespie, 2012; Tufekci, 2015). Indeed, algorithmic systems can produce harm or injustices by automating pre-existing biases

shared by the people who develop, implement, or use them (Garcia, 2016), especially as there are still no clear guidelines overseeing algorithmic innovation or governance (Ananny, 2016; Introna, 2016).

A scholarship on the implications that platform algorithms – and platform affordances more generally – raise for LGBTQ+ users has emerged over the past few years. For example, researchers have suggested that the predictive nature of platform algorithms can result in outing LGBTQ+ people online by promoting default settings like publicness and visibility (Cho, 2018; Werbin et al., 2017). Some scholars have illustrated how digital platforms, guided by industry imperatives, rely on binary classification systems that reproduce cisheteronormative assumptions about gender and sexual orientation (Bivens & Haimson, 2016; Lingel & Golub, 2015) that can further lead to online harassment against LGBTQ+ users (Albury et al., 2020; Blackwell et al., 2017). Other researchers have focused on the cisheteronormative biases reproduced by automated content moderation strategies used by digital platforms, documenting how they label LGBTQ+ content or users as being "questionable" or "offensive" (Anderson & Roth, 2020). For example, platforms like YouTube appear to be more likely to flag LGBTQ+ channels as 'inappropriate' and demonetize them by employing recommendation algorithms positively biased toward "family-friendly" content (Fredenburg, 2020; Southerton et al., 2020; Wilkinson & Berry, 2020). In this context, cisheteronormative biases refer to sets of beliefs and attitudes that normalize heterosexuality and cisgender identity (i.e., gender identity matching sex assigned at birth), while also making non-heterosexual and transgender/non-binary individuals abnormal in ways that legitimize deliberate or inadvertent discrimination against them (Adam, 2015).

Other researchers have examined how changes in content moderation policies, especially those targeting "explicit" or "inappropriate" content, negatively impact LGBTQ+ online communities developed to experiment and socialize outside of cisheteronormative environments (Byron, 2019; Pilipets & Paasonen, 2020). Furthermore, these platforms' terms of service often fail to ensure the safety of their LGBTQ+ users and can participate in further censoring the content they produce online (Duguay et al., 2020; Oliva et al., 2020). This emerging scholarship raises important issues in matters of algorithmic governance, as LGBTQ+ communities exert little control over the Internet regulations that oversee their online activities (DeNardis & Hackl, 2016). As such, these issues are indicative of broader trends in the automation of dataveillance by digital platforms that raise important issues in terms of social equity and self-determination (Wood & Monahan, 2019), especially for LGBTQ+ communities (Kafer & Grinberg, 2019).

3 CONDUCTING AN STS-INFORMED PRESS REVIEW AND ANALYSIS

This paper is primarily grounded in media and communication studies and mobilizes key insights from the fields of science and technology studies (STS) as well as gender and sexuality studies. First, STS examine technologies as socially and historically situated artefacts (Sismondo, 2010), with a particular interest in the human values that are embedded in technological design (Friedman & Kahn, 2003; Nissenbaum, 2005). Communication scholars often mobilize STS scholarship to understand how communication technologies are shaped by humans and how they shape humans in return (MacKenzie & Wajcman, 1985). Feminist STS critique is particularly relevant for this study, as it examines the interplay between the social and material processes that participate in constructing technologies (Wajcman, 2010). Since the 1990s, this scholarship has unpacked the gender and sexual norms that are embedded in digital technologies and their tendency to exacerbate social inequalities (Suchman, 2008). This study also mobilizes key theoretical and methodological insights from controversy mapping and analysis (Marres, 2015). It investigates how different stakeholders make sense of Internet regulatory issues by shaping them as public controversies, namely, to strategically highlight (or underplay) some of their sociocultural or political implications (Musiani, 2018). Controversies have long been used in STS to examine the ruptures in seemingly seamless technologies, that is, to reveal the values embedded in them by paying attention to moments of failure (Star, 1999). In this paper, algorithmic controversies were identified through the press review and analysis detailed below.

Second, this study is informed by gender and sexuality studies that deconstruct and examine the performative nature of gender and sexual identities (Butler, 1990; Giffney & O'Rourke, 2016). This scholarship is grounded in the pivotal work of Foucault (1978) and his successors who revealed the institutional processes through which certain gender and sexual identities have become historically stigmatized and how LGBTQ+ individuals were prevented from fully exercising their civic privileges (Evans, 2007; Richardson, 2017). Furthermore, Foucault's (1975) investigation into the institutional mechanisms of control and discipline has largely contributed to the field of surveillance studies (Lyon, 2002), whose queer-informed subfield (Kafer & Grinberg, 2019; Phillips & Cunningham, 2007) also guides this study. Today, digital platforms and the algorithms they operate have become new sociotechnical institutions (Napoli, 2013). These emerging regulatory bodies are actively enacting sexual and gender categories whose performative power raises issues for LGBTQ+ communities in terms of social exclusion and discrimination (Southerton et

al., 2020). In turn, critical inquiries into the performative nature of gender and sexual categories bridge canonical work in STS on the role of technological categories in sorting and classifying ideas, things, and people (Bowker & Star, 2000).

In light of our STS-informed theoretical framework, we adapted the press review and analysis method (Chartier, 2003), which consisted of a detailed search in newspapers and online magazines through the use of keyword lists. We narrowed our review via three inclusion criteria and their related keywords: a) 54 keywords pertaining to digital platforms (Search field A, located anywhere in the text: e.g., 'digital media', 'dating app', 'Google'); b) 7 keywords pertaining to algorithms and datafication (Search field B, located anywhere in the text: e.g., 'algorithmic', 'artificial intelligence', 'big data'); and c) 39 keywords pertaining to sexual and gender diversity (Search field C, located anywhere excluding the body of the text: e.g., 'lesbian', 'queer', 'transgender'). Those keywords were used in ProQuest Global Newsstream, a database that counts over 2,800 news sources around the world. Our search was limited to full-text articles written in English published between January 2010 and February 2022 (articles published before this period were mostly found to be false positives). Dissertations, scholarly articles, and conference papers were excluded from this study, as our team was conducting a concomitant review of the scientific literature at the time.

This process yielded 1,243 articles from newspapers, magazines, news blogs, and other news media sources, to which our team members applied the established inclusion criteria through the collaborative reference management software Zotero. To be considered, articles needed to explicitly address digital platforms, algorithmic and/or datafication processes, as well as LGBTQ+ users. Our team also removed duplicates and false positives during that step. For example, numerous articles about Alan Turing were removed, as well as news wires summarizing the last 24-hour news cycle. Through that process, 276 articles were identified through ProQuest. A second step of our review consisted of a search through the Google and Google News search engines with a reduced version of our keyword list to comply with the platform's limitations. After eliminating duplicates, 67 articles were added to our corpus, for a total of 343 articles. By using algorithmic functions as a guiding concept (adapted from Latzer et al., 2016), we classified our corpus into five categories: sorting (n=106), recommending (n=81), filtering (n=112), and searching (n=35), as well as a fifth meta-category examining issues related to cisheteronormativity in the tech industry more generally (n=24), with certain articles recurring in more than one category.

This review process was not without challenges and limitations. The selection of our guiding concept to categorize our corpus was the subject of ongoing debate within our team. Various concepts could have been used to perform this classification with varying results. For example, we could have sorted articles by platform types, by social activities, by user populations, or by social justice issues. Ultimately, we decided to use algorithmic functions as our main guiding concept because it best served the objective of this paper, which is to highlight the implications that platform algorithms raise for LGBTQ+ users. Furthermore, using algorithmic functions allowed us to identify transversal algorithmic issues across platforms and user populations, an output that our team found particularly valuable given the variety of people and sites covered in the corpus.

