



Article

Restricted modes: Social media, content classification and LGBTQ sexual citizenship

new media & society
1–19

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1461444820904362

journals.sagepub.com/home/nms



Clare Southerton 
UNSW Sydney, Australia

Daniel Marshall
Deakin University, Australia

Peter Aggleton
UNSW Sydney, Australia

Mary Lou Rasmussen
Australian National University, Australia

Rob Cover
RMIT University, Australia

Abstract

In the context of recent controversies surrounding the censorship of lesbian, gay, bisexual, transgender and queer online content, specifically on YouTube and Tumblr, we interrogate the relationship between normative understandings of sexual citizenship and the content classification regimes. We argue that these content classification systems and the platforms' responses to public criticism both operate as norm-producing technologies, in which the complexities of sexuality and desire are obscured in order to cultivate notions of a 'good' lesbian, gay, bisexual, transgender or queer

Corresponding author:

Clare Southerton, Centre for Social Research in Health and Social Policy Research Centre, Faculty of Arts & Social Sciences, UNSW Sydney, Goodsell Building, Sydney, NSW 2052, Australia.

Email: c.southerton@unsw.edu.au

sexual citizen. However, despite normative work of classification seeking to distinguish between sexuality and sex, we argue that the high-profile failures of these classification systems create the conditions for users to draw attention to, rather than firm, these messy boundaries.

Keywords

Censorship, content classification, LGBTQ, platform governance, restricted mode, safe mode, Tumblr, YouTube

Introduction: on social media and queer utopias

With many lesbian, gay, bisexual, transgender and queer (LGBTQ) young people identifying social media platforms like YouTube and Tumblr as prominent spaces in the formation of their sexual identity (Cho, 2015, 2017; Wuest, 2014), these communities have been identified as key sites of queer¹ expression (Byron and Robards, 2017; Ciechalski, 2017; Duguay, 2014; Robards et al., 2019; Wargo, 2015). Recent controversies surrounding the censorship of LGBTQ content on these sites (Castello, 2017; Hunt, 2017; Perez, 2017) have challenged these narratives, revealing underlying governance mechanisms that were, at best, indifferent to queer communities, or, at worst, hostile to various forms of their expression. With an eye to examining how young people craft understandings of themselves as sexual citizens from within the discursive and material environments they inhabit, this article examines these restrictions and classifications of social media content and the way these policies construct normative sexual citizenship.²

In general terms, sexual citizenship can be understood as designating diverse ‘sexual claims of belonging’ (Aggleton et al., 2018: 4). Prominently associated with the work of David Evans (1993), Jeffrey Weeks (1998) and Diane Richardson (2000, 2018), the concept has been deployed in a variety of critical contexts. Our engagement with it here reflects the view that ‘sexual citizenship is useful in that it recognizes and situates individuals and society as intrinsically sexual, thereby contesting more traditional notions of citizenship that have relegated the sexual to the private and the domestic in favour of new possibilities’ (Aggleton et al., 2018: 5). As vibrant spaces for the renegotiation of how people understand ‘public’ and ‘private’ and the links between these ideas and expressions of intimacy and desire, social media are dynamic contexts for observing the renovation of contemporary understandings of sexual citizenship. In particular, these spaces provide opportunities for observing how citizenship is regulated in formal and informal ways by the features of these new contexts, such as their classification practices, which are the focus of this article. This article will examine the platform policies, platform responses to public criticism and responses from users to these modes. In doing so, we interrogate the forms of governance mobilised by algorithmic-enabled filters in order to examine the forms of sexual citizenship they attempt to cultivate – one which we argue identifies responsible LGBTQ subjects as largely devoid of sexual desire.

New social media LGBTQ subjectivity and sexual citizenship

To pursue our investigation regarding the ways in which communication technologies might shape LGBTQ subjectivity and sexual citizenship, we examine controversial changes to classification practices and viewing modes on both YouTube and Tumblr platforms in 2017 and 2018. In the first half of 2017, YouTube and Tumblr made a number of changes to their classification practices and viewing modes, which impacted the accessibility of LGBTQ content across both platforms, which elicited significant criticism from LGBTQ producers and consumers of content (Bell, 2017; Hunt, 2017). These changes in classification troubled expectations of the Internet as a utopic space for queer expression, setting the scene for old debates in sexual citizenship about rights, identity, representation and expression to play out in the new digital landscapes.

By focusing on the changes that YouTube and Tumblr made to their classification practices and viewing modes, this article's contribution is twofold. First, we contribute to a growing literature that conceptualises online classification practices as discrete governance technologies embedded within social media (see, for example, Crawford and Gillespie, 2016; Duguay et al., 2018; Gillespie, 2018; Olszanowski, 2014). Second, we examine how these practices of classification and restriction, functioning in the mode of governance technologies, construct specific ways of being a sexual citizen and how recent changes in classification practices dramatise the productive work of such governance technologies because what they make visible are different understandings of what LGBTQ sexual citizenship is and how it is enacted online. Essentially, we consider the content classification regimes governing YouTube and Tumblr content as norm-producing technologies and examine how normative understandings of LGBTQ sexual citizenship that seek to remove queer desire from the public performance of LGBTQ sexual identity are a key outcome of these technologies.

Conditions of emergence: recent classification practices on YouTube and Tumblr

Restriction on the kinds of content available on social media platforms and practices such as permitting users to 'flag' inappropriate content for removal, or the use of automated systems to detect inappropriate content, have become commonplace. These automated detection mechanisms can involve processes like identifying offensive language, either in the text typed in by users or by converting spoken word in videos to text, or identifying nudity in images (Gillespie, 2018). The problems of flagging have been well established, with scholars identifying the ways this practice, while seeming to shift power to the users of a platform, lacks transparency and works against minority user groups by subjecting them to evaluation by dominant values which may be discriminatory (Crawford and Gillespie, 2016; Duguay et al., 2018; Olszanowski, 2014). These distributed governance processes not only involve users identifying content but also are connected to platform policy documents like the 'terms of service' and the 'community guidelines' that determine what can and cannot be hosted on the sites and how disputes can be managed (Gillespie, 2018). These guidelines, which each platform refines, are

increasingly enforced through algorithmic sorting systems, as well as human content evaluations (Crawford and Gillespie, 2016). While, at times, these guidelines can be violated by social media platforms themselves in their decision-making, both human and automated, they are a reflection of the platforms' current content restriction practices and, more importantly, discursive performances of a platform's values (Gillespie, 2018). Human evaluation is often presented by platforms as recourse to address limitations of algorithmic sorting; however, human moderators often work for extremely low wages in factory-like settings with only seconds to evaluate images, working long shifts of exposure to hours of sometimes highly traumatic material and making nuanced assessments that may relate to cultural contexts removed from their own (Gillespie, 2018; Roberts, 2019). In addition, human content moderators rely on computational tools to cope with the volume of content moderators are expected to assess (Roberts, 2019).

