

Dear Reader,

Thank you for taking the time to read this working paper. Any feedback is very welcome and appreciated. Please do not cite without contacting the authors.

Rachel and Nóra

## **Annual Cambridge International Law Conference 2019**

### **New Technologies: New Challenges for Democracy and International Law**

Addressing Indirect Discrimination and Gender Stereotypes

in AI Virtual Personal Assistants:

The Role of International Human Rights Law

#### **ABSTRACT**

Virtual Personal Assistants are increasingly becoming a common aspect of everyday living. However, with female names, voices, and characters, these devices appear to reproduce harmful gender stereotypes about the role of women in society and the type of work women perform. Designed to “assist”, virtual personal assistants – such as Apple’s Siri and Amazon’s Alexa – reproduce and reify the idea that women are subordinate to men, and exist to be “used” by men. Despite their ubiquity, these aspects of their design have seen little critical attention in scholarship, and the potential legal responses to this issue have yet to be fully canvassed. Accordingly, this article sets out to critique the reproduction of negative gender stereotypes in virtual personal assistants and explores the provisions and findings within international women’s rights law to assess both how this constitutes indirect discrimination and possible remedies for redress. In this regard, this article explores the obligation to protect women from discrimination at the hands of private actors under the Convention on the Elimination of All Forms of Discrimination Against Women, the work of the Committee on Discrimination Against Women on gender stereotyping, the role of the United Nations Guiding Principles on Business and Human Rights, as well as domestic enforcement mechanisms for international human rights norms and standards.

Key words: gender stereotypes; indirect discrimination; AI; Virtual Personal Assistants; women’s rights; CEDAW; UN Guiding Principles; GDPR; data protection impact assessments.

## INTRODUCTION

Virtual Personal Assistants (VPAs) are increasingly becoming a common aspect of everyday living.<sup>1</sup> However, with female names, voices, and characters, these devices appear to reproduce harmful gender stereotypes about the role of women in society and the type of work women perform. Designed to “assist”, virtual personal assistants – such as Apple’s Siri and Amazon’s Alexa – the gendering of such technologies constitutes a broader societal harm insofar as they reproduce and reify negative gender stereotypes about the role of women as submissive and secondary to men, and consequently serve to perpetuate indirect discrimination against women. Within this context, we take gender stereotypes to refer broadly to a set of expectations and assumptions about the behaviour, nature and role of women. Yet, despite their ubiquity, these aspects of their design has seen little critical attention in scholarship, and the potential legal responses to this issue have yet to be fully canvassed. Accordingly, this article proceeds in three parts. First, we explore how VPAs reproduce gender stereotypes about the role of women in society, the work women perform, and the personality traits women are expected to display. The second part of this article turns to analyse how the international women’s rights canon can be drawn on as part of a broader agenda of legal responses to the issue.

In particular, we examine the obligation of state parties to prevent indirect discrimination at the hands of private actors, as set out under the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW) and the prescriptions of the Committee on the Elimination of Discrimination Against Women (the Committee) with respect to gender stereotypes and addressing discrimination against women in emerging technologies. This section additionally analyses the United Nations Guiding Principles on Business and Human Rights (UN Guiding Principles), including the obligation to conduct human rights impact assessments and the recommendations made in terms of data protection and new technologies. The third and last part of this paper explores relevant mechanisms for upholding international human rights norms and standards at a domestic level, including local regulatory and oversight functions and their role in securing, promoting and protecting human rights.

Our examination demonstrates the continued relevance of the international human rights law canon, and considers the ways in which the standards and mechanisms established at an

---

<sup>1</sup> John Danaher, ‘Toward an Ethics of AI Assistants: An Initial Framework’, (forthcoming) *Philosophy and Technology* <<https://philpapers.org/archive/DANTAE-2.pdf>> accessed 5 May 2019.

international level can be better enforced and integrated through local governance structures, including national ombudsmen.