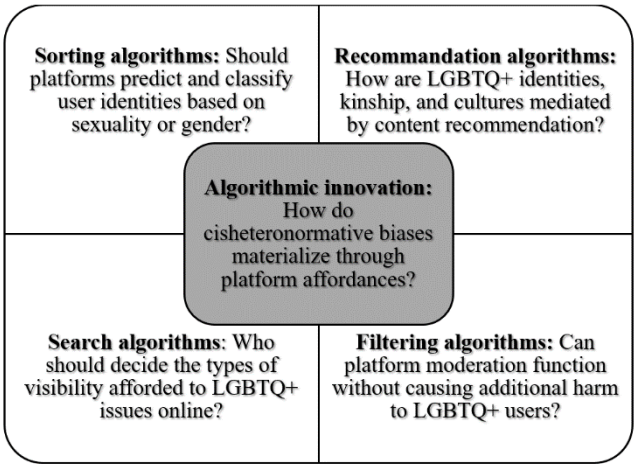

Figure 1 Platform algorithms and their social implications for LGBTQ+ users.

Moreover, the use of algorithmic functions as a guiding concept raised a significant ontological issue, namely, because those functions are not clear-cut but often interrelated and distributed. For example, it can be difficult to distinguish between recommendation and filtering algorithms, especially when recommendation algorithms are used as a filtering mechanism (Gillespie, 2018). Like all algorithmic decision-making, sorting remains a highly arbitrary task (Bowker & Star, 2000). Through our process, our goal was to identify the algorithmic function that was at the heart of the controversy or case study exposed in each article. When an article addressed multiple algorithmic functions in significant ways, it was classified accordingly. Effectively, our sorting system aimed at distributing our corpus coherently between conceptual categories in a way that would yield important questions and support meaningful discussions. These questions have been summarized in Figure 1 (see above). Most importantly, our team members sought consistency in their decision-making, and

ensured that the reasoning behind our decisions was stated clearly and transparently in each of the sections that follow[1].

3.1 Sorting algorithms and the prediction of LGBTQ+ identities

Sorting algorithms monitor and quantify platform users' activities to predict or infer individual qualities associated with them that can later be commodified. As such, they directly relate to platform dataveillance (Van Dijck, 2014), which has become the topic of increased public interest following the Snowden revelations and the Cambridge Analytica scandal. Algorithmic sorting relates to foundational considerations in platform and surveillance studies over the role of technology in the labelling and classification of social identities and relations (Lyon, 2005). Sorting is often accomplished through the use of predetermined sociotechnical categories that may reproduce binary conceptions of sexuality and gender (Bivens, 2017). Overall, the articles in this first category fall into two subsets: those that address the safety and privacy implications raised by sorting algorithms in a more general fashion; and those that specifically highlight the increasing capacity for digital platforms to classify or predict the gender and/or sexual identity of their users (as well as for external actors to train their algorithms by exploiting accessible social media data).

A first subset of articles offers critical reflections on the general implications that digital platforms raise for LGBTQ+ communities in terms of safety and privacy. Platforms were especially criticized for their propensity to provide hostile governments or organizations with the means to identify, track, and oppress LGBTQ+ users[2]. Those articles highlighted an important paradox: on the one hand, digital platforms play an emancipatory role, as they allow LGBTQ+ users to shape their own communities online; on the other, sorting algorithms and the surveillance imperatives they serve increase the vulnerability of LGBTQ+ citizens by potentially exposing them to state-sponsored or corporate mistreatment. Those concerns were particularly expressed toward dating and hookup apps that cater to LGBTQ+ users, likely because of the sensitive nature of the data they produce or manage (e.g., intimate pictures, sexual preferences, health-related information)[3]. Unsurprisingly, the hookup app Grindr was most commonly cited in this subset of articles, as the app was at the center

[1] Our list of journalistic references, which we have ordered alphabetically according to each analytical section of this article, can be accessed here: https://borealisdata.ca/dataset.xhtml?persistentId=doi:10.5683/SP3/QOFTIC

[2] Hackl (2014); Harford (2019); Hassett (2014); Mohd Yatid (2018); Nangia (2020); Shezaf & Jacobson (2018); Wareham (2021a; 2021b).

[3] Anonymous (2020a); Harari (2019); Ramos (2016); Rudder (2014); Wood (2014); Wood (2018).

of various controversies over the past decade. Its acquisition by the Chinese corporation Kunlun Tech between 2016 and 2018 garnered fears over the misuse of sensitive user data by China[4], as previously discussed in the scientific literature (Kokas, 2022; Myles, 2022). The app was also criticized for its role in the public outing of athletes and religious figures, as well as for sharing HIV-related data to third-party companies[5].

A second issue on dating apps arose around the introduction of new gender categories that cater to trans and gender nonconforming users, especially on Tinder[6]. Those debates relate to broader considerations surrounding gender inclusivity on digital platforms. An important tension has emerged between the need for platforms to proliferate gender categories to ensure user inclusivity and the belief that such categories should be deleted altogether in favor of gender neutrality or fluidity. This tension was also discussed in relation to linguistic-based platform services[7]. Overall, those articles point to an ongoing debate around social sorting and queerness (Schram, 2019). Namely, does the introduction of new gender and sexual categories by digital platforms truly aim for a better representation and inclusion of LGBTQ+ users or do these initiatives conceal corporate schemes seeking to better predict and commodify sexual- and gender-diverse users through the training of high-performing sorting algorithms?

A second subset of articles in this category addresses controversies in which digital platforms operate sorting algorithms to predict the sexual and/or gender identity of their users, either through the use of behavioral data or facial recognition software. First, several articles addressed how Facebook can predict the sexual orientation of their users with 'just a few likes' or by analyzing user friends lists[8]. Published in the wake of the 2013 Cambridge Analytica controversy, the case of sexual prediction by Facebook was mainly characterized as a privacy rights issue. It was sometimes referred to as a sort of algorithmic 'gaydar', a reflection that was also extended to other platforms[9]. A second corpus of articles addressed the 2017 Kosinski experiment, in which a Stanford researcher collected tens of

[4] AFP (2019); Anonymous (2020a); Stone Fish (2019).

[5] Boorstein et al. (2021); Jefferson (2020).

[6] Burke (2016); Mallenbaum (2016); Newmark (2015); O'Brien (2017); OkCupid (2018); Stack (2016a; 2016b); Tanna (2021).

[7] Fingas (2018); Reuters (2018).

[8] AFP (2013); Anonymous (2014); Anonymous (2018a); Alloway (2015); Anderson (2013); Ball (2013); Cadwalladr & Graham-Harrison (2018); Daily Telegraph (2017); Franklin (2013); Goldbeck (2015); Kendall (2013); Jayson (2013); McCarthy & Cookson (2013); Naughton (2017); Preston (2014); Rainey (2015); Taylor (2013); The Telegraph (2013); Tinker (2018); The Herald (2013); Tufekci & King (2014); Tyree (2017).

[9] Anonymous (2020b); Khalaf (2018).

thousands of pictures from a popular dating site and supposedly trained a facial recognition software to predict sexual orientation based on key facial features[10]. This researcher claimed that his software could predict the sexual orientation of users with a high level of certainty. If some articles have painted Kosinski as a sort of whistleblower against LGBTQ+ platform surveillance, his research was heavily criticized both for its conceptual limitations (e.g., for its propensity to reproduce binary conceptions of sexual orientation) and for its implications in terms of user safety and privacy, especially if such software were to fall into the wrong hands.

In 2019, a related controversy arose around the propensity for facial recognition software operated by digital platforms to misgender trans and non-binary users[11]. This controversy echoes a paradox in public discourse between the need for facial recognition software to acknowledge trans and gender nonconforming users to ensure their full inclusion in a digital society versus the inherent privacy and safety issues that such surveillance apparatus raises[12] (Lingel, 2020). This paradox is best illustrated by the Facebook 10-year challenge, which was seen as a way for trans users to visually narrate their own transitioning story online or, alternately, as a ploy to help digital platforms train their facial recognition software to better predict the effects of gender-affirming surgeries or therapies.