These automated filters and their sorting decisions are controversial, with their errors regularly making news with accusations of bias. However, users often have little information about how their content is filtered, seeing only the final decision of the system presented as objective. However, the exact workings are not shared due to their proprietary nature; the rules they apply are unable to be interrogated by the community subject to them. In particular, popular video-sharing website YouTube and micro-blogging website Tumblr have both faced significant criticism from their LGBTQ users, who found their content was inappropriately being classified as 'adult' simply for being associated with LGBTQ issues (Bell, 2017; Hunt, 2017). While both platforms had been filtering content for many years, changes to their automated content filtering in 2017 and 2018 brought these processes to public attention.

The trouble with YouTube

In February 2017, YouTube's search-engine parent company, Google, faced public outrage and millions of dollars lost in withdrawn advertising business after it was found that videos promoting extremist views and terrorism were being hosted on YouTube (Mostrous, 2017; Solon, 2017). In the wake of the controversy, Google pledged to improve its systems for identifying offensive content to ensure it was not monetized and to allow advertisers greater control over what kinds of content their brands were paired with (Harris, 2017). Soon after this incident, LGBTQ creators reported that their videos were hidden when 'Restricted Mode' was turned on (Hunt, 2017). YouTube had introduced Restricted Mode in 2010 as an opt-in setting for those seeking to restrict mature content such as profanity, sexual content, nudity or violence (Wright, 2017a), although it was only in early March 2017 that LGBTQ content creators and others began to notice their videos being restricted on the platform (Hunt, 2017).

LGBTQ YouTubers reported that videos on their channels were being restricted for the reason that they featured LGBTQ content and stressed that their videos did not include explicit content (Hunt, 2017). YouTuber NeonFiona reported that videos in which she discussed bisexuality were being restricted, while videos that discussed sex explicitly without reference to her bisexuality were not similarly affected (Watson, 2017). The LGBTQ youth organisation Everyone is Gay (2017) then reported on Twitter that all their advice videos had been restricted, and prominent YouTuber and LGBTQ

activist Tyler Oakley (2017) tweeted that a video he had posted called ‘8 Black LGBTQ + Trailblazers Who Inspire Me’ had also been restricted. His tweet was retweeted by thousands.

Following the complaints, the platform responded with a statement apologising for the ‘confusing and upsetting’ mistake (Wright, 2017a). The statement, posted on the YouTube Creator Blog on March 20, 2017, explained,

The bottom line is that this feature isn’t working the way it should. We’re sorry and we’re going to fix it . . . Our system sometimes makes mistakes in understanding context and nuances when it assesses which videos to make available in Restricted Mode. (Wright, 2017a)

A follow-up blog post published in April emphasised the company had ‘fixed an issue that was incorrectly filtering videos for this feature’, positioning the error as a technical problem ‘[o]n the engineering side’ (Wright, 2017b). By June 2017, YouTube announced hundreds of thousands of videos featuring LGBTQ content were now unrestricted (Google, 2017):

[s]haring stories about facing discrimination, opening up about your sexuality, and confronting and overcoming discrimination is what makes YouTube great, and we will work to ensure those stories are included in Restricted Mode. (YouTube, 2017)

A statement provided in the Policy Centre of YouTube’s website provides a working definition of ‘mature content’ in its description of what may still be restricted, which includes drugs and alcohol, sexual situations (but offers an exclusion for ‘some educational’ content), violence, terrorism, war, crime, profane language, and demeaning or incendiary content (YouTube, 2017). YouTube’s Policy Centre also explains that the process by which most videos are classified is primarily automated, with only a small number undergoing human evaluation, with this being determined by metadata, title and language within the video, detected upon uploading of the video (YouTube, 2017).

YouTube denied any connection between the restriction and the recent advertiser boycott, stating that ‘[t]hese were separate issues that unfortunately happened at the same time’ (Google, 2017). However, the increased vigilance around the monetisation of controversial content has undoubtedly adversely impacted LGBTQ creators, with some now reporting that their content is not only restricted but classified as ‘not suitable for all advertisers’, resulting in vastly reduced earnings from their videos (Alkhatib and Bernstein, 2019; Priddy, 2017). Since the initial rollout of Restricted Mode in 2017, LGBTQ content creators have continued to voice their frustrations about inappropriate censorship, with five creators even filing a joint lawsuit against the platform in August 2019 alleging discriminatory practices (Strapagiel, 2019).

Playing it ‘safe’ on Tumblr

In 2017 and 2018, Tumblr encountered similar criticism from LGBTQ users after introducing significant changes to the way ‘adult’ content was filtered on their platform. On 20 June 2017, the social media platform launched its ‘Safe Mode’ in the form of a filter

for ‘sensitive content’ which operated within both a user’s search results and dashboard (Tumblr Staff, 2017b). While users had been encouraged to flag their own content as sensitive on the site, the platform now undertook a more automated screening of content. Although the exact mechanics of this screening were not made transparent, this does include automated image analysis to detect nudity (Tumblr Staff, 2017a).

Since introducing Safe Mode, Tumblr has gone on to introduce a total ban on adult content in December 2018 – a move that has seen the site lose 30% of its web traffic (Liao, 2019). This is the latest move in Tumblr’s long war on adult content on the site, with Safe Mode being an upgrade of existing restrictions on the visibility of adult content on the platform, which had previously pertained only to search functions (Baker-Whitelaw, 2013). These changes began in 2013 when the platform was purchased by search-engine company Yahoo, which sought to clean up Tumblr, which hitherto had hosted a significant amount of pornographic material on its site,³ to make it more attractive to advertisers (Perez, 2013).

Each incremental increase of content filtering has been accompanied by significant outcry about inappropriate censorship. Soon after the introduction of search restrictions in 2013, users began reporting that searches in Safe Mode for terms like ‘gay’ or ‘lesbian’ were returning no results (Baker-Whitelaw, 2013). Tumblr emphasised this was to prevent adult content being presented in these search results but stated that the company was working towards ‘more intelligent filtering’ (Tumblr Staff, 2013). Extensions of Safe Mode during 2017 may also have been motivated by a revenge pornography controversy that emerged months before the platform announced the new Safe Mode, in which the site was criticised by victims for being slow to take down non-consensual nude images (Marsh, 2017).