## **2. Reproducing Gender Stereotypes in AI-VPAs**

In 2013, a film written, produced and directed by Spike Jonze was released, starring Scarlett Johansson as the lead female. Johansson's character features only as her voice in the film simply entitled *Her*. She plays an AI-driven VPA with whom the male lead (Joaquin Phoenix) falls in love. Her voice also caught the attention of Alex Acero, the lead for the speech team of Apple's Siri.<sup>2</sup> Inspired by *Her*, Acero sought to make Siri's voice more human-like, indeed a voice with which someone could fall in love.<sup>3</sup> To make Siri more human-like meant feeding its algorithms on huge swaths of data, training it to recognise its user's voice and provide a more personalised experience, and, of course, ensuring that Siri's voice is pitched at the optimal tone to induce feelings of affability and helpfulness.<sup>4</sup> All of this required - in following with *Her* and in order to ensure Apple's VPA was human-like – Siri to be represented and characterised by a female voice.

The year following the release of *Her*, Amazon released its VPA, Alexa. The use of VPAs in the home or on smart phones are now increasingly in popularity, with a recent survey finding 52% of their participants from the UK, US and Canada, using VPAs, such as Siri or Alexa.<sup>5</sup> Alexa is now available for your car,<sup>6</sup> such that, together with BMW's new VPA, you can 'never leave home without her' (Amazon Alexa Marketing).<sup>7</sup> Given the ubiquity of these technologies, and the likelihood that the use of VPAs in every aspect of our daily lives will only increase, the propensity of VPA companies to draw on gender stereotypes in the design and marketing of their products must be critically reviewed and addressed.

There are a number of ways in which the design, marketing, and indeed use of, VPAs draws on discriminatory and harmful gender stereotypes insofar as VPAs are presented as female. Their presentation as female largely occurs through female names (Siri means "the

---

<sup>2</sup> David Pierce, 'How Apple Finally Made Siri Sound More Human', 2017 WIRED  
<<https://www.wired.com/story/how-apple-finally-made-siri-sound-more-human/>> accessed 5 May 2019.

<sup>3</sup> Ibid.

<sup>4</sup> Pierce (n 2).

<sup>5</sup> Amanda Zantal-Wiener, 'The State of Voice: Looking Ahead to 2019' (2018) Hubspot  
<<https://blog.hubspot.com/news-trends/the-state-of-voice-2018-2019>> accessed 19 March 2019.

<sup>6</sup> Amazon Alexa Vehicles homepage, <<https://www.amazon.co.uk/b?ie=UTF8&node=16236511031>> accessed 20 May 2019.

<sup>7</sup> Ibid.

beautiful woman that leads you to victory” in Nordic<sup>8</sup>), the use of a default female voice, and the characterisation of VPAs as “assistants”. Indeed, given their presentation as women (and, as we have seen above Amazon prompts its customers to utilise the object pronoun ‘her’ in reference to Alexa), the association is that the work they perform is then female work and, given that that the intrinsic purpose of VPAs is to perform work that is less important such that the user is freed up to conduct the more important work,<sup>9</sup> this gendering of VPAs then makes a critical value statement about the cultural worth of “women’s work”. In all, it reproduces a harmful stereotype of the female assistant or secretary that, like Siri’s name, leads you (read: men) to victory, and is inferior to the man (or “user”) they serve.

In addition, VPAs are often marketed in ways that implicitly suggest that they take the place of an absent domestic female figure, typically a mother or wife.<sup>10</sup> Amazon’s 2018 advert for Alexa – entitled ‘Dad’s Day’ – sees a new dad at home with his baby, *just for the day*.<sup>11</sup> He is presented (uncritically) as unaccustomed to the role of primary caregiver to the baby, but Alexa is on hand to guide him through his day, and even to praise and comfort him when he does well. Apple’s 2017 advert for Siri and the Apple iPhone 7 stars Dwayne Johnson and is entitled ‘The Rock x Siri Dominate the Day’.<sup>12</sup> This advert depicts Dwayne Johnson accomplishing various astonishing tasks (painting the Sistine Chapel, launching a fashion line) during the course of a day, with the help of Siri to execute tasks such as reading emails or ordering a Lyft ride. Indeed, the design of VPAs as female is a critical part of their characterisation as helpful assistants who help you succeed, but who will not get in your way. They play into the stereotype of women as inferior and secondary to men and men’s work, thus reproducing discriminatory and historical gender norms about the place and value of women in society.