As illustrated by the controversies above, the majority of articles in this section characterize social sorting algorithms as being predominantly oppressive toward LGBTQ+ communities. However, a subset highlighted the potential of platform dataveillance for LGBTQ+ health and safety. For example, the Trevor Project developed an AI-based surveillance system to identify youth at risk of suicide in collaboration with Google[13], researchers used big datasets to support the health of trans users via Twitter[14], while other researchers talked about the dataveillance potential of the dating app Blued to develop sexual health initiatives[15]. Similarly, one article evoked the potential of platform dataveillance to identify LGBTQ-phobic personnel in the US Army in collaboration with the Rand corporation[16]. Other articles

[10] Ahmed (2017); Anonymous (2017); Baska (2017); Coldewey (2017; 2021); Daily Telegraph (2017); Dialani (2021); Fernandez (2017); Harari (2019); Hawkins (2017); Hindustan Times (2017); Holden (2018); Kalaichandran (2017); Kayvon (2017); Kuang (2017); Levin (2017); Lewis (2018); Morgan (2018); Murphy (2017); Rose (2017); Segal (2021a; 2021b); ; Sulleyman (2017); Sweeney (2019); Volokh (2017).

[11] Burt (2020); Crockford (2019); Fried (2019); Haggard (2019); Khalid (2019); Metz (2019); Rose (2019); Swerling (2019); Thalheim (2019); Trout (2017).

[12] Watling (2021).

[13] Jones et al. (2020); Srikanth (2021); The Trevor Project (2021a; 2021b).

[14] Rivero (2015); Sharma (2015).

[15] Blued (2019).

[16] Wentling (2021).

explored the use of dataveillance schemes to support the safety of increasingly mobile LGBTQ+ consumers, especially in the field of tourism[17]. For example, the app GeoSure was developed to help LGBTQ+ people identify the safest neighborhoods when travelling[18].

Overall, the articles cited in this section illustrate ongoing debates in the literature at the intersection of dataveillance and queerness (Schram, 2019). The sorting of sexual and gender identities by platform algorithms can be perceived as a powerful tool to afford visibility to LGBTQ+ users and provide them with legitimacy by shaping them as a recognizable social group. Yet, algorithmic sorting and its path to commodification can also be seen as irreconcilable with queer political agendas that seek to challenge the use of gender and sexual categories by public and corporate actors altogether. A final subset of articles reversed this logic by underscoring how anti-LGBTQ+ attitudes could be used by sorting algorithms to predict conservative affiliations among platform users[19], either in the context of romantic matchmaking or to identify potential voters during electoral campaigns. As such, the instrumentalization of LGBTQ+ issues by digital platforms to predict the personal values of cisheterosexual users deserves more scrutiny, as it uniquely exploits queerness as a metric and not as a predictive output.

3.2 Recommendation algorithms and the remediation of LGBTQ+ cultures

Recommendation algorithms are built to provide platform users with personalized content or ads based on their past preferences or identity markers to maintain a high level of engagement and, by extension, to help platforms generate a profit. As such, they constitute key mechanisms in a system of surveillance capitalism (Zuboff, 2019) and have transformed how LGBTQ+ publics are constituted and reached by advertisers, businesses, and non-profit organizations alike (Sender, 2018). Recommendation algorithms are predictive systems that directly relate to culture and identity, in the sense that what users consume and whom they interact with online co-construct their own social identity (Cheney-Lippold, 2017). The articles that were included in this category address issues that fall in one of these three subsets: the increasing role that recommendation algorithms play in reshaping LGBTQ+ cultures, the algorithmic promotion or radicalization of anti-LGBTQ+ stances, and the consumerist imperatives under which algorithmic ad promotion operates.

[17] Phataranawik (2019).

[18] GeoSure (2018a; 2018b).

[19] Cookson & McCarthy (2013); Goldbeck (2015); Lord & Potter (2015); Pothier (2015).

A first subset of articles examines how algorithmic recommendation systems have become new curators of LGBTQ+ cultures, tastes, and aesthetics. In recent years, TikTok's recommendation algorithms have been particularly discussed for their propensity to help LGBTQ+ users find their communities of peers online, as well as to support the remixing of LGBTQ+ cultural codes[20], as also discussed in the scientific literature (Simpson & Semaan, 2021). In particular, the app's For You page, which offers users a personalized list of trending videos that are tweaked based on their preferences, has received a lot of media attention for its eerie ability to predict users' interests. Instagram's recommendation algorithms have also been identified as key mediators of gay visual cultures and beauty standards, resulting in a new category of "Instagays" that intersects with influencer culture[21]. Other articles have identified how apps are transforming dating and hookup cultures within LGBTQ+ communities[22]. Indeed, the task of romantic or sexual coupling has been increasingly undertaken by matchmaking algorithms that combine recommendation and user rating systems (Myles, 2020).

Recommendation algorithms were also discussed in their propensity to reshape processes of production and circulation within cultural and media industries. For example, some articles underlined how emerging LGBTQ+ musical artists, like Lil Nas X, ElyOtto, and Saucy Santana, have trended by harnessing the power of recommendation algorithms[23]. Some articles addressed how LGBTQ+ journals and magazines have changed in recent years to follow trend hypes imposed by recommendation algorithms to ensure commercial success[24]. Some have highlighted how Netflix's recommendation algorithms can recreate "an experience akin to having a queer owned video store"[25], while others have warned that their predictive power could also unintentionally out users to family members by revealing their viewing habits[26]. Related articles argued that recommendation algorithms play an increasing role in mediating user identities, namely, because of their predictive capability. To that effect, TikTok's recommendation algorithms were particularly identified as being able to predict the sexual and/or gender identity of users before users themselves, with some users realizing or coming to terms with their own queerness

[20] Jeffers (2021); George (2022); Hess (2021); Ohlheiser (2020); Noyce (2022); Savage (2020); Scriver (2021); Wilson (2020).
[21] Davis (2022).
[22] Anonymous (2020b; 2020c; 2020d; 2020e); Hanau (2017); Headero (2021); Heffernan (2021); Light (2010); O'Brien (2016); Wells (2015).
[23] Hardy (2022); Hughes (2021); Hitt (2019); Volmers (2021a; 2021b).
[24] Anonymous (2018b); Czyzselska (2018).
[25] Chilton (2019).
[26] Ellis (2019).

through their interactions with the app's recommended content[27]. Similarly, an article evoked an older controversy where the digital video recorder TiVo would "think" certain users were gay based on their media preferences[28], which echoes studies that have underlined how digital platforms increasingly intersect recommendation algorithms, identity, and queerness (Cohn, 2016).

A second subset of our corpus highlights how recommendation algorithms may participate in radicalizing anti- or pro-LGBTQ+ stances. This relates to the extensive scholarship that has examined (or challenged) the claim that digital platforms can act as echo chambers or filter bubbles (Bruns, 2019). Articles in this subset criticized digital platforms and their quest for virality that allegedly exacerbate forms of LGBTQ-phobias[29]. Platforms like Facebook and YouTube were particularly criticized for operating recommendation systems geared at promoting the most controversial – and often hateful – videos to other users, while failing to consider the implications that those videos may raise in terms of LGBTQ+ safety and discrimination[30]. Similarly, TikTok was criticized for promoting anti-LGBTQ+ videos during the US month of Pride[31]. Articles also underlined how music streaming platforms, like Spotify, Apple, and Deezer, were being investigated after it was disclosed that they were recommending homophobic and racist music to their users[32]. The role of recommendation algorithms in radicalizing political stances was also a key issue. Some articles discussed how recommendation algorithms could exacerbate hate speech against LGBTQ+ politicians, like US Democrat Pete Buttigieg[33], while others posited that recommendation systems played a significant role in sustaining both hate speech campaigns against LGBTQ+ folks and, inversely, "cancel campaigns" against individuals who disagree with LGBTQ+ rights campaigns[34].