Three days after announcing the extension of Safe Mode in June 2017, Tumblr received a significant volume of complaints from users that LGBTQ content was being inappropriately censored (Tumblr Staff, 2017a). Users reported that the filtering system was identifying non-explicit LGBTQ content as ‘sensitive’, which meant it was not visible in Safe Mode, as well as failing to identify pornography as such (Castello, 2017). The platform posted an apology, stating that

We’ve heard from a bunch of you that Safe Mode was filtering posts from the LGBTQ+ community even though they were completely innocuous and totally safe-for-work. Please know that was never our intention, and we appreciate you letting us know so quickly – and forcefully! We’re deeply sorry. Tumblr will always be a place where everyone is welcome and protected, so we want to explain what happened. (Tumblr Staff, 2017a)

The platform identified a number of reasons for the error, primarily that users who had self-flagged their blogs as explicit were incorrectly having all their posts flagged as sensitive regardless of content and that any post reblogged from an originally explicit post was being automatically classified as sensitive. Consequently, all their content was not visible to users who had activated Safe Mode regardless of the nature of the post. Tumblr also blamed the nudity detecting algorithm, explaining that ‘[w]hen you make a photo post, a computer algorithm classifies the image as safe or sensitive. It’s a machine so it’s not perfect’ (Tumblr Staff, 2017a). In doing so, like YouTube, Tumblr sought to distance

itself from accusations of bias and affirm that the error had been corrected by adjustments to the mechanics of the screening process.

In February 2018, Tumblr further extended Safe Mode by making it the default setting for all users, meaning it had to be turned off by every user who was not already using it (Cole, 2018). Then in December 2018, Tumblr announced that it would ban all adult content (Tumblr Staff, 2018). The ban follows recent incidents involving child pornography being found on the site, which resulted in Apple banning Tumblr from the App Store (Porter, 2018). Tumblr updated their user guidelines to include the prohibition of ‘real-life human genitals or female-presenting nipples’ as well as ‘any content, including images, videos, GIFs, or illustrations, that depicts sex acts’ (Tumblr Staff, 2018). Users expressed outcry at the decision, with some taking to rival platform Twitter to vent their frustrations using the hashtags #TumblrIsDead and #BoycottTumblr (Gremore, 2018). LGBTQ users and creators whose content explored sexuality argued their community would be particularly marginalised by the move, identifying Tumblr as a place that had previously offered a ‘safe space’ for them (Braidwood, 2018; Lee, 2018; Liao, 2018).

Content classification and the platform

Social media platforms have placed themselves strategically in content regulation debates, with platforms being reluctant to engage in the practice of content moderation, fearing backlash on the basis of infringing on ‘free speech’ but forced to do so in the face of social, financial and legal obligations (Gillespie, 2018). Gillespie (2010) argues that the term ‘platform’ itself offers online content providers a way of strategically positioning themselves within policy to reduce their liability, both legally and culturally, for the content on the sites. This desire is an understandable one given the difficult task that is moderation, making decisions on highly political and sensitive issues with extremely limited time and information and on topics which are messy to navigate in the best of circumstances.

Classification systems operate to identify objects as belonging to distinct categories and consequently play a significant role in knowledge production. As Bowker and Star (2000) argue, classification is an often-invisible ordering of human interaction based on dominant moral values. These moral values frequently reinforce existing inequalities and power relations and as such classification systems tend to neglect marginalised groups (Blackwell et al., 2017). There is a long-standing relationship between classification practices and the understandings of youth sexual citizenship which are shaped by such practices. As Grealy and Driscoll (2015) have argued, classification practices have long played a significant role in ‘managing relations between “youth” and “culture” as a pedagogy of citizenship’ (p. 63). Grealy and Driscoll (2015) point out that it is important, however, to draw a distinction between classification and censorship as governance technologies, arguing that ‘classification-as-governance depends on concepts relevant to the cultural training of youth as much as their social protection, and as such functions as a cultural pedagogy’ (p. 64). Framed as ‘cultural pedagogy’, the shifting practices of classification on YouTube and Tumblr which we have outlined can be seen as disciplinary mechanisms which incorporate the subject through experiences of restriction and access to diverse materials and produce an understanding of the subject which relies on these experiences.

The algorithmic sorting of content classification has, in particular, been subject to criticism for reinforcing discrimination while obscuring the workings of such discrimination (Noble, 2018). Noble (2018), examining the racism and sexism of algorithms, argues that 'algorithmic oppression is not just a glitch in the system but, rather, is fundamental to the operating system of the web' (p. 10). Software code, in collaboration with users, platform policies and broader social norms, plays a role in constituting *normative* gendered sexual citizenship. Indeed, as Bivens (2017: 881) outlines in relation to Facebook, 'software can produce the conditions for gendered existence', noting that although the platform has moved towards much more open categories of gender identification, binary gender persists in the code of the platform and in the deeply embedded software structures. Furthermore, when it comes to regulatory mechanisms like algorithms, as Bowker and Star (2000), emphasise these 'algorithms for codification do not resolve the moral questions involved, although they may obscure them' (p. 24).

Normative sexual citizenship

Classification governs interactions with texts through the journey to 'adulthood', and as such the assessment of LGBTQ content in general as 'adults only' material or not suitable for a 'general' audience both normalises heterosexuality and reinforces negative historical associations of LGBTQ life with the illicit. LGBTQ young people are thus placed in an inevitably antagonistic relationship to such classifications and restrictions, which block their access to expression, representation and even, given the implications of monetization, the material resources of LGBTQ life. The events surrounding YouTube's *Restricted Mode* and Tumblr's *Safe Mode* highlight the effects of codifying practices in which the complexities of sexuality and sex must be reduced for efficiency – explorations of non-normative sexuality and LGBTQ expressions of desire automatically get placed within the restricted category. The protection of children from content deemed 'inappropriate' remains a central aim of contemporary classification practices (Flew, 2012; Grealy and Driscoll, 2015; Leitch and Warren, 2015). In these ways, the recent moves by platforms like YouTube and Tumblr that configure LGBTQ sexual citizenship as inappropriate for children is in line with historical discourses of LGBTQ people as dangerous outsiders or, conversely, as people who need to exercise stealth to circulate in contemporary economies of visibility and knowledge. In addition, insofar as these classification practices are *technologies* of citizenship, they are technologies of *sexual* citizenship because classification is preoccupied with sex. Furthermore, content classification, as a norm-producing technology, works against the diversity of uses and users that contemporary platforms like YouTube and Tumblr make more and more explicit. Sex emerges as a difficult case, but one which points out classification's reach as well as its limits.

