INTERNET POLICY REVIEW
Journal on internet regulation

Volume 10 **Issue** 4

RESEARCH ARTICLE

OPEN ACCESS

PEER REVIEWED

Whiteness in and through data protection: an intersectional approach to anti-violence apps and #MeToo bots

Renee Shelby *Northwestern University*
Jenna Harb *Australian National University*
Kathryn Henne *Australian National University*

DOI: https://doi.org/10.14763/2021.4.1589

Published: 7 December 2021
Received: 28 October 2020 **Accepted:** 12 April 2021

Funding: This research benefited from funding provided by the Australian National University Futures Scheme and by a Mellon/American Council of Learned Societies Fellowship.
Competing Interests: The author has declared that no competing interests exist that have influenced the text.
Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. https://creativecommons.org/licenses/by/3.0/de/deed.en
Copyright remains with the author(s).

Citation: Shelby, R., Harb, J., & Henne, K. (2021). Whiteness in and through data protection: an intersectional approach to anti-violence apps and #MeToo bots. *Internet Policy Review*, *10*(4). https://doi.org/10.14763/2021.4.1589

Keywords: Data protection, Race, gender, Artificial intelligence, Intersectionality

Abstract: This article analyses apps and artificial intelligence chatbots designed to offer survivors of sexual violence with emergency assistance, education, and a means to report and build evidence against perpetrators. Demonstrating how these technologies both confront and constitute forms of oppression, this analysis complicates assumptions about data protection through an intersectional feminist examination of these digital tools. In surveying different anti-violence apps, we interrogate how the racial formation of whiteness manifests in ways that can be understood as the political, representational, and structural intersectional dimensions of data protection.

This paper is part of **Feminist data protection**, a special issue of *Internet Policy Review* guest-edited by Jens T. Theilen, Andreas Baur, Felix Bieker, Regina Ammicht Quinn, Marit Hansen, and Gloria González Fuster.

Introduction

The use of mobile applications and artificial intelligence (AI) chatbots are an emerging strategy against sexual violence or gender-based violence (GBV). [1] While affordances vary, apps are generally "reactive" to violence. The majority aim to provide people experiencing GBV with one-time emergency assistance and a means to digitally report and build evidence against perpetrators (Eisenhut et al., 2020). Proponents frame these technologies—hereafter referred to as "anti-violence" apps—as empowering tools that support women through the accessible and anonymous transfer and processing of data. Global advocacy bodies, multinational corporations, and national governments increasingly endorse anti-violence apps as promising responses to GBV, [2] further legitimising their use and portrayal as a neutral and universal mode of protection.

Existing research on anti-violence apps focuses on their communication affordances (Eisenhut et al., 2020) and cryptographic strategies for reducing victims' hesitancy in reporting violence (Ayers & Unkovic, 2012). The few existing critical studies on anti-violence apps suggest they reinforce rape myths (Bivens & Hasinoff, 2018), strengthen surveillance structures (Mason & Magnet, 2012), and commodify violence within neoliberal capitalism (White & McMillan, 2020). As there is little scrutiny of data protection and inequality in relation to these technologies, this article deconstructs how anti-violence apps designed to report violence interact with forms of oppression in ways that complicate assumptions about privacy and data protection. Through an intersectional feminist examination, we interrogate how these technologies reflect and instil whiteness, the unnamed racial grammar that reinforces white logics as normative and, in turn, upholds racial hierarchies (Bonilla-Silva, 2011).

In the pages that follow, we outline how whiteness is embedded in the apps' de-

1. We use GBV to signal the connections between power, systemic oppressions, and acts of violence, recognising that GBV intersects with axes of domination, including classism, colonialism, homophobia, racism, and transphobia.
2. For example, the World Bank Group funds Armenia's Safe YOU mobile app, the United Nations Development Programme (UNDP) has endorsed Montenegro's 'Be Safe' mobile app, and the Yukon Human Rights Commission has adopted Spot AI to document and respond to various forms of abuse.

sign and materialises through user engagement, constitutively becoming the organising basis for data protection in these applications. This argument builds upon anti-racist and anti-colonial analyses that have examined how digital apps and AI chatbots default to whiteness in tone, language, and cultural tendencies by erasing racialised markers (Cave & Dihal, 2020; Phan, 2019). It also draws on critiques of feminisms that uphold whiteness by embracing legislative and criminal legal responses that increase precarity for Black, poor, trans, and women of colour (Combahee River Collective, 1982; Kim, 2018; Richie, 2000) over strategies that address interlocking oppressions and increase women's access to material resources to ameliorate vulnerability (Coker, 1999).

Aligned with past Critical Race Theory (CRT) analyses that underscore the importance of intersectionality, this article captures both the structural and discursive modes through which interlocking inequalities contribute to oppression (Crenshaw, 1991). We map the political, representational, and structural intersectional dimensions of data protection. In doing so, it illustrates how seemingly progressive mainstream feminist strategies to address GBV through technological interventions do not capture how gendered concerns intersect with racist, heteronormative, and classist logics. These concerns are especially relevant to the extractive features of digital platforms, which require personal information from users, emerge from public-private partnerships, and depend on problematic relationships with the criminal and civil legal systems. We conclude by explaining these tensions and discussing how anti-violence apps reveal a form of "colourblind intersectionality" (Carbado, 2013), which may yield dangerous unintended consequences that existing forms of data protection do not prevent.

The development of apps and AI for gender-based violence reporting

The nine reporting technologies examined here encompass smartphone and web applications designed to support victims, bystanders, and institutional advocates by carrying out GBV prevention and response strategies (see Table 1). Though each app's affordances vary, the most common functionality enables users to digitally report details of harassment or violence (e.g., location, nature of abuse, time, perpetrator characteristics, pictures of evidence). The resulting data is extracted and stored by a third party or the app's parent company and then processed using automated 'big data' analytics that employ algorithms to categorise and aggregate data for key stakeholders—such as human resources professionals, investigators, lawyers, and police—who serve as gatekeepers for institutional responses. While

there is a longer history of anti-violence movements creating technologies to prevent and respond to violence, including wearable defensive clothing, rape whistles, and evidence collecting devices (see Shelby, 2020), anti-violence apps in their current form have circulated for at least a decade with support from government initiatives, such as the Obama White House administration's (2011) *Apps Against Abuse Challenge*.