A third subset of articles specifically examines the implications that targeted ads raise for LGBTQ+ communities. Articles included in this subset offered critical reflections on the dataveillance and consumerist imperatives to which targeted ads respond and the implications they generate in terms of consumer rights[35]. Those articles often underlined how targeted ads

[27] Bokody (2021); Joho (2022); Kammerer (2021); MacGowan (2020); Oliver (2021); Simpson (2020); Singal (2018).
[28] Kikidis (2019).
[29] Beaty (2019); Godwin (2017).
[30] Balkissoon (2019); Crerar (2021); January (2020).
[31] Colombo (2021); Little (2021); Zakrzewski (2021); Zitser (2021).
[32] Fingas (2020).
[33] Hicklin (2020).
[34] Aberle (2018); Cook (2019); Maddox (2021); Metz (2021); Susarla (2020).
[35] Brown (2016); Murgia (2019); Wood (2018a; 2018b).

could lead to further discrimination against LGBTQ+ communities. For example, some addressed how targeted ad services could be used by conservative religious groups to reach young LGBTQ+ users and promote conversion therapies, as well as the difficulty for digital platforms to implement policies differentiating anti-LGBTQ hate speech from legitimate freedom of religious speech [36]. Other articles have identified how recommendation algorithms could lead to the promotion of targeted anti-LGBTQ+ ads and disinformation during political campaigns[37], as well as highlighted how targeted ads, coupled with automated gender recognition software, could further lead to the misgendering of trans and non-binary users[38].

Facebook's failings in the implementation of clear and safe policies regarding their targeted ads service was the topic of several analyses. On the one hand, because of the sexual and gender datafication schemes they operate, targeted ads raise significant privacy risks for LGBTQ+ users who might not want to have their personal information disclosed to third-party companies; on the other hand, LGBTQ+ users may also be unfairly excluded from these ad services by companies that do not wish to include LGBTQ+ folks in their consumer base, thus further leading to their social discrimination[39]. LGBTQ-themed ads on Facebook were also more prone to be the targets of hateful comments[40]. Additionally, Facebook was criticized by advertisers for allegedly blocking gay-themed ads that were erroneously labelled as sexually explicit, as well as by unhappy businesses whose ads were recommended under LGBTQ-phobic videos[41]. Those controversies came in the wake of various policy and bill propositions seeking to prevent or better oversee the use of sexual orientation and/or gender identity to target platform users (among other protected identity categories), like in the US and in Europe [42]. In response to these controversies, Facebook announced in 2021 that it would likely remove sensitive ad-targeting categories from its service[43].

Overall, our review indicates that LGBTQ+ ad targeting is emerging as a lucrative market, with companies like Gay Ad Network specifically catering to the needs of digital advertisers seeking to reach those publics[44].

---

[36] Asher-Schapiro & Gebeily (2021); Cuthbertson (2018); Horton & Cook (2018); Lidman (2019); Pandey (2018); Rosa (2018); Toce (2018); Whyte (2018).
[37] Buonaiuto (2018); Legon (2021).
[38] Sadagopan (2019).
[39] McIntyre (2019).
[40] Piper (2013).
[41] Dwoskin & Timberg (2017); Rosenberg (2018).
[42] Consumer Watchdog (2020); Dag (2022); Nuttall (2019).
[43] Isaac & Hsu (2021).
[44] Gay Ad Network (2012); Nocera (2014); Qvist (2016).

Other articles identified how new platforms are launched to provide marketers with safe ways to reach LGBTQ+ publics [45] or promoted opportunities for marketers to help them improve their skills to reach niche publics through targeted ads training[46]. Our analysis underlines how the topic of targeted ads is a particularly complex issue. The use of targeted ads by corporations to reach LGBTQ+ publics based on their personal data is especially criticized for its implications in matters of consumer rights and user privacy. That said, these services can also be used by LGBTQ+ non-profit organizations and cultural industries to reach the same publics. Mainstream platforms' decision to remove LGBTQ+ targeted ads could significantly hinder their operations, especially those of smaller organizations or businesses that may not have the resources to build their own LGBTQ+ publics, an issue that deserves additional scrutiny.

3.3 Filtering algorithms and the politics of LGBTQ+ acceptability

Filtering algorithms are typically geared at identifying undesirable content on digital platforms and preventing users from accessing it, either by removing that content or by making it more difficult to find. As such, filtering algorithms are inherently related to platform content moderation strategies (Gerrard & Thornham, 2020; Gillespie, 2018). In our review, articles that addressed the topic of algorithmic filtering in relation to LGBTQ+ issues were generally interested in how digital platforms are becoming new moral authorities that redefine the boundaries of social acceptability. The articles that were included in this category are divided into two interrelated subsets: those addressing the need for platforms to better support the algorithmic detection and removal of anti-LGBTQ+ hate speech and those highlighting, somewhat paradoxically, how those content moderation mechanisms often censor or harm the LGBTQ+ users they are supposed to protect (Cobbe, 2020).

A first subset of articles addresses the contentious topic of anti-LGBTQ+ hate speech on digital platforms. Several articles shared the reports and studies prepared by human rights agencies and organizations that highlight the need for digital platforms, like Facebook and Twitter in particular, to develop better filtering algorithms in the hope of mitigating LGBTQ+ harassment online. Those organizations include GLAAD, the anti-defamation league, Media Matters for America, the United Nations, and the California Attorney General[47]. Their reports identify LGBTQ+ users, among

[45] Grier (2020).

[46] Mckelvey (2015).

[47] Collins (2021); Fried (2021); Guynn (2019); Monifa (2021); TNS (2019); TNS (2020a); Zitser (2021).

other vulnerable populations, as being disproportionally targeted by hate speech and argue that those users deserve safe spaces to interact online. Articles were particularly critical of Facebook's alleged laxity in the matter of LGBTQ+ hate speech detection[48], with certain celebrities, like singer Elton John, inviting users to boycott the platform in 2018[49]. Facebook was also criticized for introducing its "real name" policy in 2014, which negatively impacted LGBTQ+ communities, and drag artists in particular, who often employ online alter egos[50]. Furthermore, the platform was criticized for its alleged unwillingness to follow local laws and its inability to protect LGBTQ+ users outside of Western countries but was also celebrated for its ban on targeted conversion therapy ads[51].

Elsewhere, the video-sharing platform YouTube came under heavy criticism when it decided to reinstate the channel of a homophobic creator in 2019, opting to demonetize his channel rather than banning it altogether[52]. Similarly, Twitter was identified as a hotbed of anti-LGBTQ+ discourse[53], whereas Google's and TikTok's[54] efforts to combat hate speech online received more positive responses[55]. The failures of various types of platforms in protecting their LGBTQ+ users from hate speech were also raised through cases relating to the blackmail of North African queer dating app users, the increase of LGBTQ-phobic trolls on Instagram and Facebook, and the presence of homophobic bullying on gaming and sports platforms[56]. Overall, these articles highlight the prevalence of anti-LGBTQ+ hate speech online and the limitations of content moderation policies – and their related filtering algorithms – in removing discriminatory or threatening content, as well as call for more effective automated hate speech removal systems.

A second subset of articles examines the limitations and dangers associated with the implementation of content moderation policies per se. Notably, this second subset documents the flaws of AI-based detection systems that have had the tendency to censor LGBTQ+ content in favor of more 'family-friendly' content[57]. Several articles addressed how sex and

---

[48] Dwoskin et al. (2021); Foufas (2014); Hindman et al. (2022); Kuchler (2018); Scott (2021); Seetharaman et al. (2021); The Washington Post (2021); TNS (2020b).

[49] Naffi & Davidson (2018).

[50] Allen (2015); Foufas (2014); Michaelson (2014a; 2014b); Rushton (2014a).

[51] Al-Khal (2020); Asher-Shchapiro & Gebeily (2021); Fae (2017); Roeder (2020); Rosa (2018); Whyte (2018).

[52] Farrell (2019); Fitzgerald Rodriguez (2019); Gilbert (2019); Kinney (2019); McCarthy (2019); Sharman (2019); Weill (2019a); Yurieff (2019).

[53] Desmarais (2019).

[54] Prang (2022).

[55] Besen (2019); LaMagna (2018); Smith (2017).

[56] Al-Khal (2020); Chang (2021); Farrell (2021); Hardy (2022); Jammot (2021); Levesque (2020); O'Leary (2012); Porter (2021); Ryall (2021).