The troubled relationship between sex and classification is aptly depicted in the difficulties encountered in classifying online material that is the focus of these platform scandals. For example, in many of the issues raised by critics of recent changes in online classification, it is clear that what passes as 'sexual' content has, in practice, not been clearly defined. As users have pointed out, non-pornographic representations of 'sexuality' are often classified in the same way as explicit depictions of sex, demonstrating the

failure by the classifying ‘system’ (algorithm, user rating, screening staff or a combination of these) to make clear distinctions between signs of sexuality and sex, or pornographic and non-pornographic depictions. This brings attention to the messiness of these distinctions, exacerbated in the context of automated classification applied to the significant volumes of content circulating on social media sites. These classification systems struggle to deal with the ever-growing and diversifying content they manage because algorithmic modes only collect and register content based on what can be recognised in a manner suitable for automated evaluation. Policies of what is acceptable, significantly influenced by the processing of historical content, dictate classification protocols. However, what cannot be captured by these filters, what is not reducible, cannot be registered as ‘sexual’ (or indeed ‘violent’, for example). This means that classification is biased towards historical depictions and to depictions which can be expressed linguistically. Offensive content which eschews straightforward depiction based on pre-existing categories can evade the classification process.

Beyond content classification, social media platforms more broadly have long been implicated in conflict with the LGBTQ community and other marginalised groups over biases in their platform structures and policies. For example, Lingel and Golub’s (2015) study of Brooklyn’s drag community and its battle with Facebook’s ‘real name policy’ highlights the platform’s design incompatibility with diverse approaches to gender performances and naming practices. Indeed, DeNardis and Hackl (2016) argue that debates about LGBTQ rights often arise around Internet governance systems, noting that these platforms are not neutral and can reproduce or even exacerbate the marginalisation of LGBTQ people through the ways they act as ‘control points’ through which information is accessed, and which voices are heard.

The ‘good’ queer sexual citizen without desire

Although the algorithmic sorting systems of the platforms reinforce long-standing associations between LGBTQ communities and illicit practices, there are more subtle but no less normative constitutions of sexual citizenship at play in the policies of platforms and in their responses to public criticism. While the classification regimes struggle to establish the nuanced distinction between sex and sexuality, and between what is explicit and what is acceptable for wider distribution, dialogue between the platforms and their community in the wake of these censorship scandals further attempts to stabilise the boundaries between these categories, presenting them as discoverable only if the correct systems are created. These restrictive modes create the conditions in which gender and sexual difference are readily distinguishable from desire, and in which ‘good’ LGBTQ sexual citizenship practices are oriented around preserving this distinction.

The responses from social media platforms to criticism of their content moderating systems seek to preserve in their user communities a hope that systems of classification will be corrected such that the ‘right’ forms of queer life will be permitted through the filter. The platforms need their users to buy into and aspire to their visions of acceptable LGBTQ expression, rather than challenge these heteronormative distinctions. YouTube’s (Wright, 2017a) public statement about the errors emerging in Restricted Mode, in which it was stated that there was ‘nothing more important to us than being a platform where

anyone can belong' and Tumblr Staff's (2017a) assurance that '[i]t might take some time to get it perfect, but we're committed to getting there with your help' reflect a desire by both platforms to contain the issue and incorporate gender and sexual difference, though within their own limited and normative terms.

Statements from the platforms construct the algorithmic classification process as sophisticated, yet inept in the face of the 'problem' of sexuality. To date, platforms have not revealed the workings of their filtering systems beyond the most simplistic of descriptions. As Crawford (2015) explains, the negotiations that take place are 'nonnegotiable and kept far from view, inside an algorithmic "black box"' (p. 77). While there is some reverence for the sophistication of the algorithm in the language adopted by the platforms, both platforms undermine the intelligence of their systems in order to justify the 'errors' made. Tumblr Staff (2017a) describes the mistakes made by its system as 'silly', while YouTube (Wright, 2017a) stated that their system makes 'mistakes in understanding context and nuances'. Although the algorithm is mobilised here by social media platforms as an objective sorting mechanism, this idea has been widely problematised (Bozdag, 2013; Gillespie, 2017; Kitchin, 2017; Ziewitz, 2016). In contrast to their positioning of the algorithm as value-free, the platforms both significantly downplay the role of human judgements in the analysis and restriction of content. However, this does not reflect this significant degree to which human intervention is a part of these processes (Bozdag, 2013). Bozdag (2013) argues this is a common tactic of online web services to distance themselves from the judgements they make.

While preserving the image of the foolish and clumsy machine, the platforms also retain hope and desire for a future in which the sorting mechanisms will 'get it right' where only appropriate content will be censored. Both platforms cultivated this hope by emphasising the steps they had already taken to fixing their system shortly after users responded with outrage (Tumblr Staff, 2017a; Wright, 2017b). The platforms also work to carefully preserve distinctions between 'good' LGBTQ content, described as 'completely innocuous' (Tumblr Staff, 2017a), and 'bad' LGBTQ content and affirm the ease by which these two types of content are distinguishable, despite the errors the filter may have made. It is significant that in their apology statement, YouTube (Wright, 2017a) specifically identifies a number of videos by LGBTQ creators where the platform's filter 'got it wrong', using the example of a queer Youtuber's video about their wedding vows and another in which a young gay man comes out to his grandmother. These examples serve to demonstrate the normative queer subject, in the form of the one who participates in and seeks validation from heteronormative social institutions like marriage and the family. Such subjects are presented in opposition to the more 'mature topics' to which queer desire belongs. It is significant too, that the initial impetus for both YouTube and Tumblr to introduce and extend filtering systems was connected to pressure from advertisers to avoid advertisements appearing alongside content not deemed to be 'suitable for all audiences' (Perez, 2013; Solon, 2017). These pressures and practices can be compared to the prevalence of 'multicasting' in television production, in which content seeks to appeal to the widest possible audience, rather than fill a niche (Himberg, 2014). Platform responses to LGBTQ users' complaints that seek to identify what is acceptable and unacceptable queerness in the mainstream can thereby be situated within a broader history of media representations of queer life that are oriented towards heterosexual tastes – via a process that Ng (2013) has termed 'gaystreaming'.

Ultimately, these classification regimes produce a responsible sexual citizen through processes that seek to keep sex and queer desire contained and reserved for the 'adult'. This attempt to obscure queer desire from public digital space can be situated within a history of heteronormativity in which heterosexual sex is rendered ordinary (Warner, 2000). As Rubin (1984) argues, queerness is acceptable only to the extent that it resembles 'good' and 'natural' heterosexual sex. Yet, as Lauren Berlant (1997) reminds us, the protection of privacy is for heterosexuality, while 'all queers have is that closet' (p. 63). Putting this in the context of the normative sexual citizenship these social media platform policies constitute, while forms of LGBTQ life may be eventually deemed acceptable for general viewing by these restricted modes, the queer desires that underpins this life remains suspect. Responsible sexual citizenship here, for the LGBTQ community, relies heavily on homonormativity (Duggan, 2002), in which heterosexual privilege must be maintained and queer desire must remain 'private' lest it threaten 'acceptable' LGBTQ content.