As the #MeToo movement renewed conversations about sexual violence, it accelerated broader calls to prevent and address GBV using new technologies. In 2018, the United Nations Human Rights Council (UNHCR) adopted the "Preventing and Responding to Violence Against Women and Girls in Digital Contexts" Resolution by a 73-country consensus. The Resolution supports funding initiatives aimed at "preventing, responding to, and protecting women and girls from violence in digital contexts" (UNHCR, 2018, p. 5). That same year the G7 committed to promoting anti-violence strategies and educational approaches that "keep pace with technological development" (G7 Research Group, 2018).Proponents often imagine anti-violence reporting apps as innovative solutions to GBV, particularly in terms of facilitating the confidential and accessible transfer of data. In a 2019 publication, UN Women asserted *Sis Bot*, a Thai AI chatbot offering information about legal services, provides "access to information safely and confidentially" (UN Women, 2019, para. 9). This view presumes digital GBV interventions transcend problems associated with offline approaches (Bivens & Hasinoff, 2018), and is grounded in the familiar construction of data privacy as the protection of individual privacy and liberties (e.g., Gellert et al., 2013), a construction codified for instance in the OECD's (2013) *Guidelines on the Protection of Privacy and Transborder Flows*.

Anti-violence apps, however, are not neutral mediators of justice. They can magnify entrenched cultural and political values, and the choices made by app developers reflect stereotypical assumptions about the causes and solutions to GBV. The features of anti-violence apps often reify pervasive racialised rape myths, such as "stranger danger" and victims' responsibility in managing GBV (Bivens & Hasinoff, 2018). They also tend to direct surveillance onto victims rather than perpetrators (Beaton, 2015) and enhance the relationships between anti-violence movements and the criminal legal system (Mason & Magnet, 2012), a system that perpetuates violence against women of colour. White and McMillan (2019, p. 1134) suggest other limitations, including device failure and misuse, evidentiary ambiguities, unanticipated uses, and the possibility of data being used against the survivor. In sum, while anti-violence technologies address some needs, they also compound and create new challenges.

Unveiling whiteness through intersectional analysis

Whiteness can be understood as a normative set of values that sustain and advance racial hierarchies (Bonilla-Silva, 2011), predicated on white supremacy (Sue, 2006). For beneficiaries of whiteness, its hegemonic qualities are often unseen, even though they actively participate in its maintenance through "racial stereotypes and biases (a beliefs aspect), racial metaphors and concepts (a deeper cognitive aspect), racialised images (the visual aspect), racialised emotions (feelings), interpretive racial narratives, and inclinations to discriminate within a broad racial framing" (Feagin, 2013, p. 91). White supremacist power dynamics have co-opted the mainstream #MeToo movement (Phipps, 2019; Tambe, 2018), and in public digital feminisms more broadly, reproduced race-neutral ideologies (Patil & Puri, 2021) and centred white, middle-class women as a universal subject position (Travers, 2003). In the pages that follow, we explore how whiteness manifests in the designs of anti-violence apps, a mainstream feminist intervention, encoding race neutrality and heteronormativity within assumptions about victims of GBV and their needs.

Building on Crenshaw's (1991) earlier work on GBV against women of colour, this article focuses on three dimensions of intersectionality: political, representational, and structural. Political intersectionality captures the discursive practices of marginalisation as they cross lines that surpass narrow understandings of gender. In this case, race-neutral data protection policies and practices reveal tacit investments in whiteness. Representational intersectionality encompasses imaginary modalities that render women of colour—and, in this analysis, also nonbinary and poor persons—invisible in anti-violence discourse. Instead, anti-violence app companies draw on high-profile and stereotypical representations of #MeToo perpetrators to produce software that protects racialised institutions as much, if not more, than survivors. Structural intersectionality captures how structural conditions and socio-economic disadvantage disproportionately affect individuals and groups who experience multiple forms of marginalisation. Here, anti-violence apps offer modest and short-term solutions that obscure the need for broader social transformation to address racialised precarity and vulnerability to violence. As these dimensions are often interrelated in terms of how they are experienced, our conclusion reflects on how they cumulatively reveal a form of colourblind intersectionality that feminist data protection must address.

Political intersectionality: how intersecting inequalities are relevant to data protection

The privacy laws governing internet and communication platforms treat data as property by setting the terms of data transactions and the resulting "data subject"—which, according to the EU's General Data Protection Regulation (GDPR), is a person whose "personal data" are processed by a business or organisation (Article 4.1). Political intersectionality draws attention to how inequalities intersect across race, class, gender, and sexuality. They are relevant to data protection policies, their political agendas, and the *data capital* they produce, which refers to how collected data accumulates and circulates as a business asset (Sadowski, 2019). In this case, universalist data protection policies purport to be equal by offering blanket protections to all platform users; however, they reveal tacit racial investments by reaffirming whiteness through race-neutral logics. In other words, whiteness contributes to the re-creation of a universal subject position, but its modes of doing so are often rendered invisible.

As explored in this section, data protection policies and practices organised around a single axis of inequality—in this case, gender—universalise whiteness as an attribute of survivors using these apps. They do so in three interrelated ways: (1) as data protection laws and policies that individualise protection and negate the intersections of race and sexuality, (2) through their privacy functionality, which produces a colourblind aesthetic of anonymity, and (3) through the erasure of race and sexuality in anti-violence data. These sociotechnical solutions for addressing and intervening in such violence, in turn, disregard the experiences of people of colour and LGBTQ+ individuals, because they fail to capture how ableism, heteronormativity, racism, and sexism collude to sustain forms of discrimination and violence (Gray, 2020).

The nine reporting apps we examined in depth, based in Australia, Canada, and the United States, are governed by a constellation of laws and regulations. [3] Adding further complexity, data protection law can traverse national boundaries. In Europe, for example, the GDPR applies in geographically expansive ways, covering the territory of the EU and any companies offering goods and services or monitor-

3. In the United States, for example, hundreds of laws, including the 1974 Privacy Act (5 U.S.C. § 552a), the Electronic Communications Privacy Act (Pub.L. 99–508§), the California Consumer Privacy Act (effective through 2022), and the Computer Fraud and Abuse Act (18 U.S. Code § 1030), have been enacted at the state and federal level. In Canada, digital technologies and related data issues are governed by the Personal Information Protection and Electronic Documents Act (PIPEDA), and in Australia, The Privacy Act (1988) and the Australian Privacy Principles provide the foundation for Australia's regulation of personal information through digital platforms at the time of publication.

ing behaviour of users in the EU (e.g., tracking cookies or IP addresses) (see Article 3.1). Thus, some US-based applications, such as AllVoices and Talk to Spot, can be subject to GDPR provisions. Further scrutiny of the GDPR demonstrates how its regulatory logic does not prevent "data sanitation" and creates conditions of "intersectional invisibility" where dominant group identities come to stand in as categorical norms (Purdie-Vaughns & Eibach, 2008, p. 378). Because certain categories become omitted when collecting information about users, whiteness becomes central to shaping a cognisable, yet seemingly invisible intersectional data subject position. [4] Here, we explain how the GDPR's concern with *what* is processed overlooks *how* data categories are used at any given moment or *how* protection is racialised.