[57] Snow (2022); York (2021).

nudity bans on digital platforms, as well as age verification policies have become very contentious issues[58]. This is best illustrated by the so-called Tumblr and OnlyFans purges, during which both platforms announced they would ban adult content in an effort to guarantee the safety of younger users, but with the collateral consequence of banning the profiles of thousands of LGBTQ+ creators who contributed to the early popularity of those platforms (OnlyFans later rescinded this decision)[59]. The most prevalent cases of LGBTQ+ censorship by content moderation policies, however, related to the restricted mode implemented by YouTube in 2017, which seemingly affected the ability for LGBTQ+ users to monetize their videos[60]. Similarly, a second controversy arose in 2019 when a group of LGBTQ+ YouTubers sued the platform, alleging that its algorithms were disproportionally demonetizing videos pertaining to LGBTQ+ subjects or keywords[61], as those videos would stop being suggested by the platform. While this controversy technically relates to recommendation algorithms, it was included in this section because those algorithms incidentally function as a content moderation mechanism (Gillespie, 2018).

Other controversies surrounding the issue of content moderation and censorship included Facebook preventing lesbian users to employ the word 'dyke', TikTok censoring the use of certain LGBTQ+ keywords in its Disney filter, as well as Twitter's propensity to disproportionally flag the profiles of drag artists[62]. Furthermore, Twitter and Instagram were criticized for enacting content moderation policies that disproportionally censor and penalize trans and non-binary users[63], whereas review platforms, like Yelp and Zomato, were criticized for providing transphobic labels or censoring reviews retelling homophobic experiences[64]. Finally, the use of deepfake technology to protect the identities of LGBTQ+ participants in a documentary on Chechnya also garnered reflections on the inability of digital platforms to moderate this emerging technology[65].

---

[58] Anonymous (2019); Belton (2021); Blue (2019); Drake (2018); He (2019); Robertson (2021); Turban (2018).

[59] Beachum (2021); Heater (2018); Mitchell & Akhtar (2021); Perez (2017); Powell (2018); Robertson (2022).

[60] Allen (2017a); Cooper (2017); Lewis (2017); Ogles (2017); Ryan (2017); Shu (2017); Whilem (2017).

[61] Banerjee (2019); Bensinger & Albergotti (2019); Dodgson (2019); Ellis (2019); Farokhmanesh (2018); Fisher (2019); Forrester (2020); Guilfoil (2019); Hutton (2019); Kleeman (2019); Lothian-McLean (2019); Murdock (2019); Oremus (2019); Romano (2019); Steiner (2019); Stokel-Walker (2018), TNS (2020c).

[62] Sharpe (2017); Smith (2021); Trigger (2017).

[63] Savage (2021); Smith (2018); Walker (2019); Weill (2019b).

[64] Allen (2018a); Ramnani (2017).

[65] Coyle (2020); Needham (2020); Rothkopf (2020).

Overall, this section highlights how changes in content moderation policies, especially those targeting 'inappropriate' content, negatively impact the online spaces built for LGBTQ+ socializing, as previously discussed in the scientific literature (Byron, 2019; Pilipets & Paasonen, 2020; Southerton et al., 2020). Indeed, platform policies often fail to protect their LGBTQ+ users and can even participate in censoring the content these users produce (Duguay et al., 2020; Oliva et al., 2020). This paradox was the topic of several articles included in this section[66]. In response, start-up companies are marketing new social media platforms around this paradox, arguing that their services are expressly developed to provide LGBTQ+ users with safe spaces while also preventing the censorship of their content[67]. Thus, the failure of mainstream platforms in safely catering to their LGBTQ+ users appears to have become somewhat of a marketing opportunity for LGBTQ+ tech entrepreneurs.

3.4 Search algorithms and the struggle over LGBTQ+ visibility online

Search algorithms are typically built to help users navigate the Internet by offering a ranked or ordered lists of results in response to their specific queries. To that effect, search engines have become a central topic of interest in the scientific literature on algorithmic oppression, especially as they can reproduce supremacist and cisheteronormative conceptions about sexuality, gender, and race (Noble, 2018). The articles in this fourth category generally discuss case studies and controversies showcasing how the Internet has become the site of significant struggles over LGBTQ+ self-expression and mediated visibility.

Unsurprisingly, Google emerged as the main topic of discussion in this category, with a specific interest in the racial, sexual, and gender biases that can be reproduced by the company's proprietary search engines[68]. Google was particularly criticized for its tendency to sexualize and fetishize search results on lesbians and for catering to the interests and taste of white, heterosexual men, while preventing LGBTQ+ users – and queer women in particular – from obtaining search results that meet their individual or collective needs (Google tweaked its search algorithms to that effect in 2019)[69]. Google also came under heavy criticism in 2017 when it was disclosed that an AI bot it developed, which relied on sentiment analysis,

[66] Allen (2018b); Christy, (2018); Garland (2019); Lidman (2019); O'Neill (2019); Vecchietti (2020).

[67] ETX Studio (2022); Grier (2020); Rushton (2014a; 2014b).

[68] Davis (2014).

[69] Bondioli (2019); Burkholder (2019); Ehrenkranz (2019); Marr (2019); Parkinson (2019); Ritschel (2019); Simister (2019).

had learned to automatically label LGBTQ+ keywords as negative[70]. Other articles discussed how Google's autocomplete function, which offers users suggestions to complete their queries based on the most popular searches, reproduced discriminatory assumptions about LGBTQ+ people[71]. Similarly, other controversies emerged after learning that the natural language processing models upon which several of Google's online tools operate often filter out or minoritize gender and sexual diverse voices[72].

Other articles addressed Google's search algorithms at the intersection of politics and policy. In 2011, several articles reported how Rick Santorum, an ultra-conservative US politician, demanded that Google remove undesirable search results associated with his name[73]. This news story relates to a 2003 controversy when the Republican party member compared homosexuality to incest and bestiality. In response, the American journalist Dan Savage organized an online campaign with LGBTQ+ rights supporters to manipulate Google's search results so that queries related to the politician's name resulted in an explicit sexual term, a political tactic now known as Google bombing (Gillespie, 2017). Other articles addressed the ethical intricacies for Google to comply or not with anti-LGBTQ+ policies in countries like Russia, Turkey and Thailand, as well as the implications of the corporation's decisions in matters of free speech[74]. Another controversy came in the wake of the 2017 Australian Marriage Law Postal Survey, when Google was accused of censoring anti-marriage equality results in its search engine and, thus, of being biased in favor of marriage equality[75], an accusation that was also made against Facebook[76].

To conclude, algorithmic searchability also directly relates to the labels and hashtags that facilitate and accelerate the provision of meaningful search results through the use of predetermined sociotechnical categories. For example, the introduction of new LGBTQ+ labels on the gaming platform Steam in 2019 was seen as a useful addition for users in search of more diverse games[77], whereas Twitter's perplexing decision to prevent the search of certain keywords associated with bisexuality was condemned[78]. Interestingly, other articles used LGBTQ+ search results on adult streaming platforms, like Pornhub, to demonstrate how sexual and gender diverse

[70] Jackman (2017); Lumb (2017); Morse (2017); Thompson (2017); Wood (2018).
[71] Cain Miller (2015); Quinn (2016); Morozov (2012); Nicas (2017).
[72] Anderson (2021); Winchel (2018).
[73] Cohen (2012); Goodman (2012); Krieg (2016); Rozsa (2011).
[74] Gordon Crovitz (2015).
[75] Davidson (2017); Fox Koob (2017); Tomlinson (2018).
[76] Mitchell (2016).
[77] Grayson (2019).
[78] Allen (2017b).

citizens exist globally, thus refuting the conservative claim that LGBTQ+ people are the exclusive products of Western societies[79].

Together, the articles included in this section raise important questions as to who possesses the legitimacy to shape and control LGBTQ+ representations online amid new algorithmic regimes of visibility (Bucher, 2018). Articles highlighting sexualized, stereotypical, and negatively biased search results also reflect queer scholars' assertion that visibility can be a double-edged sword for LGBTQ+ people (Barnhurst, 2007). In the past decade, algorithmic searchability has emerged as a key concept, particularly in the hashtag activism scholarship (Khoja-Moolji, 2015; Yang, 2016). This is also true for LGBTQ+ communities, for which the hashtag feature has supported the rapid constitution of highly visible and searchable LGBTQ+ publics that are leveraged to formulate sets of cultural (Navar-Gill & Stanfill, 2018) and/or political demands (Duguay, 2016; Myles & Lewis, 2019; Woods & McVey, 2016). That said, articles pertaining to LGBTQ+ hashtags have not emerged as a significant subset in our review. This points to some of the methodological limitations of our review that are discussed in this papers' conclusion.