Scandalising Weibo

The production of such limited options for LGBTQ expression can be seen in a censorship scandal that occurred in April 2018 in relation to the Chinese social media site Weibo. The site faced intense public outcry after it issued new community guidelines that sought to remove homosexual content from its platform, along with pornographic and violent material (Koetse, 2018). Weibo's change in policy served to bring the site into line with regulations governing Internet content issued by the mainland Chinese government (Koetse, 2018). However, only 2 days after making the announcement, Weibo reversed the ban on homosexual content, following a grassroots viral campaign by its users organised around the hashtags #iamgay and #iamgaynotapervert (Kuo, 2018). Although Weibo's controversy needs to be understood within its specific cultural and political context (for further information, see Koetse, 2017), like Tumblr and YouTube these incidents involving the restriction of content made similar connections between LGBTQ content and violence or pornography. And, as with the LGBTQ campaign against censorship on Tumblr and YouTube, the changes proposed by Weibo were countered by protests from within the LGBTQ community which mobilised so as to constitute queerness as 'safe'. This is reflected in the hashtags used to proclaim the distinction between queerness and perversion, constructing social networking spaces as accessible to the LGBTQ community *conditionally* if they demonstrate normative sexual behaviour. Weibo's censorship scandal and the resulting resistance campaign demonstrates the ways such classification technologies help induct LGBTQ subjects into normative sexual citizenship, even while they may seek to resist the censorship itself.

The horse is naked: potential for resistance beyond normative sexual citizenship

Despite the normative imperative driving the work of classification, we argue that the high-profile failures of these classification systems create the conditions for users to dwell on the messy boundaries between sex and sexuality. By drawing attention to the

failure of rigid categories of classification, the current debate can be seen as an opportunity for fashioning more nuanced and different approaches to the assessment of online content, based on context, contingency and mobility of meaning. Through this distribution of meaning and opportunity to more and more online users, these shifts in classification invite reflections on how diversified practices of access and restriction can give rise to different experiences and expressions of sexual citizenship.

In the wake of the 2017 and 2018 Safe Mode scandals on Tumblr, users posted about the system's failure to appropriately classify posts as well as shared memes mocking the platform. One user posted a list of 'Things Tumblr Safemode has blocked from my infant eyes' (sic), which was liked and reblogged by thousands of users (Nyan, 2017). The list included a 'gif of a parrot getting brushed (sic)' and a '[v]ideo of Fireworks'. A widely reshared post by another user showed an image of a tan coloured horse being ridden by a dog; the user offers a look into the source code of the image to reveal that it was classified as 'unsafe' (dongulusdisgustus, 2017).⁴

In 2018, following the ban of adult content on Tumblr, users played with the language of the new guidelines referring to 'female presenting nipples', posting images that were flagged as inappropriate and speculating that they may have contained these nipples (Bright, 2018). One user posted a topless cartoon image of Nintendo videogame character Mario with the caption 'at least we still have Mario-presenting nipples', only to show a screenshot of the image flagged as explicit (osha-watt, 2018). These and other similar posts invite playful reflection on the part of the viewer as to what the algorithm might see in the image. Put simply, was the horse naked? Could the algorithm see Mario's nipples? The tan of the horse's coat is refigured for the viewer through the perspective of the Safe Mode algorithm, to consider what nudity might mean for the filter. Though some of these posts were circulated by users to make light of the classification mechanism, at the same time the viewer may, albeit playfully, imagine how the image may be interpreted digitally as nakedness or sexuality. In this sense, we might interpret these algorithmic sorting systems as also capable of producing queer encounters with platform users, which undo attempts to make these categories of sex, sexuality, explicit and 'safe' solid.

In the statements made by YouTube and Tumblr, platform policies attempt to establish a false opposition between algorithmic failure to attend to nuance and the sophistication of human judgement. In contrast, the reality of both these processes involves the reduction of dynamic sexual practices, desires and pleasures to blunt representational frames. In this article, we have explored the capacity of these restrictive modes to produce normative sexual citizenship through stabilising boundaries between sex and sexuality and preserving adulthood and maturity around these boundaries. However, within the constant negotiation ongoing in navigating what constitutes sex, sexuality or nudity, there remains the potential for resistance in the excessive experience of content that cannot be captured by these representational systems. 'Sexuality refuses demystification', as Edelman (2004: 28) proclaims, and classification practices always fail to understand it fully, try as they might to grasp its principal characteristics.

By sharing moments of the algorithm's failure, users also defy the representational terms of the policies and incline towards the excessive and messy nature of classification practices. In these parodic moments, we find the fraying of the authority and credibility of online classification. As such, they create opportunities for cultivating an ethics of

access and restriction in relation to sexual and gender difference which is of a different order to the corporate, systems-level work of classification currently in place. The users encounter with the content classification algorithm here might be best understood through Bucher's (2017) concept of the *algorithmic imaginary*, which she defines as 'the way in which people imagine, perceive and experience algorithms and what these imaginations make possible' (p. 31). Bucher argues for an understanding of the affective dimensions of entanglements with algorithmic logics and seeks to attend to the *felt* dimension of algorithms. In taking such an approach to the playful responses to Tumblr's Safe Mode, we can observe users reflecting on images through the imagined eye of the algorithm, becoming sensitised to parts of the image that might have drawn its attention. In their encounter with this content classification algorithm, users may reflect on how one might differentiate between a cartoon nipple that is erotic versus one that is 'non-sexual'. Although these responses by Tumblr users identify the faults with the algorithm they also highlight the ambiguities and always-more-than-representational nature of the subject matter.

Conclusion

The restricted modes that we have discussed here draw attention to the normative production of sexual citizenship, of adulthood and maturity through the connection of sex and violence to 'adult' content, but also more subtly to the messiness of distinctions between sex and sexuality. By seeking to affirm the possibility of these mechanisms being corrected to adequately distinguish between sex and sexuality, the responses from YouTube and Tumblr construct a queer subject without desire as the 'good' sexual citizen in their online communities. Their content moderation policies attempt to cultivate a belief in the possibility of a future of LGBTQ inclusion that is predicated on separating sexuality from queer sexual identity, such that queer can be acceptable 'content' only if queer bodies are not sexualised.