The EU Charter of Fundamental Rights (Article 8) enshrines the right to the protection of personal data and for personal data to be processed fairly. Predicated on this right, the primary project of "protection" in the GDPR is to regulate the processing of personal data. Article 9 of the GDPR prohibits the processing of data related to race or ethnic origin and sexual orientation, asserting the need for special protections on grounds data could expose one to discrimination and create risks to fundamental rights and freedoms—with exceptions only allowed with user consent or in exceptional situations (European Union, 2016). As GDPR protections are transactional and identity-based, only certain identity data categories are accounted for essential sources of discrimination. Other categories, including those that might capture social formations that contribute to oppression, are rendered unnecessary and become framed as privacy risks.

By design, anti-violence apps are gynocentric technologies; that is, based on analysis of language used in advertisements and on their platforms, the imagined prototypical user is typically a cisgender woman. None of the anti-violence apps we analysed collect information about ethnicity, race, or sexuality (see Table 2). In line with the GDPR and other race-neutral data protection laws, for example, Callisto notes in their privacy policy they "strongly discourage" users from "emailing [them] any sensitive information about you or anyone else, including . . . race or ethnicity. . . and sexual orientation" (Callisto, 2020, You Contact Us Directly section, para. 2). The apps do, however, collect other personal information, including gender, age, date of birth, name, and occupation. While choosing not to report would render someone "missing" in anti-violence app data sets, race-neutral data protection takes-up a formal conception of equality whereby ignoring race remedies discrimination. Marking ethnicity, race, and sexual orientation as special data cate-

4. For further analysis on data protection and the GDPR, see Gellert et al. (2013).

gories has consequences for how data subjects are understood as reflections of users—in this case, they appear similar to members of dominant groups. Anti-violence reporting apps' demographic flattening of its users reifies the false notion there is a common experience of violence. As such, they uphold white heteronormative logics that have become a central critique of mainstream digital anti-violence movements in European countries and settler states, such as Australia, Canada, and the United States (Daniels, 2016). In doing so, protection centres the presumption of white cis-gender women's experience, rendering trans, non-binary, and non-white users invisible.

These concerns are often side-stepped in reporting apps through a promise of data protection as "anonymity". Six of the nine anti-violence apps we reviewed (#NotMe, AllVoices, Botler, Hello Cass, JDoe, Talk to Spot) frame themselves as anonymous reporting platforms, providing a type of protection that can be understood as the system knowing the actions, but not the identity, of a user. Callisto positions itself as a confidential platform, using end-to-end encryption to protect user data. On one hand, there is evidence to suggest anonymous platforms may increase self-disclosure through online channels (Taddei & Contenta, 2013), as face-to-face help-seeking can become a disciplining process where survivors must work to conform to cis-gendered and white racialised normative beliefs that a victim is hyper-feminine and passive (Guadalupe-Diaz & Jasinski, 2017). On the other hand, in researching survivors' concerns with online reporting platforms, Obada-Obieh, Spagnolo, and Beznosov (2020) found women worry abusers could find out they accessed the system, subjecting them to further harm. Conceptualising data protection as anonymity does not preclude abusers from digitally tracking their victims' online movements, nor does it offer a mechanism to transcend racial privilege and other forms of oppression when institutional authorities review reports submitted through these platforms. In this way, reporting app users may still discipline their recounting of violence to align with (white) racialised expectations of survivorship.

Moreover, anti-violence apps' claims of anonymity are more a discursive promise than meaningful privacy practice. Even when the stated purpose of the app is to facilitate basic information gathering, "anonymous" apps collect significant personal information. Although #NotMe and JDoe present themselves as "anonymous" platforms, the first step to using their products is a prompt to provide first and last name, date of birth, and phone number. Botler [5]—an AI that scans through Canadian and US sexual harassment cases to determine if there is cause for legal action—collects first and last name, gender, email address, birth date, residential lo-

5. See https://botler.ai/privacy-policy

cation, and employment title. These practices of collecting identifying information, which can, in turn, be circulated as data capital are commonplace (see Tables 2 and 3). The use of data as an asset attests that the drive to produce and accumulate data is as important as—and may even supersede—the anti-violence service itself. As many anti-violence apps rely on for-profit business models, this common tendency reflects one way these technologies undergo "private sector symbiosis" (Hess, 2005, p. 516) in which social movement technologies come to reflect goals of entrepreneurs and industry innovators.

In complying with data protection policies, anti-violence apps undermine the potential to acknowledge and mobilise against the interconnected ways racism, sexism, and heteronormativity produce violence. As Crenshaw (1991, p. 1242) notes, "ignoring differences *within* groups frequently contributes to tension *among* groups, another problem of identity politics that frustrates efforts to politicise violence against women". Race-neutral data protection practices become yet another means to uphold the hegemony of white feminism. [6] Although non-white, trans, and queer persons are representationally invisible in anti-violence user data sets, algorithmic profiling may still produce the discriminatory processing that Article 9 of the GDPR seeks to prevent. As Mann and Matzner (2019, p. 2) note, "discriminatory effects also occur if data on discriminatory features like gender, race, ethnicity, etc. are not directly processed… [as] algorithmic profiling can easily identify 'proxies'". While there are no known algorithmic bias studies on anti-violence apps to assess whether the software produces inequitable outcomes, without a race-conscious design, it is not difficult to imagine how these apps reproduce offline biases. For example, authorities may vet digital reports in ways that tacitly endorse racialised rape myths, resulting in the same dismissal outcomes as offline reporting. Further, anti-violence apps can become sites of political disempowerment when they make queer persons and women of colour invisible in reports and data sets. Thus, they undermine the liberatory potential of these apps to contest on-the-ground practices of domination and oppression.

Representational intersectionality: institutional "bad apples" and protection as legal recourse

With the exception of Callisto, the anti-violence apps we examined emerged against the backdrop of the mainstream #MeToo movement and its goals of em-

6. White feminism centers the experiences of white women by focusing solely on patriarchy and failing to examine how the experiences of Black, Indigenous, and other women of colour are shaped by racism, settler colonialism, and other interlocking forms of oppression. In articulating gender equality without attention to racialisation, white feminism upholds white supremacy.

powering survivors by bringing repeat offenders to account, often through a punitive justice lens (Hayes & Kaba, 2018; Phipps, 2019). The apps draw explicit connections to mainstream #MeToo narratives in which institutional bad actors that enable perpetrators are an exceptional, yet ubiquitous problem given their ability to harm survivors *and* institutions. Anti-violence apps' self-constructed hero narratives suggest their technologies empower survivors and safeguard institutions. This section reflects on how these portrayals are etched and shaped by whiteness. Our point is not to minimise how institutional gatekeepers can be hostile to survivors. Rather, we assert this framing of "bad actors" masks how institutions are racialised power structures and how the legibility of violence in institutional gatekeeping is shaped by racialised images of gender and sexuality.