3.5 Algorithmic innovation and the reproduction of cisheteronormative biases

The last subset of our corpus deals with the overarching issue of representation in the tech industry and its role in reproducing gender and sexual biases through the digital platforms that stem from it. The articles included in this category offer meta-commentaries to understand the origins of algorithmic oppression and its implications for LGBTQ+ users. In line with the scholarship on algorithmic justice (Noble, 2018), some articles identified Silicon Valley's toxic technoculture (Massanari, 2017) as a main source of cisheteronormative biases[80], which ultimately materialize into platform affordances and result in LGBTQ+ discrimination[81].

Two related controversies arose in this subset: the resignation of Mozilla's ex-CEO over homophobic remarks in 2014[82], as well as the nomination of a seemingly transphobic individual to sit on Google's newly formed AI Council in 2019[83]. To that end, Google was particularly criticized for the alleged mistreatment of its LGBTQ+ employees[84]. Despite the

[79] Doig (2014); Foxton (2014); Gupta (2017).
[80] Metz (2021); Norton (2021); Segal (2021); TNS (2021); Watling (2021); Webb (2017).
[81] Allen (2018); Bahl (2022).
[82] Tan (2014).
[83] AFP (2019); Chapman (2019); Harris (2019); Hatmaker (2019); Metz (2019); Risi (2019); Sumagaysay (2019).
[84] Daily Telegraph (2019); Grind & MacMillan (2018).

adoption of equity, diversity and inclusion policies, the under-representation of LGBTQ+ workers in STEM research and the tech industry was identified as a key obstacle to reach algorithmic justice[85]. Hence, some articles discussed the necessity to increase LGBTQ+ hires in the tech industry[86] or stressed the importance for LGBTQ+ professionals to build their own AI-related skills[87]. This relates to recurrent calls to diversify the tech industry, as similarly argued by Crawford (2016b): "artificial intelligence [reflects] the values of its creators. So, inclusivity matters — from who designs it to who sits on the company boards and which ethical perspectives are included. Otherwise, we risk constructing machine intelligence that mirrors a narrow and privileged vision of society, with its old, familiar biases and stereotypes".

As such, this last subset – the smallest of our corpus – highlights how most of the newspaper articles included in our review tend to focus on downstream considerations, as they generally contend with the effects that platform algorithms have on the lives of LGBTQ+ users after their implementation. This points to the importance for researchers, activists, and journalists alike to also engage with more upstream concerns when dealing with LGBTQ+ algorithmic justice, namely, by considering digital platforms as culturally situated artefacts that reflect and exacerbate the power imbalances of the industrial sector from which they spawn.

4 CONCLUSION

In this paper, we conducted a review and analysis of journalistic articles to identify significant issues that platform algorithms raise for LGBTQ+ users and analyze their implications, especially in terms of social justice and equity. The results of our review were divided into four algorithmic functions (i.e., sorting, recommending, filtering, and searching), as well as a fifth meta-category relating to cisheteronormativity in the tech industry more broadly. In light of our analysis, our paper points to five important issues that platform algorithms raise for LGBTQ+ communities that merit additional scrutiny: the ability for sorting algorithms to identify, categorize, and predict the sexual orientation and/or gender identity of users; the role that recommendation algorithms play in mediating LGBTQ+ identities, kinship, and cultures; the development of automated anti-LGBTQ+ speech detection/filtering software and the collateral harm caused to LGBTQ+ users; the power struggles over the nature and effects of visibility afforded

[85] Flickinger (2018); Sahoo (2021); Telford (2020).
[86] Collins (2021).
[87] Croxon (2019).

to LGBTQ+ issues/people online; and the overall enactment of cisheteronormative biases through platform affordances.

Our press review presented some methodological limitations mainly associated with the use of keyword searches to identify relevant controversies. First, our review is limited to articles containing specific LGBTQ+ keywords. Evidently, AI-related issues that can be applied to the general population also apply to LGBTQ+ communities but may not have been identified by our review process. Second, the use of predetermined keywords represented a challenge to encapsulate dynamic or contested concepts (e.g., 'algorithms', 'digital platforms', 'LGBTQ+ communities'). Third, our review process did not allow us to fully engage in an intersectional analysis of the literature, that is, to examine sexual and gender identity at the intersection of other equally important social identity markers such as race, disability, and/or social class. Fourth, our review was limited to newspaper articles published in English, a methodological decision that diminishes the scope of our results. While our analysis included articles from English-language newspapers from the Americas, the Middle East, Europe, Africa, Asia, and Oceania, its results and their implications should in no way be understood as universal, given this linguistic limitation.

Still, our review points to key avenues for future research. Researchers must develop innovative empirical studies to account for the various forms of algorithmic oppression LGBTQ+ users experience on and by digital platforms. Importantly, future studies should examine how platform algorithms affect and oppress LGBTQ+ users in differentiated ways, depending on a variety of intersecting and complex factors (e.g., race, disability, social class, religion, age, local cultures, indigeneity, geopolitical contexts, legal jurisdictions), in line with studies that apply intersectional theories to digital research objects (Geerts & Rahbari, 2022). Furthermore, our review shows how platform algorithms have so far been apprehended as being predominantly oppressive in nature toward LGBTQ+ users. Beyond algorithmic oppression or harm (Tufekci, 2015), future research should consider how platform algorithms are not solely oppressive but also productive in nature (Bucher, 2018). Indeed, algorithmic surveillance does not solely repress LGBTQ+ identities, kinship, and cultures or prevent them from emerging in the first place; rather, it shifts and reshapes them in differentiated and complex ways.

To that end, more research is needed to examine platform algorithms beyond their constraining properties to understand how they are actively reshaping the lives and practices of LGBTQ+ users, while simultaneously unpacking the capitalistic and surveillance imperatives to which these algorithms respond. Importantly, future research should account for the agency LGBTQ+ users exert vis-à-vis platform algorithms. For example, in

what creative ways do LGBTQ+ users resist, counter, or hijack platform algorithms (Grison & Julliard, 2021)? Finally, our review stresses the need for more industry-related investigations into the underrepresentation of LGBTQ+ professionals in positions of power within the tech industry, and how this may contribute to the development of cisheteronormative algorithmic systems. Together, these recommendations lay the foundation of a research agenda on algorithmic justice that can steer future technological innovation and policy, as well as guide research on LGBTQ+ communities and other marginalized communities facing similar challenges online.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

This research has received funding from the Social Sciences and Humanities Research Council (SSHRC) of Canada.

REFERENCES

Abreu, R. L., & Kenny, M. C. (2018). Cyberbullying and LGBTQ youth: A systematic literature review and recommendations for prevention and intervention. *Journal of Child & Adolescent Trauma*, *11*(1), 81-97.

Adam, B. D. (2015). Homophobia and heterosexism. In Ritzer, G. *The Blackwell Encyclopedia of Sociology*, Blackwell Publishing, 2154-2157.

Albury, K., Dietzel, C., Pym, T., Vivienne, S., & Cook, T. (2020). Not your unicorn: Trans dating app users' negotiations of personal safety and sexual health. *Health Sociology Review*, 1-15.

Ananny, M. (2016). Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Science, Technology, & Human Values*, *41*(1), 93-117.

Anderson, A., & Roth, A. L. (2020). Queer erasure: Internet browsing can be biased against LGBTQ people, new exclusive research shows. *Index on Censorship*, *49*(1), 75-77.

Ayoub, P. M., & Brezinska, O. (2015). Caught in a web? The Internet and deterritorialization of LGBT activism. In Paternotte, D. & M. Tremblay (Eds.), *The Ashgate Research Companion to Lesbian and Gay Activism*, New York: Routledge, 225-243.