Sexual citizenship is, we argue, negotiated through many digital practices, including these content classification processes. These flawed modes disrupt the safe spaces on these platforms in which queerness can flourish, revealing them to be subject to some of the same kinds of censorship of queerness that occur in other spaces. However, they also reveal the futility of the *fantasy* of these algorithmic censors – namely, that we can easily 'get it right' or distinguish between these messy categories. In cultivating this awareness, these moments of censorship create a space in which users may dwell on what constitutes nudity, sexual content or desire in ways that are sometimes playful and oriented towards resisting the normative modes of sexual citizenship that imagine these distinctions to be solid and self-evident. By attending to these emerging differences, we can conceptualise new ways of doing queer ethics online, reflecting on the flaws within contemporary classification practices and embracing desire through these mediated constraints.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a Discovery Grant from the Australian Research Council.

ORCID iD

Clare Southerton  <https://orcid.org/0000-0002-5812-9785>

Notes

1. 'Queer' and LGBTQ are used interchangeably in this article to refer to people who are of diverse genders and sexualities, to reflect current popular terminology in which queer operates as an umbrella term for a proliferating group of more specific sexual and gender identities (Robards et al., 2018).
2. For a discussion of the historical circulation of 'sexual citizenship' as an idea, and in particular its relationship to ideas about 'youth', please see the introduction to *Youth Sexuality and Sexual Citizenship* (Aggleton et al., 2018).
3. At the time of Yahoo's purchase of the site, adult content represented 11.4% of Tumblr's top 200 most-visited domains and 22% of referral traffic to the site was from adult websites (Perez, 2013).
4. In line with the Australian National Health and Medical Research Council's (2018) guidelines for research involving social media content, we considered the context in which these social media posts (on Tumblr) were located before choosing not to anonymise the content. Because the content was presented with the intention of shareability and had reached a wide audience, we wanted the authors to retain authorship of their content.

References

- Aggleton P, Prankumar SK, Cover R, et al. (2018) Introduction. In: Aggleton P, Cover R, Leahy D, et al. (eds) *Youth, Sexuality and Sexual Citizenship*. London & New York: Routledge, pp. 1–16.
- Alkhatib A and Bernstein M (2019) Street-level algorithms: a theory at the gaps between policy and decisions. In: *CHI conference on human factors in computing systems proceedings (CHI 2019)*, Glasgow, 4–9 May. Available at: <https://doi.org/10.1145/3290605.3300760>
- Baker-Whitelaw G (2013) New NSFW content restrictions enrage Tumblr users. *The Daily Dot*, 18 July. Available at: <https://www.dailydot.com/irl/tumblr-nsfw-content-tags-search/> (accessed 14 November 2017).
- Bell K (2017) Why Tumblr's new 'safe mode' is a bigger deal than you think. *Mashable*, 22 June. Available at: <http://mashable.com/2017/06/21/tumblr-safe-mode/> (accessed 14 November 2017).
- Berlant L (1997) *The Queen of America Goes to Washington City: Essays on Sex and Citizenship*. Durham, NC; London: Duke University Press.
- Bivens R (2017) The gender binary will not be deprogrammed: ten years of coding gender on Facebook. *New Media & Society* 19(6): 880–898.
- Blackwell L, Dimond J, Schoenebeck S, et al. (2017) Classification and its consequences for online harassment: design insights from HeartMob. *Proceedings of the ACM on Human-Computer Interaction* 1: 24:1–24:19.
- Bowker GC and Star SL (2000) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.

- Bozdag E (2013) Bias in algorithmic filtering and personalization. *Ethics and Information Technology* 15(3): 209–227.
- Braidwood E (2018) Tumblr's porn ban is an 'erasure' of the LGBT community. *PinkNews*, 5 December. Available at: <https://www.pinknews.co.uk/2018/12/05/tumblr-porn-ban-erasure-lgbt-community/> (accessed 12 December 2018).
- Bright P (2018) Tumblr's porn ban is going about as badly as expected. *Ars Technica*. Available at: <https://arstechnica.com/gaming/2018/12/tumblrs-porn-ban-is-going-about-as-badly-as-expected/> (accessed 12 December 2018).
- Bucher T (2017) The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication and Society* 20(1): 30–44.
- Byron P and Robards B (2017) There's something queer about Tumblr. *The Conversation*. Available at: <http://theconversation.com/theres-something-queer-about-tumblr-73520> (accessed 5 March 2018).
- Castello J (2017) Why are Tumblr, Twitter, And YouTube blocking LGBTQ+ content? *The Establishment*, 20 July. Available at: <https://theestablishment.co/why-are-tumblr-twitter-and-youtube-blocking-lgbtq-content-e54e7acf4b5c> (accessed 17 November 2017).
- Cho A (2015) Queer reverb: Tumblr, affect, time. In: Hillis K, Paasonen S and Petit M (eds) *Networked Affect*. Cambridge, MA: MIT Press, pp. 43–58.
- Cho A (2017) Default publicness: queer youth of color, social media, and being outed by the machine. *New Media & Society* 20(9): 3183–3200.
- Ciechalski S (2017) YouTube debuts inspiring Pride Month video to highlight queer creators. *Mashable*, 28 June. Available at: <https://mashable.com/2017/06/28/youtube-pride-campaign-proudtobe/> (accessed 5 March 2018).
- Cole S (2018) Tumblr is turning 'safe mode' on by default for everyone. *Motherboard*, 9 February. Available at: https://motherboard.vice.com/en_us/article/3k7qe5/turning-tumblr-safe-mode-on-by-default (accessed 21 February 2018).
- Crawford K (2015) Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology & Human Values* 41(1): 77–92.
- Crawford K and Gillespie T (2016) What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media & Society* 18(3): 410–428.
- DeNardis L and Hackl AM (2016) Internet control points as LGBT rights mediation. *Information, Communication and Society* 19(6): 753–770.
- dongulusdisgustus (2017) *Sensitive Tumblr Help*. Tumblr. Available at: <http://dongulusdisgustus-archive.tumblr.com/post/162131621628/sensitive-tumblr-help> (accessed 4 May 2018).
- Duggan L (2002) The new homonormativity: the sexual politics of neoliberalism. In: Castronovo R and Nelson DD (eds) *Materializing Democracy: Toward a Revitalized Cultural Politics*. Durham, NC: Duke University Press, pp. 175–194.
- Duguay S (2014) 'He has a way gayer Facebook than I do': investigating sexual identity disclosure and context collapse on a social networking site. *New Media & Society* 18(6): 891–907.
- Duguay S, Burgess J and Suzor N (2018) Queer women's experiences of patchwork platform governance on Tinder, Instagram, and Vine. *Convergence*. Available at: <https://doi.org/10.1177/1354856518781530>
- Edelman L (2004) *No Future: Queer Theory and the Death Drive*. Durham, NC; London: Duke University Press.
- Evans D (1993) *Sexual Citizenship: The Material Construction of Sexualities*. London; New York: Routledge.
- Everyone is Gay (2017) *RT if you are as outraged as we are that now, in restricted mode, young people cannot access our library of LGBTQ advice videos on @YouTube*. Available at: <https://twitter.com/everyoneisgay/status/843631133245956096> (accessed 19 April 2018).