Here, anti-violence reporting apps become part of a #MeToo techno-solutionism for addressing the persistence of violence. They illustrate how data protection practices mutually shape and are shaped by stereotypical representations, particularly perceived notions of who survivors and perpetrators are. In the apps, the imagined wound created by repeat offenders and bad actors extends victimisation beyond the survivor to harm institutions; this wound can be sutured through well-placed technological intervention. While "bad apple" institutional accomplices appear in many #MeToo narratives (Orgad & Gill, 2019), they feature prominently in the high-profile Larry Nassar sex abuse scandal. Nassar, a former USA Gymnastics and Michigan State University doctor, has been accused by more than 350 young women and girls of assault over the course of nearly 40 years. Condemnation of Nassar has been eclipsed only by outrage over the institutional actors that protected Nassar.

When the story broke in 2016, the media and criminal investigations into why institutional accomplices at Michigan State University, USA Gymnastics, the US Olympic Committee, local law enforcement, and even the FBI failed to hold Nassar accountable focused primarily on specific self-interested personalities (McPhee & Dowden, 2018).The findings of a US Senate investigation into the Nassar case illustrates how the narration of institutional power slips into condemnation of individual "bad apples". According to the Subcommittee:

> *Nassar committed his criminal sexual conduct by himself, but multiple institutions responsible for keeping amateur athletes safe failed to adequately respond to credible allegations against Nassar. . . Repeatedly, institutions failed to act aggressively to report wrongdoing to proper law enforcement agencies. Repeatedly, men and women entrusted with positions of power prioritized their own reputation*

> *or the reputation of an NGB over the health and safety of the athletes.* (The Courage of Survivors, p. 2)

This framing of bad institutional actors does little to protect the interests of survivors, especially those most likely to be ignored by institutions—people who are socially and legally precarious, including LGBTQ+, poor, and people of colour (Richie, 2013). On one hand, this depiction of "perpetrator" in reporting apps overlooks how white male heterosexuality has historically retained a place of legitimate sexual dominance; on the other, it permits a framing of the institution as largely benevolent bureaucratic structure merely in need of disciplining. In doing so, reporting technologies allow what Ray (2019) terms "racialised decoupling" in which organisations decouple commitments to equity from practices that reinforce existing hierarchies that disadvantage non-white employees.

This depiction of the institutional accomplice plays out in response to the Nassar case through two forms of response: (1) the effort to replace institutional accomplices, which locates the accountability problem among individual actors; and (2) the introduction of a new third-party governing body and anti-violence app, SafeSport. Imagined as a reliable authority, SafeSport allows survivors to perform a kind of technically sanctioned victimhood—where their reports of violence are captured, preserved, and circulated as data capital. Rather than a success story, SafeSport illustrates how technical solutions often lack the capability of delivering institutional accountability.

Within a year of launching SafeSport, the organisation received over 1,800 reports through its reporting platform. An investigation by USA Today found underfunding of SafeSport hindered its ability to investigate those 1,800 reports, and there is little to no enforcement of sanctions for sexual misconduct (Axon & Armour, 2019). Although the design of this reporting technology and its data protection practices were to discipline the institutional "bad apple" figure, its embedded imagining casted villainous gatekeeping as a problem that could be resolved by using so-called algorithmic expertise to demand institutional accountability through the streamlining of reporting and investigation. However, in using the extreme case of Nassar as the impetus for accountability, efforts to hold less prolific perpetrators accountable may fall short.

Numerous anti-violence apps tout the affordances of anonymity (#NotMe, AllVoices, Botler, Hello Cass, SafeSport, Talk to Spot), matching algorithms and the information escrow (Callisto, JDoe) for holding repeat offenders accountable. JDoe, for

example, states it is:

> *constantly grouping reports together that concern the same perpetrator. This is done algorithmically, without ever storing the reports in a form that we could access. When two or more users report the same perpetrator, they are given the option to contact a local law firm.* (n.d., para. 1)

While legal remedies to violence have never been a panacea, especially for precarious survivors, these affordances present technological expertise as capable of countering the biases of institutional gatekeepers. Talk to Spot co-founder Julia Shaw reaffirms this point, stating, "By taking an evidence-based approach to help break down the barriers to reporting harassment and discrimination, Spot allows companies to provide timely, transparent, and unbiased responses and mitigate the negative consequences of harassment" (Dickey, 2018, para. 3). Yet, race and sexual orientation particularise narratives of sexual violence. People of colour and sexual minorities are more likely to be blamed for their assaults than their heterosexual counterparts and authorities have historically rendered these survivors' claims of assault suspect (Richie, 2012).

To date, no anti-violence app companies have provided information about how they design their algorithms or train their AIs to detect and mitigate bias. While this approach may be understood as a model of accountability via algorithmic transparency, unveiling the black box is not the same as holding a socio-technical system accountable; in fact, it arguably obscures how algorithms enact social complexity (Ananny & Crawford, 2016). As Louise Amoore (2020) explains further, making algorithms transparent with the aim of removing bias does not fully attend to the nature of how algorithms operate: they are not only predicated on categorisation, value judgments, and assumptions, but they also learn and adapt through iterative engagement. In this case, reporting apps adapt by shifting the site of protection from the precarious person to the institution. In fact, the marketing of anti-violence apps tacitly conveys them as uniquely able to discipline survivors who pose a threat to institutions by failing to report behaviours of abusers—actions that are presumed to undermine the integrity of organisations and to carry financial implications.

Consider, for example, how #NotMe explains how employee inaction threatens the productivity of the institution:

> *Our deep experience in this space has led us to understand that misconduct is not an HR [human resources] problem; it is a people problem. People are often frozen in fear of speaking up and seeking help from their organisations; thus perpetuating unchecked repeat offenses, reducing workplace productivity, and ultimately causing high turnover and high legal costs that often go hand-in-hand with unpleasant media headlines.* (n.d., Why We're the Experts section, para. 2 & 3)

The primary function of these platforms, which is to transform user inputs into data capital valuable and suitable for legal and institutional discipline (e.g., reports), shores up the figure of the institutional accomplice and centrality of the imagined white, cisgender female user. They do so by conducting interviews with survivors or bystanders, reformatting them into evidentiary reports with time stamps and offering guidance on how to interact with legal and institutional authorities.