The stereotyped feminisation of VPAs is also apparent in their speech. In addition to the way in which they are marketed, the programmed speech of VPAs is a key component of their characterisation. While primarily VPAs, and particularly Siri and Alexa, have been

---

<sup>8</sup> Adam Cheyer, ‘How Did Siri Gets Its Name?’ (2012) Forbes <<https://www.forbes.com/sites/quora/2012/12/21/how-did-siri-get-its-name/#19fd7c57376b>> accessed 7 May 2019.

<sup>9</sup> Danaher (n 1) also engages in ethical questions around the cognitive capacities VPAs supposedly free up by performing less important work.

<sup>10</sup> R Adams, “‘Make Google Do It’: Interpellating Digital Object/Subjects of Desire in the Gendering of AI-driven Virtual Personal Assistants’ (2019) <<https://www.youtube.com/watch?v=0Q9mzoEXa64>> accessed 20 May 2019.

<sup>11</sup> Amazon, ‘Dad’s Day’ advert, <<https://www.ispot.tv/ad/d2jz/amazon-echo-dads-day>> accessed 20 May 2019.

<sup>12</sup> Apple, ‘The Rock x Siri Dominate the Day’, <<https://www.youtube.com/watch?v=2DWUuKII09Y>> accessed 20 May 2019.

designed to assist with everyday tasks, such as shopping (Alexa) and diary organisation (Siri), their designers were also keen to ensure that their natural language processing skills were sufficiently sophisticated to hold general conversations, and to respond to a wide variety of questions and comments. Indeed, Heather Zorn – Director of Alexa Engagement at Amazon – has stated that Alexa’s design team was keen to ensure that Alexa could respond to statements such as “‘Alexa, I love you’”.<sup>13</sup> (Alexa’s response: “that is a nice thing to say”; Siri’s response: ‘you are the wind beneath my wings’<sup>14</sup>). Yet, they have also been programmed to respond to more sordid language, including “you are a bitch!”.<sup>15</sup> To this Siri previously responded “I’d blush if I could”,<sup>16</sup> – it later became “I don’t know how to respond to that”<sup>17</sup> – and Alexa responds with ‘well, thanks for the feedback’.<sup>18</sup> It is through their programmed responses to questions, such as these, that are distinctly gendered, that Siri and Alexa reproduce male-orientated stereotypes of female behaviour, suggesting that male advances should be responded to – whether virtually or in the real world – with polite deflections, rather than calling this out as problematic and wrongful.

It is also through these programmed responses that Alexa and Siri, in particular, have been developed with what Woods describes as ‘stereotypical personas’.<sup>19</sup> Woods describes thus:

Alexa performs a form of digital domesticity with gender as a central facet. Whether managing a user’s calendar, modelling good etiquette for children, or providing (sexualized) companionship, these artefacts reveal how Alexa performs a gendered persona, leveraging her significant technological achievements towards traditionally feminine roles of mother and care-taker.

---

<sup>13</sup> Quoted in Heather Suzanne Woods, ‘Asking More of Siri and Alexa: Feminine Persona in Service of Surveillance Capitalism’, (2018) *Critical Studies in Media Communication*, 11. See also, James Risley, ‘One Year After Amazon Introduced Echo, Half A Million People Have Told Alexa, “I Love You”’, (GeekWire, 17 November 2015) <<https://www.geekwire.com/2015/one-year-after-amazon-introduced-echo-half-a-million-people-have-told-alexa-i-love-you/>> accessed 12 May 2019.

<sup>14</sup> Researcher own research with Siri device, April 2019.

<sup>15</sup> Nóra Ni Loideain and Rachel Adams, ‘From Ava to Siri and the GDPR’ (2018) King’s College London Dickson Poon School of Law Legal Studies Research Paper Series:

<[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3281807](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3281807)> accessed 18 April 2019, 5.

<sup>16</sup> For an analysis on Siri’s response to this question, see also Hillary Bergen, ‘I’d Blush if I Could’: Digital Assistants, Disembodied Cyborgs and the Problem of Gender’, (2016) *Word and Text: A Journal of Literary Studies and Linguistics* Vol VI, 95-113.

<sup>17</sup> Researcher own research with Siri device, May 2019.

<sup>18</sup> Ibid.

<sup>19</sup> Woods (n 13) 7.