Baker, P., & Potts, A. (2013). 'Why do white people have thin lips?' Google and the perpetuation of stereotypes via auto-complete search forms. *Critical discourse studies*, *10*(2), 187-204.

Barnhurst, K. (2007). Visibility as paradox: Representation and simultaneous contrast. In K. Barnhurst (Ed.), *Media Queered: Visibility and its Discontents,* pp. 1-22. Peter Lang.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, *11*(6), 985-1002.

Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. John Wiley & Sons.

Bivens, R. (2017). The gender binary will not be deprogrammed: Ten years of coding gender on Facebook. *New Media & Society*, *19*(6), 880-898.

Bivens, R., & Haimson, O. L. (2016). Baking gender into social media design: How platforms shape categories for users and advertisers. *Social Media+ Society*, *2*(4), DOI: 2056305116672486.

Blackwell, L., Dimond, J., Schoenebeck, S., & Lampe, C. (2017). Classification and its consequences for online harassment. *Proceedings of the ACM on Human-Computer Interaction*, *1*, 1-19.

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge: MIT press.

Bruns, A. (2019). *Are filter bubbles real?*. John Wiley & Sons.

Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford: Oxford University Press.

Bucher, T., & Helmond, A. (2018). The Affordances of Social Media Platforms. In Sloane, L. & A. Quan-Haase. (Eds), *SAGE Handbook of social media,* SAGE, 233-253.

Butler, J. (1990). *Feminism and the Subversion of Identity*. New York: Routledge.

Byron, P. (2019). 'How could you write your name below that?' The queer life and death of Tumblr. *Porn Studies*, *6*(3), 336-349.

Campbell, J. E. (2005). Outing PlanetOut: surveillance, gay marketing and internet affinity portals. *New Media & Society*, *7*(5), 663-683.

Cardon, D. (2015). *A quoi rêvent les algorithmes. Nos vies à l'heure: Nos vies à l'heure des big data*. Paris: Le Seuil.

Chartier, L. (2003). *Mesurer l'insaisissable: Méthode d'analyse du discours de presse*. PUQ.

Cheney-Lippold, J. (2017). We are data. In *We Are Data*. New York University Press.

Cho, A. (2018). Default publicness: Queer youth of color, social media, and being outed by the machine. *New Media & Society*, *20*(9), 3183-3200.

Cobbe, J. (2020). Algorithmic censorship by social platforms. *Philosophy & Technology*, 1-28.

Cohn, J. (2016). My TiVo thinks I'm gay: Algorithmic culture and its discontents. *Television & New Media*, *17*(8), 675-690.

Crawford, K. (2016a). Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology, & Human Values*, *41*(1), 77-92.

Crawford, K. (2016b). Artificial Intelligence's White Guy Problem. *The New York Times.* https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html?_r=0

Crawford, K., & Gillespie, T. (2016). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society*, *18*(3), 410-428.

Daniels, J. (2013). Race and racism in Internet studies: A review and critique. *New Media & Society*, *15*(5), 695-719.

de Filippi, P. (2016). Gouvernance algorithmique: Vie privée et autonomie individuelle à l'ère des Big Data. In P. De Filippi & D. Bourcier (Eds.), *Open Data & Data Protection : Nouveaux défis pour la vie privée*. Paris: Mare & Martin, 1-22.

DeNardis, L., & Hackl, A. M. (2016). Internet control points as LGBT rights mediation. *Information, Communication & Society*, *19*(6), 753-770.

Duguay, S. (2016). Constructing public space|"legit can't wait for# toronto# worldpride!": Investigating the twitter public of a large-scale LGBTQ festival. *International Journal of Communication*, *10*, 25.

Duguay, S. (2016). Investigating the Twitter public of a large-scale LGBTQ festival. *International Journal of Communication*, *10*, 274-298.

Duguay, S. (2019). "Running the Numbers": Modes of Microcelebrity Labor in Queer Women's Self-Representation on Instagram and Vine. *Social Media+ Society*, *5*(4), DOI: 2056305119894002.

Duguay, S., Burgess, J., & Suzor, N. (2020). Queer women's experiences of patchwork platform governance on Tinder, Instagram, and Vine. *Convergence*, *26*(2), 237-252.

Elipe, P., de la Oliva Muñoz, M., & Del Rey, R. (2018). Homophobic bullying and cyberbullying: Study of a silenced problem. *Journal of Homosexuality*, *65*(5), 672-686.

Evans, D. T. (2007). Sexual citizenship. In Ritzer, G. (Ed.), *The Blackwell Encyclopedia of Sociology*. Hoboken: Wiley-Blackwell.

Ferris, L., & Duguay, S. (2020). Tinder's lesbian digital imaginary: Investigating (im)permeable boundaries of sexual identity on a popular dating app. *New Media & Society*, *22*(3), 489-506.

Foucault, M. (1975). *Surveiller et punir*. Gallimard.

Foucault, M. (1978). *Histoire de la sexualité, volume 1: La Volonté de savoir*. Gallimard.

Fredenburg, J. N. (2020). *YouTube as an Ally of Convenience: The Platform's Building and Breaking with the LGBTQ+ Community*, Georgetown University.

Friedman, B., & Kahn Jr, P. H. (2003). Human values, ethics, and design. In Jacko, J.A. (Ed.), *The Human-Computer Interaction Handbook*, Boca Raton: CRC Press, 1177-1201.

Garcia, M. (2016). Racist in the machine: The disturbing implications of algorithmic bias. *World Policy Journal*, *33*(4), 111-117.

Geerts, E., & Rahbari, L. (2022). Gender, Sexuality, and Embodiment in Digital Spheres. Connecting Intersectionality and Digitality. *Journal of Digital Social Research*, *4*(3).

Gerrard, Y., & Thornham, H. (2020). Content moderation: Social media's sexist assemblages. *new media & society*, *22*(7), 1266-1286.

Giffney, N., & O'Rourke, M. (Eds.). (2016). *The Ashgate Research Companion to Queer Theory*. New York: Routledge.

Gillespie, T. (2010). The politics of platforms. *New Media & Society*, *12*(3), 347-364.

Gillespie, T. (2012). Can an algorithm be wrong? *Limn*, *1*(2), n.p.

Gillespie, T. (2017). Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem. *Information, Communication & Society*, *20*(1), 63-80.

Gillespie, T. (2018). *Custodians of the Internet*. Yale University Press.

Gillespie, T., & Seaver, N. (2016). Critical algorithm studies: A reading list. *Social Media Collective*.

Gray, M. L. (2009). "Queer Nation is dead/long live Queer Nation": The politics and poetics of social movement and media representation. *Critical Studies in Media Communication*, *26*(3), 212-236.

Grison, T., & Julliard, V. (2021). Les enjeux de la modération automatisée sur les réseaux sociaux numériques: les mobilisations LGBT contre la loi Avia. *Communication, technologies et développement*, (10).

Hanckel, B., Vivienne, S., Byron, P., Robards, B., & Churchill, B. (2019). 'That's not necessarily for them': LGBTIQ+ young people, social media platform affordances and identity curation. *Media, Culture & Society*, *41*(8), 1261-1278.

Helmond, A. (2015). The platformization of the web: Making web data platform ready. *Social Media+ Society*, *1*(2), 1-11.

Hoffmann, A. L. (2019). Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, *22*(7), 900-915.

Introna, L. D. (2016). Algorithms, governance, and governmentality: On governing academic writing. *Science, Technology, & Human Values*, *41*(1), 17-49.

Kafer, G., & Grinberg, D. (2019). Queer Surveillance. *Surveillance & Society*, *17*(5), 592-601.

Khoja-Moolji, S. (2015). Becoming an "intimate publics": Exploring the affective intensities of hashtag feminism. *Feminist media studies*, *15*(2), 347-350.

Kokas, A. (2022). Data Trafficking and the International Risks of Surveillance Capitalism: The Case of Grindr and China. *Television & New Media*, 15274764221137250.

Latzer, M., Hollnbuchner, K., Just, N., & Saurwein, F. (2016). The economics of algorithmic selection on the Internet. In Bauer, J. & M. Latzer (Eds.), *Handbook on the Economics of the Internet*, Edward Elgar Publishing.