Five of the platforms (Botler, Callisto, JDoe, SafeSport, Vesta) uphold the dominant paradigm that reporting is primarily for the purposes of punishment. They automate legal connections but fail to encode access to emotional and support services. #NotMe, AllVoices, and Talk to Spot strengthen HR reporting channels and promote workplace disciplinary action by sending reports to HR rather than to direct supervisors. Without grounding accountability in a theory of transformative justice (Kim, 2018), these automating practices may have unintended consequences, particularly for women of colour who are least likely to be believed by police (Campbell et al., 2015) and more vulnerable to workplace discrimination despite Civil Rights Protections (Light et al., 2011). Making survivors more legible to HR systems may make them more likely to be unfairly targeted or even punished. By favouring a disembodied universalism of white womanhood, these data practices fail to account for embodied inequalities that survivors experience.

Structural intersectionality: material constraints that challenge data protection

Emphasis on data protection in anti-violence apps implies various avenues of defence for survivors of violence, such as data safeguards, including encryption, and autonomous access to services. This sense of digitally facilitated security obfuscates intersecting structural inequalities. By design, technological interventions are not equipped to address entrenched causes of violence or the societal systems shaping them. As Sokoloff and Dupont (2005, p. 55) explain, a structural approach to combating GBV offers material resources "to the poorest and most disadvantaged" survivors to meet their immediate needs alongside institutional reform

through advocacy coalitions and "monitoring police, prosecutorial, and judicial responses". In contrast, anti-violence apps offer modest and short-term relief. Their emphasis on data protection may justify their efficacy, but it also shifts attention away from deeper structures that enable GBV. By eliding these structural dimensions, anti-violence apps can sustain and reinforce oppression by failing to challenge "routinised forms of domination that often converge" (Crenshaw, 1991, p. 1245).

The emphasis on privacy as a key element of safety offers a case in point. Anti-violence apps reinforce messages that users play a key part in being the solution to ending GBV. #NotMe states, "If you don't report, it will likely continue. And possibly get worse — for you and others. Speaking up is in service to all" (n.d., Speak Up Safely section, para. 3). In doing so, the focus on individual onus not only shifts responsibility and risk onto disempowered actors rather than institutions (see McDermott, Henne, & Hayes, 2018), but also reinforces the idea that interaction with the legal system is a desirable outcome—even though there is ample evidence law enforcement personnel and jurists discriminate against survivors, particularly those who are people of colour, when they report violence, attend court, and attempt to access services (Campbell et al., 2015; Crenshaw, 1991).

The Brazilian organisation Think Eva is an exception: in addition to documenting, tracking, and analysing forms of harassment over the internet, it actively tackles structurally oriented dimensions of GBV through activism and the dissemination of information meant to enhance social awareness (Das, 2020). Its sister organisation, a non-governmental organisation (NGO) called Think Olga, participates in social justice campaigns and offers free resources on the complexity of GBV, which highlight the experiences of those who occupy disenfranchised and marginalised social positions.

Many anti-violence proposals reliant on technological solutions take access to secure internet and quality digital services for granted. In the United States, for instance, 26% of adults with an annual income below US$30,000 are smartphone dependent, meaning they do not own a computer or have broadband internet at home (Anderson & Kumar, 2019). This poses challenges for accessing anti-violence apps that are not optimised for mobile devices, such as Botler, Callisto, and Talk to Spot. Considered alongside evidence that more than 12% of Black and Hispanic Americans rely on smartphones for internet access, and 63% of rural adults do not have home broadband (Anderson & Kumar, 2019), it becomes clear the obstacles to using digitised anti-violence platforms retain structural intersectional contours. Reliance on public wireless internet or public computers enhances the risk of hav-

ing details of their personal information intercepted or used (Sombatruang, Sasse, & Baddeley, 2016). Further, inequities, such as precarious employment, feminised low-wage employment, and housing insecurity, can undermine one's ability to retain control over one's data, a core tenet of data protection (Bellanova, 2017). Hello Cass, for instance, allows users to "request details of the personal information that we hold about you" (n.d., Your Rights and Controlling Your Personal Information section, para. 5), noting they may be required to pay a fee to receive such information.

Several structural intersectional concerns link to issues of representational intersectionality. For example, "stereotypes of women of colour, lack of trust of outsiders and public officials, and fear of reporting due to the potential for discriminatory treatment and further violence" (Bent-Goodley, 2009, p. 263) materialise as significant barriers for racialised groups, which reporting through an app cannot overcome. Representational stereotypes often inform the avenues for support presented by these applications. Callisto, for instance, has a webpage for navigating options, including others to consult, such as an attorney, a therapist, or the "Whisper Network", which provides informal channels of information, such as shared online spreadsheets, to document experiences of violence and flag perpetrators to make others aware of their behaviour (Tuerkheimer, 2019). Despite listing fifteen options, Callisto does not capture services that may be able to assist in managing discriminatory effects of ableism, heteronormativity, or racism. In doing so, these and other design features fail to account for the diverse realities of those who experience violence.

JDoe's incident hotspots mapping, which visually charts incident locations based on geolocation data recorded when users report, offers another example. It enables users to set geofencing alerts to notify them when they are close to areas with high levels of reported incidents. Although hotspots mapping is justified as a tool for developing "a new level of clarity about your surroundings" (n.d., Visualize section, para. 1) [7], such features fuel tacit beliefs that certain areas are unsafe. It communicates a form of "stranger danger" (Bivens & Hasinoff, 2018), a misguided understanding of GBV that posits public spaces as particularly dangerous for women without drawing critical attention to structural patterns that contribute to GBV.

In thinking about the design of anti-violence apps in relation to inequitable condi-

7. Given the lack of transparency around the data informing maps and the low levels of reporting GBV, there are significant questions about the accuracy and reliability of these features.

tions, the role of data capital is an important consideration. Collecting volumes of personal information, whether through cookies (AllVoices, Callisto), web beacons (AllVoices), geolocation (JDoe), or targets modes of gathering "a range of Personally-Identifying Information" (JDoe) is a condition of use; it applies when registering for anti-violence services, engaging in surveys or contests, communicating with the company, or receiving newsletters. As noted in Callisto and Spot's privacy policies, even if users enable 'Do Not Track' on their browsers, Google continues to use persistent cookies and Callisto and JDoe continue to log and store metadata (Callisto Privacy Policy; Spot Privacy Policy). Moreover, the emphasis on data collection as care often presents a coercive exchange: if survivors want to access quality services, they must allow the collection of more data—for example, by enabling cookies (AllVoices) or providing personally-identifying information (JDoe). These practices not only go against protection imperatives, such as the principle of 'data minimisation', as enshrined in the GDPR and other international data protection guidelines (e.g., Kuner and Marelli, 2020; Privacy International, 2018), they reflect a conscious choice to selectively adhere to data protection principles, negating those that conflict with the generation of data capital and potential profits.