Yet Alexa is not alone in this charge. Alexa's AI VA peer, Siri, is asked to perform a similar, stereotypical persona.<sup>20</sup>

While the digital domesticity of Siri and Alexa clearly both reproduces and reifies gendered stereotypes about the role of women, their personas are also grounded in traditional and stereotyped characteristics of femininity. In an essay on traditional gender stereotypes, Mary Kite, Kay Deaux and Elizabeth Haines describe how 'women [...] are viewed as more emotional, gentle, understanding, and devoted, whereas men are seen as more active, competitive, independent, and self-confident'.<sup>21</sup> Comparably, Woods notes how Alexa's character was developed to be 'smart, approachable, humble, enthusiastic, helpful and friendly'<sup>22</sup> – traits, as noted above, broadly aligned to stereotyped notions of women. Indeed, the use of the female voice was a critical tool through which the designers of VPAs sought to distance the personality of VPAs from more male-associated traits, such as independence and even aggression.<sup>23</sup>

By reproducing stereotyped ideas about the personality traits of women, the way in which women behave, and the work which women perform, is not just harmful in and of itself, but has material consequences for women and the expectations of women in society. As Naomi Ellemers describes, gender stereotypes 'reinforce perceived boundaries between women and men and seemingly justify the symbolic and social implications of gender for role differentiation and social inequality'.<sup>24</sup> One of the critical ways in which gender stereotypes have material effects on women's lives is in their careers and life choices, both in terms of the career options and abilities women believe they have, and the career options broader society believe are available to them.<sup>25</sup> Indeed, this has been what has led to the continued association of women with administrative work and thus the prevalent occupation by women of secretary jobs. In addition, Ellemers notes how 'gender stereotypes not only influence the perceived

---

<sup>20</sup> Woods (n 13) 7.

<sup>21</sup> Mary Kite, Kay Deaux and Elizabeth Haines, 'Gender Stereotypes', in Florence L. Denmark and Michele A. Paludi (eds), *Psychology of Women: A Handbook of Issues and Theories* (Second Edition, Praeger, London) 2008, 226

<sup>22</sup> Woods (n 13) 11.

<sup>23</sup> Ni Loideain and Adams (n 15).

<sup>24</sup> Naomi Ellemers, 'Gender Stereotypes', (2018) *Annual Review of Psychology* Vol 69, 275-298, 278

<sup>25</sup> Ibid.

potential of men and women when they are being selected for future careers, but also impact how the work actually performed by men and women is rated and valued'.<sup>26</sup>

What can also be added here, is that gender stereotypes can impact the very work that is produced. VPAs, which encompass stereotyped ideas about women, women's work and their roles in society, are the product of the IT and tech sector which is now increasingly well recognised to contain critical gender inequalities.<sup>27</sup> Research has shown that women occupy less than 20% of positions in the tech sector,<sup>28</sup> and further, that 'from 1980 to 2010, 88 percent of all information technology patents were by male-only invention teams'.<sup>29</sup> Thus, with design teams largely made up of men, it is inevitable that the products they design will reflect their biases, beliefs and value systems, including stereotyped ideas of women.

## INTERNATIONAL WOMEN'S RIGHTS LAW

The development and use of technologies based on normative ideas about women's behaviour, personality and roles in society, can serve to perpetuate discrimination against women. In September 1981 CEDAW came into force, with the overall objective of ending discrimination against women.<sup>30</sup> Since coming into force it has been ratified by 189 states parties. Article 2 of the Convention pertains to the core obligations of states parties with respect to the eradication of discrimination, including:

Article 2. States Parties condemn discrimination against women in all its forms, agree to pursue by all appropriate means and without delay a policy of eliminating discrimination against women and, to this end, undertake:

[...]

(e) To take all appropriate measures to eliminate discrimination against women by any person, organization or enterprise.<sup>31</sup>

---

<sup>26</sup> *Ellemers* (n 24) 279.

<sup>27</sup> S M West, Meredith Whittaker and Kate Crawford, 'Discriminating Systems: Gender, Race and Power in AI' (2019) AI Now Institute <<https://ainowinstitute.org/discriminatingsystems.html>> accessed 21 May 2019.

<sup>28</sup> Shubhomita Bose, 'Only 20% of Tech Jobs are Held by Women' (26 December 2018 Smallbiztrends) <<https://smallbiztrends.com/2018/03/women-in-technology-statistics.html>> accessed 20 May 2019.