- Flew T (2012) Media classification: content regulation in an age of convergent media. *Media International Australia* 143(1): 5–15.
- Gillespie T (2010) The politics of ‘platforms’. *New Media & Society* 12(3): 347–364.
- Gillespie T (2017) Algorithmically recognizable: Santorum’s Google problem, and Google’s Santorum problem. *Information, Communication and Society* 20(1): 63–80.
- Gillespie T (2018) *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.
- Google (2017) *A Message on Pride and LGBTQ Initiatives*. Available at: <https://youtube-creators.googleblog.com/2017/06/a-message-on-pride-and-lgbtq-initiatives.html> (accessed 16 November 2017).
- Grealy L and Driscoll C (2015) The plastic adolescent: classification and minority. In: Driscoll C, Noble G and Watkins M (eds) *Cultural Pedagogies and Human Conduct*. Abingdon: Routledge, pp. 81–95.
- Gremore G (2018) People are taking Tumblr’s ban on x-rated content really, really personally. *Queerty*, 5 December. Available at: <https://www.queerty.com/people-taking-tumblrs-ban-x-rated-content-really-really-personally-20181205> (accessed 12 December 2018).
- Harris R (2017) Improving our brand safety controls. *Google*, 17 March. Available at: <https://blog.google/topics/google-europe/improving-our-brand-safety-controls/> (accessed 16 November 2017).
- Himberg J (2014) Multicasting: lesbian programming and the changing landscape of cable TV. *Television & New Media* 15(4): 289–304.
- Hunt E (2017) LGBT community anger over YouTube restrictions which make their videos invisible. *The Guardian*, 20 March. Available at: <http://www.theguardian.com/technology/2017/mar/20/lgbt-community-anger-over-youtube-restrictions-which-make-their-videos-invisible> (accessed 14 November 2017).
- Kitchin R (2017) Thinking critically about and researching algorithms. *Information, Communication and Society* 20(1): 14–29.
- Koetse M (2017) New rules for online videos in China: ‘no displays of homosexuality’. *What’s On Weibo*, 30 June. Available at: <https://www.whatsonweibo.com/new-rules-online-videos-china-no-displays-homosexuality/> (accessed 19 April 2018).
- Koetse M (2018) Weibo’s new online guidelines: no homosexual content allowed. *What’s On Weibo*, 14 April. Available at: <https://www.whatsonweibo.com/weibos-new-online-guidelines-no-homosexuality-allowed/> (accessed 18 April 2018).
- Kuo L (2018) China’s Weibo reverses ban on ‘homosexual’ content after outcry. *The Guardian*, 16 April. Available at: <http://www.theguardian.com/world/2018/apr/16/china-weibo-bans-homosexual-content-protest> (accessed 18 April 2018).
- Lee D (2018) Tumblr’s porn ban abandons the marginalised. *BBC News*, 4 December. Available at: <https://www.bbc.com/news/technology-46435975> (accessed 12 December 2018).
- Leitch S and Warren M (2015) Applying classification controls to Internet content in Australia. *Journal of Information, Communication & Ethics in Society* 13(2): 82–97.
- Liao S (2018) Tumblr’s adult content ban means the death of unique blogs that explore sexuality. *The Verge*, 6 December. Available at: <https://www.theverge.com/2018/12/6/18124260/tumblr-porn-ban-sexuality-blogs-unique> (accessed 12 December 2018).
- Liao S (2019) After the porn ban, Tumblr users have ditched the platform as promised. *The Verge*. Available at: <https://www.theverge.com/2019/3/14/18266013/tumblr-porn-ban-lost-users-down-traffic> (accessed 30 October 2019).
- Lingel J and 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.

- Marsh J (2017) 'I haven't stopped crying since': Tumblr revenge porn victim. *New York Post*, 20 March. Available at: <https://nypost.com/2017/03/20/tumblr-under-fire-for-ignoring-revenge-porn-posts/> (accessed 17 November 2017).
- Mostrous A (2017) Google faces questions over videos on YouTube. *The Times*, 9 February. Available at: <https://www.thetimes.co.uk/article/google-faces-questions-over-videos-on-youtube-3km257v8d> (accessed 16 November 2017).
- National Health and Medical Research Council, Australian Research Council and Universities Australia (2018) *National Statement on Ethical Conduct in Human Research 2007 (Updated 2018)*. National Health and Medical Research Council. Available at: <https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018>
- Ng E (2013) A 'post-gay' era? Media gaystreaming, homonormativity, and the politics of LGBT integration. *Communication, Culture and Critique* 6(2): 258–283.
- Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- Nyan (2017) *Things Tumblr Safemode has blocked from my infant eyes*. Tumblr. Available at: <http://hit-a-rock.tumblr.com/post/162104163557/things-tumblr-safemode-has-blocked-from-my-infant> (accessed 4 May 2018).
- Oakley T (2017) *still not fixed. one of my recent videos '8 Black LGBTQ+ Trailblazers Who Inspire Me' is blocked because of this. i'm perplexed, @YouTube*. Available at: <https://twitter.com/tyleroakley/status/843544801916010496> (accessed 19 April 2018).
- Olszanowski M (2014) Feminist self-imaging and Instagram: tactics of circumventing sensorship. *Visual Communication Quarterly* 21(2): 83–95.
- osha-watt (2018) OH COME ON. Tumblr. Available at: <https://osha-watt.tumblr.com/post/180764030960/oh-come-on> (accessed 12 December 2018).
- Perez S (2013) Tumblr's adult fare accounts for 11.4% of site's top 200K domains, adult sites are leading category of referrals. *TechCrunch*, 20 May. Available at: <https://techcrunch.com/2013/05/20/tumblrs-adult-fare-accounts-for-11-4-of-sites-top-200k-domains-tumblrs-adult-fare-accounts-for-11-4-of-sites-top-200k-domains-adults-sites-are-leading-category-of-referrals/> (accessed 17 November 2017).
- Perez S (2017) Tumblr says it fixed the 'Safe Mode' glitch that hid innocent posts, including LGBTQ+ content. *TechCrunch*, 24 June. Available at: <https://techcrunch.com/2017/06/24/tumblr-says-it-fixed-the-safe-mode-glitch-that-hid-innocent-posts-including-lgbtq-content/> (accessed 17 November 2017).
- Porter J (2018) Tumblr was removed from Apple's App Store over child pornography issues. *The Verge*, 20 November. Available at: <https://www.theverge.com/2018/11/20/18104366/tumblr-ios-app-child-pornography-removed-from-app-store> (accessed 10 December 2018).
- Priddy M (2017) Why is YouTube demonetizing LGBTQ videos? *Autostraddle*, 22 September. Available at: <https://www.autostraddle.com/why-is-youtube-demonetizing-lgbtqia-videos-395058/> (accessed 17 November 2017).
- Richardson D (2000) Constructing sexual citizenship: theorizing sexual rights. *Critical Social Policy* 20(1): 105–135.
- Richardson D (2018) *Sexuality and Citizenship*. Cambridge: Polity Press.
- Robards B, Churchill B, Vivienne S, et al. (2018) Twenty years of 'cyberqueer': The enduring significance of the Internet for young LGBTIQ+ people. In: Aggleton P, Cover R, Leahy D, et al. (eds) *Youth, Sexuality and Sexual Citizenship*. London & New York: Routledge, pp. 151–167.
- Roberts ST (2019) *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press.