Thus, there are structurally grounded reasons to question the role of private technology companies in anti-GBV efforts. Apps such as AllVoices actually subcontract data processing to undisclosed entities that are presented as 'trusted partners', including service providers that can be located outside one's own country of residence. Such partnerships are only one element of the ecosystem that enables anti-violence apps to collect, store, track, and analyse internet activity and personal data; compromise in these nodes can have reverberating effects. Further, people of colour often have grounded reasons for not trusting authorities with their data, which has been used for targeted surveillance and policing of racialised groups (Browne, 2015). This concern is salient considering growing evidence of poor data protection from private companies, including Google Analytics, which is used by several anti-violence platforms (AllVoices, Callisto, Hello Cass), and was exploited in 2020 to transfer sensitive personal data such as passwords and credit card numbers from hacked sites (Vlasova, 2020). Contemporary cases of malicious coding reinforce that, despite claims of state-of-the-art data protection measures, they have vulnerabilities that can be exploited and put users' information at risk. As anti-violence app users are already vulnerable, these data risks can have significant repercussions for survivors' wellbeing. Collectively, anti-violence apps operate through an interconnected ecosystem of sociotechnical components and partnerships with multi-sectoral actors. The required extraction of data for service provision is an inherently unequal exchange that is skewed towards technology companies and

their partners. Existing structural inequalities exacerbate the associated risks for users.

Conclusion

This analysis illustrates various modes through which anti-violence reporting apps contribute to the erasure of important political, representational, and structural dimensions of GBV. As the imagined survivor embedded and privileged in anti-violence apps is presumed to be white, cisgender, and heterosexual, these digital tools contribute to a form of "colourblind intersectionality" in which this unarticulated and seemingly invisible subject position does racially constitutive work (Carbado, 2013). As scholarly critiques of digital technologies attest, colourblind dynamics often exacerbate racialised stereotypes and contribute to discrimination (Cave & Dihal, 2020; Phan, 2019). Anti-violence app interventions and the data subjects they produce become whitewashed in ways that obscure how users' experiences of sexual violence and interactions with institutional support systems are rooted in interlocking systems of oppression. As a result, users who use these apps but do not fit normative whiteness are more likely to experience unintended consequences, including neglect or punitive outcomes.

In sum, anti-violence apps do not overcome long standing intersectional criticisms of legal responses to GBV: that they are poorly equipped to support survivors who occupy more than one axis of oppression and often perpetuate further violence against them (e.g., Crenshaw, 1991; Richie, 2013). Our findings therefore raise serious questions about how these platforms comprise a new front-line approach for addressing GBV. As such, they are instructive for the future of feminist data protection. Foremost, feminist conceptualisations of data protection must confront and dismantle the ways that whiteness operates as the normative lens for collecting and processing data, which requires addressing structural conditions of GBV and the imbalances exacerbated by the drive for data capital. When framed in this way, it becomes clear these apps' strategies are not sufficient for protecting persons whose data is being harvested and used through these platforms, especially in relation to those who experience marginalisation along social categories of difference, such as race, ethnicity, or sexuality. Anti-violence reporting interventions that fail to address the contexts of systemic violence and precarity will never be transformative. Without more radical corrective strategies and alternative conceptualisations of protection, anti-violence apps are bound to perpetuate interlocking forms of oppression.

References

Amoore, L. (2020). *Cloud ethics: Algorithms and the attributes of ourselves and others*. Duke University Press.

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, *20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Anderson, M., & Kumar, M. (2019). *Digital divide persists even as lower-income Americans make gains in tech adoption*. Pew Research Center. https://www.pewresearch.org/fact-tank/2021/06/22/digital-divide-persists-even-as-americans-with-lower-incomes-make-gains-in-tech-adoption/

Axon, R., & Armour, N. (2018, December 28). SafeSport CEO Shellie Pfohl will step down at year's end. *USA Today Sports*. https://eu.usatoday.com/story/sports/olympics/2018/12/28/shellie-pfohl-safesports-ceo-stepping-down/2438229002/

Ayers, I., & Unkovic, C. (2012). Information escrows. *Michigan Law Review*, *111*(2), 145–196.

Beaton, B. (2015). Safety as net work: "Apps against abuse" and the digital labour of sexual assault prevention. *MediaTropes*, *5*(1), 105–124.

Bellanova, R. (2017). Digital, politics, and algorithms: Governing digital data through the lens of data protection. *European Journal of Social Theory*, *20*(3), 329–347. https://doi.org/10.1177/1368431016679167

Bent-Goodley, T. B. (2009). A Black Experience-Based Approach to Gender-Based Violence. *Social Work*, *54*(3), 262–269. https://doi.org/10.1093/sw/54.3.262

Bivens, R., & Hasinoff, A. A. (2018). Rape: Is there an app for that? An empirical analysis of the features of anti-rape apps. *Information, Communication & Society*, *21*(8), 1050–1067. https://doi.org/10.1080/1369118X.2017.1309444

Bonilla-Silva, E. (2011). The invisible weight of whiteness: The racial grammar of everyday life in contemporary America. *Ethnic and Racial Studies*, 1–22. https://doi.org/10.1080/01419870.2011.613997

Browne, S. (2015). *Dark Matters: On the Surveillance of Blackness* (p. dup;9780822375302/1). Duke University Press. https://doi.org/10.1215/9780822375302

Callisto. (n.d.). *Privacy policy*. https://mycallisto.org/privacy-policy

Campbell, R., Shaw, J., & Fehler–Cabral, G. (2015). Shelving Justice: The Discovery of Thousands of Untested Rape Kits in Detroit. *City & Community*, *14*(2), 151–166. https://doi.org/10.1111/cico.12108

Carbado, D. W. (2013). Colorblind Intersectionality. *Signs: Journal of Women in Culture and Society*, *38*(4), 811–845. https://doi.org/10.1086/669666

Cave, S., & Dihal, K. (2020). The Whiteness of AI. *Philosophy & Technology*, *33*(4), 685–703. https://doi.org/10.1007/s13347-020-00415-6

Coker, D. (1999). Shifting power for battered women: Law, material resources, and poor women of color. *UC Davis Law Review*, *33*, 1009.

Combahee River Collective. (1982). A Black feminist statement. In G. T. Hull, P. Bell-Scott, & B. Smith (Eds.), *All the women are White, all the Blacks are men, but some of us are brave: Black women's studies*. Feminist Press.

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, *43*(6), 1241. https://doi.org/10.2307/1229039

Daniels, J. (2016). The trouble with white (online) feminism. In S. U. Noble & B. M. Tynes (Eds.), *The intersectional internet: Race, class, and culture online* (Vol. 105, pp. 41–60).