<sup>29</sup> Kasee Bailey, 'The State of Women in Tech 2019' (7 March 2019 Dreamhost) <<https://www.dreamhost.com/blog/state-of-women-in-tech/>> accessed 20 May 2019.

<sup>30</sup> International Convention on the Elimination of All Forms of Discrimination Against Women (adopted 18 December 1979, entered into force 3 September 1981) A/RES/34/180 (CEDAW).

<sup>31</sup> *Ibid*, Article 2.

The eradication of discrimination based on negative stereotypes of women is addressed in Article 5 of the Convention, which requires states parties to take appropriate measures to:

5. (a) [...] modify the social and cultural patterns of conduct of men and women, with a view to achieving the elimination of prejudices and customary and all other practices which are based on the idea of the inferiority or the superiority of either of the sexes or on stereotyped roles for men and women.<sup>32</sup>

Gender stereotyping is also noted in Article 10 on equal rights to education. Article 10 requires:

10. States Parties shall take all appropriate measures to eliminate discrimination against women in order to ensure to them equal rights with men in the field of education and in particular to ensure, on a basis of equality of men and women:

[...]

- (c) The elimination of any stereotyped concept of the roles of men and women at all levels and in all forms of education by encouraging coeducation and other types of education which will help to achieve this aim and, in particular, by the revision of textbooks and school programmes and the adaptation of teaching methods.<sup>33</sup>

While these articles of the Convention place a clear obligation on states parties to address negative gender stereotyping, and recognises how gender stereotyping occurs within the education system, the Committee on Discrimination Against Women (the Committee), established under the Convention, continued to emphasise the importance of addressing gender stereotypes as key to eradicating discrimination, in all forms, against women. In 2004, the Committee issued General Recommendation 25 under CEDAW, which provided further guidance on Article 4 of the Convention, relating to temporary special measures states parties may take in order to address discrimination against women.<sup>34</sup> Thwarting ‘the persistence of gender-based stereotypes that affect women not only through individual acts by individuals but

---

<sup>32</sup> CEDAW (n 30) Article 5.

<sup>33</sup> CEDAW (n 30) Article 10.

<sup>34</sup> UN Committee for the Elimination of All Forms of Discrimination against Women, ‘General Recommendation No 25 on article 4, paragraph 1, of the Convention on the Elimination of All Forms of Discrimination against Women, on temporary special measures’ (2004) U.N. Doc. HRI/GEN/1/Rev.7 at 282.



also in law, and legal and societal structures and institutions’,<sup>35</sup> is listed as the third key obligation of states parties under the Convention. The General Recommendation proceeds to include addressing ‘stereotypical attitudes and behaviour’<sup>36</sup> as one of the critical ‘temporary special measures’ states parties must take, under Article 4 of the Convention, in order to accelerate de facto or substance gender equality:

Temporary special measures should be adopted to accelerate the modification and elimination of cultural practices and stereotypical attitudes and behaviour that discriminate against or are disadvantageous for women.<sup>37</sup>

Another key aspect of the General Recommendation relating to gender stereotypes is its definition of indirect discrimination. The full and detailed definition is as follows:

Indirect discrimination against women may occur when laws, policies and programmes are based on seemingly gender-neutral criteria which in their actual effect have a detrimental impact on women. Gender-neutral laws, policies and programmes unintentionally may perpetuate the consequences of past discrimination. They may be inadvertently modelled on male lifestyles and thus fail to take into account aspects of women’s life experiences which may differ from those of men. These differences may exist because of stereotypical expectations, attitudes and behaviour directed towards women which are based on the biological differences between women and men. They may also exist because of the generally existing subordination of women by men.<sup>38</sup>

In short, within this definition is the understanding that gender stereotyping does not only result in discrimination against women, but also that it is a result *of* such discrimination. This definition of indirect discrimination came before General Recommendation 28 under CEDAW, issued in 2010, which provided further details for states parties on their obligations under Article 2 of the Convention.<sup>39</sup>

---

<sup>35</sup> Ibid, para 7.

<sup>36</sup> *General Recommendation 25* (n 34) para 7.

<sup>37</sup> *General Recommendation 25* (n 34) para 38.

<sup>38</sup> *General Recommendation 25* (n 34) Note 1.