Lingel, J. (2020). Dazzle camouflage as queer counter conduct. *European Journal of Cultural Studies*, 1367549420902805.

Lingel, J., & Golub, A. (2015). In face on Facebook: Brooklyn's drag community and sociotechnical practices of online communication. *Journal of Computer-Mediated Communication*, *20*(5), 536-553.

Lingiardi, V., Carone, N., Semeraro, G., Musto, C., D'amico, M., & Brena, S. (2019). Mapping Twitter hate speech towards social and sexual minorities: A lexicon-based approach to semantic content analysis. *Behaviour & Information Technology*, 1-11.

Lucero, L. (2017). Safe spaces in online places: Social media and LGBTQ youth. *Multicultural Education Review*, *9*(2), 117-128.

Lyon, D. (2002). Surveillance Studies: Understanding visibility, mobility and the phenetic fix. *Surveillance & society*, *1*(1), 1-7.

Lyon, D. (2005). Surveillance as social sorting: Computer codes and mobile bodies. In *Surveillance as social sorting* (pp. 27-44). Routledge.

MacKenzie, D., & Wajcman, J. (1985). *The social shaping of technology*, Buckingham: Open University Press.

Mainardi, A., & Pavan, E. (2020). LGBTQI Online. In Bachmann, I., Cardo, V., Moorti, S., Scarcelli, C.M. & K. Ross (Eds.), *The International Encyclopedia of Gender, Media, and Communication*, Hoboken: Wiley-Blackwell, 1-8.

Marciano, A., & Antebi-Gruszka, N. (2020). Offline and online discrimination and mental distress among lesbian, gay, and bisexual individuals: The moderating effect of LGBTQ Facebook use. *Media Psychology*, DOI: 10.1080/15213269.2020.1850295

Marres, N. (2015). Why map issues? On controversy analysis as a digital method. *Science, Technology, & Human Values*, *40*(5), 655-686.

Massanari, A. (2017). #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, *19*(3), 329-346.

Mendes, K., Ringrose, J., & Keller, J. (2018). # MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236-246.

Milan, S. (2015). When algorithms shape collective action: Social media and the dynamics of cloud protesting. *Social Media+ Society*, *1*(2), 2056305115622481.

Musiani, F. (2013). Governance by algorithms. *Internet Policy Review*, *2*(3), 1-8.

Musiani, F. (2018). L'invisible qui façonne. Études d'infrastructure et gouvernance d'Internet. *Tracés. Revue de Sciences humaines*, (35), 161-176.

Myles, D. (2022). Grindr? It's a 'Blackmailer's Goldmine'! The Weaponization of Queer Data Publics amid the US-China Trade Conflict. *Sexualities: Special Issue on Sexual Datafication.*

Myles, D. (2020). Les rencontres amoureuses et sexuelles au temps des algorithmes: Une analyse comparative de Grindr et Tinder. In Piazzesi, C., Blais, M., Lavigne, J. & C. Lavoie Mongrain (Eds), *Intimités et sexualités contemporaines: changements sociaux, transformations des pratiques et des representations*, University of Montreal Press, 73-90.

Myles, D., & Lewis, K. (2019). Constructing Injustice Symbols in Contemporary Trans Rights Activisms. *Kvinder, Køn & Forskning*, (3-4), 24-42.

Napoli, P. M. (2013). The algorithm as institution: Toward a theoretical framework for automated media production and consumption. In *Media in Transition Conference*, Fordham University Schools of Business, Massachusetts Institute of Technology, 1-36.

Nash, C. J. (2013). The age of the "post-mo"? Toronto's gay village and a new generation. *Geoforum*, *49*, 243-252.

Navar-Gill, A., & Stanfill, M. (2018). "We shouldn't have to trend to make you listen": Queer Fan Hashtag Campaigns as Production Interventions. *Journal of Film and Video*, *70*(3-4), 85-100.

Nissenbaum, H. (2005). Values in technical design. In Mitcham, C. (Ed.), *Encyclopedia of Science, Technology, and Ethics*. London: MacMillan, 66-70.

Noble, S. U. (2018). *Algorithms of oppression*. New York University Press.

Oliva, T. D., Antonialli, D. M., & Gomes, A. (2020). Fighting Hate Speech, Silencing Drag Queens? Artificial Intelligence in Content Moderation and Risks to LGBTQ Voices Online. *Sexuality & Culture*, 1-33, DOI:10.1007/s12119-020-09790-w.

Phillips, D. J., & Cunningham, C. (2007). Queering surveillance research. *Queer Online: Media Technology and Sexuality*, 31-44.

Pilipets, E., & Paasonen, S. (2020). Nipples, memes, and algorithmic failure: NSFW critique of Tumblr censorship. *New Media & Society*, 1461444820979280.

Pullen, C., & Cooper, M. (Eds.). (2010). *LGBT identity and online new media*. London: Routledge.

Ramirez, J. L., Gonzalez, K. A., & Galupo, M. P. (2018). "Invisible during my own crisis": Responses of LGBT people of color to the Orlando shooting. *Journal of homosexuality*, *65*(5), 579-599.

Renninger, B. J. (2018). Grindr killed the gay bar, and other attempts to blame social technologies for urban development. *Journal of Homosexuality*, *66*(12), 1736-1755.

Richards, L., & Morse, J. M. (2012). *Readme First for a User's Guide to Qualitative Methods*. London: Sage.

Richardson, D. (2017). Rethinking sexual citizenship. *Sociology*, *51*(2), 208-224.

Robards, B. J., Churchill, B., Vivienne, S., Hanckel, B., & Byron, P. (2018). Twenty years of 'cyberqueer': The enduring significance of the Internet for young LGBTIQ+ people. In Aggleton, P. et al. (Eds.), *Youth, Sexuality and Sexual citizenship*, Routledge, 151-167.

Schram, B. (2019). Accidental orientations: Rethinking queerness in archival times. *Surveillance & Society*, *17*(5), 602-617.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, *4*(2), DOI: 2053951717738104.

Sender, K. (2018). The gay market is dead, long live the gay market: From identity to algorithm in predicting consumer behavior. *Advertising & Society Quarterly*, *18*(4).

Simpson, E., & Semaan, B. (2021). For You, or For 'You'? Everyday LGBTQ+ Encounters with TikTok. *Proceedings of the ACM on Human-Computer Interaction*, *4*, 1-34.

Sismondo, S. (2010). *An Introduction to Science and Technology Studies* (Vol. 1). Chichester: Wiley-Blackwell.

Southerton, C., Marshall, D., Aggleton, P., Rasmussen, M. L., & Cover, R. (2020). Social media, content classification and LGBTQ sexual citizenship. *New Media & Society*, *23*(5), 920-938.

Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, *43*(3), 377-391.

Suchman, L. (2008). Feminist STS and the Sciences of the Artificial. In Hackett, E.J., Amsterdamska, O., Lynch, M., Wajcman, J. & A. Giddens. *The Handbook of Science and Technology Studies*, New York: Sage, 139-164.

Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colo. Tech. LJ*, *13*, 203.

van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & society*, *12*(2), 197-208.

Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, *34*(1), 143-152.

Werbin, Kenneth C., Mark Lipton, and Matthew J. Bowman. "The contextual integrity of the closet: Privacy, data mining and outing Facebook's algorithmic logics." *Queer Studies in Media & Popular Culture* 2.1 (2017): 29-47.

Wilkinson, W. W., & Berry, S. D. (2020). Together they are Troy and Chase: Who supports demonetization of gay content on YouTube? *Psychology of Popular Media*, *9*(2), 224.

Wood, D. M., & Monahan, T. (2019). Platform surveillance. *Surveillance & society*, *17*(1/2), 1-6.

Woods, H., & McVey, J. A. (2016). # BlackLivesMatter as A Case Study in the Politics of Digital Media: Algorithms, Hashtag Publics, and Organizing Protest Online. *Teaching Media Quarterly*, *4*(1).

Yang, G. (2016). Narrative agency in hashtag activism: The case of# BlackLivesMatter. *Media and communication*, *4*(4), 13.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama's books of 2019*. Profile books.