Das, S. (2020, January 6). AI to combat sexual harassment with chatbots, apps & trained algorithms. *Analytics India Magazine*. https://analyticsindiamag.com/ai-to-combat-sexual-harassment-with-chatbots-apps-trained-algorithms/

Dickey, M. R. (2018, October 23). Spot launches chatbot for HR departments to address harassment and discrimination. *TechCrunch*. https://techcrunch.com/2018/10/23/spot-launches-chatbot-for-hr-departments-to-address-harassment-and-discrimination/

Eisenhut, K., Sauerborn, E., García-Moreno, C., & Wild, V. (2020). Mobile applications addressing violence against women: A systematic review. *BMJ Global Health*, *5*(4), e001954. https://doi.org/10.1136/bmjgh-2019-001954

European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Feagin, J. R. (2013). *The White Racial Frame Centuries of Racial Framing and Counter-Framing, Second Edition*. Taylor and Francis. http://www.123library.org/book_details/?id=106184

G7 Research Group. (2018). *Charlevoix commitment to end sexual and gender-based violence, abuse and harassment in digital contexts*. http://www.g7.utoronto.ca/summit/2018charlevoix/violence-commitment.html

Gellert, R., de Vries, K., de Hert, P., & Gutwirth, S. (2013). A Comparative Analysis of Anti-Discrimination and Data Protection Legislations. In B. Custers, T. Calders, B. Schermer, & T. Zarsky (Eds.), *Discrimination and Privacy in the Information Society* (Vol. 3, pp. 61–89). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30487-3_4

Gray, K. L., & Sarkeesian, A. (2020). *Intersectional tech: Black users in digital gaming*. Louisiana State University Press.

Guadalupe-Diaz, X. L., & Jasinski, J. (2017). "I Wasn't a Priority, I Wasn't a Victim": Challenges in Help Seeking for Transgender Survivors of Intimate Partner Violence. *Violence Against Women*, *23*(6), 772–792. https://doi.org/10.1177/1077801216650288

Hayes, K., & Kaba, M. (2018, February 5). The sentencing of Larry Nassar was not 'transformative justice'. Here's why. *The Appeal*, 58–62.

HelloClass. (n.d.). *Privacy policy*. https://hellocass.com.au/privacy-policy/

Hess, D. J. (2005). Technology- and Product-Oriented Movements: Approximating Social Movement Studies and Science and Technology Studies. *Science, Technology, & Human Values*, *30*(4), 515–535. https://doi.org/10.1177/0162243905276499

JDoe. (n.d.). *Features*. https://jdoe.io/html/features.html

Kim, M. E. (2018). From carceral feminism to transformative justice: Women-of-color feminism and alternatives to incarceration. *Journal of Ethnic & Cultural Diversity in Social Work*, *27*(3), 219–233. https://doi.org/10.1080/15313204.2018.1474827

Kuner, C., & Marelli, M. (2020). *Handbook on Data Protection in Humanitarian Action*. International Committee of the Red Cros. https://www.icrc.org/en/data-protection-humanitarian-action-handbook

Light, R., Roscigno, V. J., & Kalev, A. (2011). Racial Discrimination, Interpretation, and Legitimation at Work. *The ANNALS of the American Academy of Political and Social Science*, *634*(1), 39–59. https://doi.org/10.1177/0002716210388475

Mann, M., & Matzner, T. (2019). Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. *Big Data & Society*, *6*(2). https://doi.org/10.1177/2053951719895805

Mason, C. L., & Magnet, S. (2012). Surveillance Studies and Violence Against Women. *Surveillance & Society*, *10*(2), 105–118. https://doi.org/10.24908/ss.v10i2.4094

McDermott, V., Henne, K., & Hayes, J. (2018). Shifting risk to the frontline: Case studies in different contract working environments. *Journal of Risk Research*, *21*(12), 1502–1516. https://doi.org/10.1080/13669877.2017.1313764

McPhee, J., & Dowden, J. P. (2018). *Report of the independent investigation: The constellation of factors underlying Larry Nassar's abuse of athletes*. Ropes and Gray.

#NotMe. (n.d.). *Individuals*. https://help.not-me.com/for-individuals

#NotMe. (n.d.). *Why we're the experts*. https://help.not-me.com/why-were-the-experts

Obada-Obieh, B., Spagnolo, L., & Beznosov, K. (2020). Towards understanding privacy and trust in online reporting of sexual assault. *Sixteenth Symposium on Usable Privacy and Security (SOUPS) 2020*, 145–164.

O.E.C.D. (2013). *The OECD Privacy Framework*. https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

Orgad, S., & Gill, R. (2019). Safety valves for mediated female rage in the #MeToo era. *Feminist Media Studies*, *19*(4), 596–603. https://doi.org/10.1080/14680777.2019.1609198

Patil, V., & Puri, J. (2021). Colorblind Feminisms: Ansari-Grace and the Limits of #MeToo Counterpublics. *Signs: Journal of Women in Culture and Society*, *46*(3), 689–713. https://doi.org/10.1086/712078

Phan, T. (2019). Amazon Echo and the Aesthetics of Whiteness. *Catalyst: Feminism, Theory, Technoscience*, *5*(1), 1–38. https://doi.org/10.28968/cftt.v5i1.29586

Phipps, A. (2019). "Every Woman Knows a Weinstein": Political Whiteness and White Woundedness in #MeToo and Public Feminisms around Sexual Violence. *Feminist Formations*, *31*(2), 1–25. https://doi.org/10.1353/ff.2019.0014

Privacy International. (2018). *A guide for policy engagement on data protection: The keys to data protection*. Privacy International. https://privacyinternational.org/sites/default/files/2018-09/Data%20Protection%20COMPLETE.pdf

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional Invisibility: The Distinctive Advantages and Disadvantages of Multiple Subordinate-Group Identities. *Sex Roles*, *59*(5–6), 377–391. https://doi.or

g/10.1007/s11199-008-9424-4

Ray, V. (2019). A Theory of Racialized Organizations. *American Sociological Review*, *84*(1), 26–53. https://doi.org/10.1177/0003122418822335

Richie, B. (2012). *Arrested justice: Black women, violence, and America's prison nation*. New York University Press.