- Rubin G (1984) Thinking sex: notes for a radical theory of the politics of sexuality. In: Vance CS (ed.) *Pleasure and Danger: Exploring Female Sexuality*. Boston, MA: Routledge & Kegan Paul, pp. 265–319.
- Solon O (2017) Google's bad week: YouTube loses millions as advertising row reaches US. *The Guardian*, 25 March. Available at: <http://www.theguardian.com/technology/2017/mar/25/google-youtube-advertising-extremist-content-att-verizon> (accessed 31 October 2017).
- Strapagiel L (2019) LGBTQ creators are suing YouTube for discrimination. *BuzzFeed News*, 14 August. Available at: <https://www.buzzfeednews.com/article/laurenstrapagiel/lgbtq-creators-youtube-lawsuit> (accessed 28 August 2019).
- Tumblr Staff (2013) *All, we've heard from a bunch of you who are concerned about Tumblr censoring NSFW/adult content*. Available at: <https://staff.tumblr.com/post/55906556378/all-weve-heard-from-a-bunch-of-you-who-are> (accessed 14 November 2017).
- Tumblr Staff (2017a) *Hi Tumblr – safe mode update*. Available at: <https://staff.tumblr.com/post/162178688374/safe-mode-update> (accessed 28 September 2017).
- Tumblr Staff (2017b) *Safe mode is here, Tumblr*. Available at: <https://staff.tumblr.com/post/162047130530/safe-mode> (accessed 28 September 2017).
- Tumblr Staff (2018) *A better, more positive Tumblr*, 3 December. Available at: <https://staff.tumblr.com/post/180758987165/a-better-more-positive-tumblr> (accessed 10 December 2018).
- Wargo JM (2015) 'Every selfie tells a story . . .': LGBTQ youth livestreams and new media narratives as connective identity texts. *New Media & Society* 19(4): 560–578.
- Warner M (2000) *The Trouble with Normal: Sex, Politics, and the Ethics of Queer Life*. Harvard, MA: Harvard University Press.
- Watson L (2017) YouTube's restricted mode is hiding some LGBT content. *Gizmodo Australia*, 18 March. Available at: <https://www.gizmodo.com.au/2017/03/youtubes-restricted-mode-is-hiding-some-lgbt-content/> (accessed 17 November 2017).
- Weeks J (1998) The sexual citizen. *Theory, Culture & Society* 15(3): 35–52.
- Wright J (2017a) *Restricted mode: how it works and what we can do better*. Available at: <https://youtube-creators.googleblog.com/2017/03/restricted-mode-how-it-works-and-what.html> (accessed 28 September 2017).
- Wright J (2017b) *An update on restricted mode*. Available at: <https://youtube-creators.googleblog.com/2017/04/an-update-on-restricted-mode.html> (accessed 28 September 2017).
- Wuest B (2014) Stories like mine: coming out videos and queer identities on YouTube. In: Pullen C (ed.) *Queer Youth and Media Cultures*. Basingstoke: Palgrave Macmillan, pp. 19–33.
- YouTube (2017) *Your content & restricted mode – YouTube help*. Available at: <https://support.google.com/youtube/answer/7354993?hl=en> (accessed 28 September 2017).
- Ziewitz M (2016) Governing algorithms: myth, mess, and methods. *Science, Technology & Human Values* 41(1): 3–16.

Author biographies

Clare Southerton is a postdoctoral research fellow in the Vitalities Lab, Social Policy Research Centre and Centre for Social Research in Health, UNSW Sydney. Her research focuses on how intimacy is formed with digital technologies and she has explored this in a number of empirical contexts including smartphone use, surveillance technologies, social media use and the intersection between everyday technologies and sexuality.

Daniel Marshall is an associate professor in the School of Communication and Creative Arts at Deakin University in Melbourne, Australia. He is also the Convenor of Deakin's Gender and Sexuality Studies Major in the Bachelor of Arts programme and of Deakin's Gender and Sexuality Studies Research Network. Current research focuses on queer youth and popular culture, queer

youth histories and queer archive studies. He is also currently a Chief Investigator in a national research team working on an ARC-funded Discovery Project called 'Queer Generations' studying how growing up LGBTQ has changed in Australia across two social generations.

Peter Aggleton is an Emeritus Scientia professor in the Centre for Social Research in Health, UNSW Sydney and a distinguished honorary professor in the College of Arts and Social Sciences at The Australian National University. He holds a visiting professorial position at UCL in the United Kingdom. He is an adjunct professor in the Australian Research Centre in Sex, Health and Society at La Trobe University in Melbourne. Peter has published more than 250 scientific papers and chapters and has authored and edited more than 50 books. He is well known internationally for his analytic work on health education and health promotion, the social aspects of HIV, sexuality and gender, and sexual and reproductive health and rights. He is editor-in-chief of three international journals: *Culture, Health & Sexuality*, *Health Education Journal* and *Sex Education* and is an associate editor of the journals *AIDS Education & Prevention*, *Global Public Health* and *Health Education Research*.

Mary Lou Rasmussen is a professor of sociology at the Australian National University College of Arts and Social Sciences. She has undertaken research in the United States, Canada, New Zealand and Australia. Her research focuses on building transdisciplinary understanding of sexuality and gender across diverse lifeworlds, taking account of issues related to sexual citizenship, cultural and religious difference and technologies of sexuality, education and health. She is co-editor, with Louisa Allen, of the *Handbook of Sexuality Education* (Palgrave).

Rob Cover is a professor in the School of Media and Communication at RMIT University. He is a social, media and cultural studies researcher whose work focuses on the implications of media and digital cultures for minorities, particularly in respect to health, social integration, diversity, ethics and belonging.