Richie, B. E. (2000). A Black Feminist Reflection on the Antiviolence Movement. *Signs: Journal of Women in Culture and Society*, *25*(4), 1133–1137. https://doi.org/10.1086/495533

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, *6*(1), 205395171882054. https://doi.org/10.1177/2053951718820549

Senate Olympics Investigation. (2019). *The courage of survivors: A call to action*. United States Senate. https://www.moran.senate.gov/public/_cache/files/c/2/c232725e-b717-4ec8-913e-845ffe0837e6/FCC5DFDE2005A2EACF5A9A25FF76D538.2019.07.30-the-courage-of-survivors--a-call-to-action-olympics-investigation-report-final.pdf

Shelby, R. M. (2020). Techno-physical feminism: Anti-rape technology, gender, and corporeal surveillance. *Feminist Media Studies*, *20*(8), 1088–1109. https://doi.org/10.1080/14680777.2019.1662823

Sokoloff, N. J., & Dupont, I. (2005). Domestic Violence at the Intersections of Race, Class, and Gender: Challenges and Contributions to Understanding Violence Against Marginalized Women in Diverse Communities. *Violence Against Women*, *11*(1), 38–64. https://doi.org/10.1177/1077801204271476

Sombatruang, N., Sasse, M. A., & Baddeley, M. (2016). Why do people use unsecure public wi-fi?: An investigation of behaviour and factors driving decisions. *Proceedings of the 6th Workshop on Socio-Technical Aspects in Security and Trust*, 61–72. https://doi.org/10.1145/3046055.3046058

Sue, D. W. (2006). The invisible whiteness of being: Whiteness, white supremacy, white privilege, and racism. In M. G. Constantine & D. W. Sue (Eds.), *Addressing racism: Facilitating cultural competence in mental health and educational settings* (pp. 15–30). John Wiley & Sons Inc.

Taddei, S., & Contena, B. (2013). Privacy, trust and control: Which relationships with online self-disclosure? *Computers in Human Behavior*, *29*(3), 821–826. https://doi.org/10.1016/j.chb.2012.11.022

Tambe, A. (2018). Reckoning with the Silences of #MeToo. *Feminist Studies*, *44*(1), 197. https://doi.org/10.15767/feministstudies.44.1.0197

The White House. (2011). *Apps against abuse*. https://obamawhitehouse.archives.gov/1is2many/apps-against-abuse

Travers, A. (2003). Parallel Subaltern Feminist Counterpublics in Cyberspace. *Sociological Perspectives*, *46*(2), 223–237. https://doi.org/10.1525/sop.2003.46.2.223

Tuerkheimer, D. (2019). *Unofficial reporting in the #MeToo era* (Vol. 2). University of Chicago Legal Forum. https://chicagounbound.uchicago.edu/uclf/vol2019/iss1/10

United Nations Human Rights Council. (2018). *Accelerating efforts to eliminate violence against women and girls: Preventing and responding to violence against women and girls in digital contexts*. Human Rights Council. https://digitallibrary.un.org/record/1654650?ln=en

United Nations Women. (2019). *Using AI in accessing justice for survivors of violence*. https://www.unwomen.org/en/news/stories/2019/5/feature-using-ai-in-accessing-justice-for-survivors-of-violence

Vlasova, V. (2020, June 22). *Web skimming with Google Analytics*. SecureList by Kaspersky. https://securelist.com/web-skimming-with-google-analytics/97414/

White, D., & McMillan, L. (2020). Innovating the Problem Away? A Critical Study of Anti-Rape Technologies. *Violence Against Women*, *26*(10), 1120–1140. https://doi.org/10.1177/1077801219856115

Appendix

Tables

TABLE 1: Anti-violence reporting apps analysed

ANTI-VIOLENCE APP	WEBSITE	HOME COUNTRY	REPORTING INTERFACE	FUNCTIONALITY	CONCEPTUALISATION OF DATA PROTECTION
#NOTME	Not-Me.com	Canada, US	E-form	Intermediary: submits anonymous reports of harassment to employers	Anonymity
ALLVOICES	AllVoices.co	US	E-form	Intermediary: submits anonymous reports of harassment to employers	Confidentiality and/or anonymity
BOTLER AI	Botler.ai	Canada, US	AI chatbot	Pre-vets cases to determine if any laws have been broken	Confidentiality
CALLISTO	MyCallisto.org	US	E-form, information escrow	Survivor enters a record into the matching system—which holds the perpetrator's name in escrow until another user names the same perpetrator	Anonymity
HELLO CASS	HelloCass.com.au	Australia	Chatbot	Provides information about family and sexual violence, available counseling services, the legal system, and safety planning	Anonymity
JDOE	JDoe.io	US	E-form, information escrow	Survivor enters a record into the Matching system—which holds the perpetrator's name in escrow until another user names the same perpetrator	Anonymity
SAFESPORT	Safesport.i-sight.com/portal	US	E-form	Intermediary: submits anonymous reports of harassment to employers	Confidentiality and/or anonymity
TALK TO SPOT	TalkToSpot.com	US	AI chatbot	Intermediary: interviews users and submits anonymous reports of harassment to employers	Anonymity
VESTA SIT	Vestasit.com	Canada	E-form	Information and Reporting Platform	Anonymity

TABLE 2: Types of account data collected in anti-violence reporting apps

	ACCOUNT DATA COLLECTED										
ANTI-VIOLENCE APP	GENDER	RACE	SEXUAL ORIENTATION	AGE OR DOB	NAME	PHONE NUMBER	EMAIL ADDRESS	RESIDENTIAL LOCATION	EMPLOYER	EMPLOYMENT TITLE	NAME OF PERPETRATOR
#NOTME				x	x	x	x		x		x
ALLVOICES						x			x	x	x
BOTLER AI	X			x	x		x	x		x	
CALLISTO					x	x	x				x
HELLO CASS						x					
JDOE	X			x	x	x	x				

	ACCOUNT DATA COLLECTED										
ANTI-VIOLENCE APP	GENDER	RACE	SEXUAL ORIENTATION	AGE OR DOB	NAME	PHONE NUMBER	EMAIL ADDRESS	RESIDENTIAL LOCATION	EMPLOYER	EMPLOYMENT TITLE	NAME OF PERPETRATOR
SAFESPORT					x	x	x				x
TALK TO SPOT							x		x	x	
VESTA SIT											x

TABLE 3: Types of visitor data collected in anti-violence reporting apps

	VISITOR DATA COLLECTED				
ANTI-VIOLENCE APP	IP ADDRESS	DEVICE AND BROWSER INFORMATION	GEOGRAPHIC LOCATION	REFERRING WEBSITE	WEBSITE USAGE
#NOTME		x			
ALLVOICES	X	x	x	x	x
BOTLER AI	X	x	x		x
CALLISTO	X	x		x	x
HELLO CASS					
JDOE	X	x		x	
SAFESPORT	X	x	x	x	x
TALK TO SPOT	X	x	x	x	x
VESTA SIT	X	x	x	x	x

Published by

Alexander von Humboldt
INSTITUTE FOR INTERNET
AND SOCIETY

in cooperation with

UNIVERSITY OF TARTU
Johan Skytte Institute of
Political Studies