<sup>39</sup> UN Committee for the Elimination of All Forms of Discrimination against Women, ‘General Recommendation No. 28 On the Core Obligations of States Parties under Article 2 of the Convention on the Elimination of All Forms of Discrimination against Women’, (2010), U.N. Doc. CEDAW/C/GC/28.

General Recommendation 28 now stands as the Committee's official position on indirect discrimination. Here, indirect discrimination is defined as occurring,

When a law, policy, programme or practice appears to be neutral in so far as it relates to men and women, but has a discriminatory effect in practice on women because pre-existing inequalities are not addressed by the apparently neutral measure. Moreover, indirect discrimination can exacerbate existing inequalities owing to a failure to recognize structural and historical patterns of discrimination and unequal power relationships between men and women.<sup>40</sup>

However, it further details states parties' duty, under Article 2 (e), to protect women from discrimination at the hands of a third party as:

The obligation to protect requires that States parties protect women from discrimination by private actors and takes steps directly aimed at eliminating customary and all other practices that prejudice and perpetuate the notion of inferiority or superiority of either of the sexes, and of stereotyped roles for men and women.<sup>41</sup>

General Recommendation 28 further notes that the principle of equality between the sexes includes the freedom to 'develop their personal abilities, pursue their professional careers and make choices without the limitations set by stereotypes, rigid gender roles and prejudices'.<sup>42</sup>

What CEDAW therefore fundamentally provides is an expansive right of non-discrimination, supported by the wholesale condemnation of all forms of discrimination against women, to which all states parties commit. Thus, while CEDAW sets out that discrimination refers to both direct and indirect discrimination resulting in the unequal treatment of women, General Recommendation No. 28 on the core obligations of States parties under article 2 (non-

---

<sup>40</sup> Ibid, para 16.

<sup>41</sup> *General Recommendation 28* (n 39) para 9.

<sup>42</sup> *General Recommendation 28* (n 39) para 22.

discrimination) of CEDAW notes that ‘the Convention anticipates the emergence of new forms of discrimination that had not been identified at the time of its drafting’.<sup>43</sup> To this end, and of relevance to the discussion here on the gendering of VPA technologies, States parties are recommended to take measures to identify and address emerging forms of discrimination, and indeed the development of new stereotypes based on assumptive and harmful notions of gender roles.

In 2010, a case was taken to the Committee which relied, amongst others, on the provisions of CEDAW outlined above relating to gender stereotypes, brought by Karen Tayag Vertido.<sup>44</sup> The case concerned the judgement of rape case in the Philippines that relied on gender stereotypes in acquitting the accused. The case was taken to the Committee under the Optional Protocol to CEDAW, to which the Philippines is a party.<sup>45</sup> These stereotypes related to the role played by victims in situations of rape, and included that a rape victim is timid and tries every opportunity to escape, that rape takes place only when the victim is directly threatened, and that rape only takes place at the hands of strangers.<sup>46</sup> While this case does not pertain to indirect discrimination, it is useful for assessing how the Committee views the right not to be subject to harmful gender stereotypes as a judiciable right, and found that the state party (the Philippines) had failed in its due diligence duty under Articles 2 and 5 of CEDAW to ‘banish gender stereotypes’, and that the author of the communication (Vertido) was subject to revictimization through the stereotypical assumptions she endured through her rape trial.<sup>47</sup>

In 2013, the Office for the High Commissioner for Human Rights (OHCHR) commissioned a report on gender stereotyping as a human right. The report noted that ‘much more work is needed to prioritise stereotypes and stereotyping as a human rights concern and, in this, there is an important and significant leadership role for the Office of the High Commissioner for Human Rights (OHCHR) to play’.<sup>48</sup> The OHCHR is the leading human rights office at the UN, and the treaty based monitoring bodies – including the CEDAW Committee – fall under its mandate. The elevation of gender stereotyping to an issue of concern for the

---

<sup>43</sup> *General Recommendation 28* (n 39) para 8.

<sup>44</sup> Committee on the Elimination of All Forms of Discrimination Against Women, Communication No 18/2008 authored by Karen Tayag Vertido (decision adopted 16 July 2010).

<sup>45</sup> Optional Protocol to the Convention on the Elimination of All Forms of Discrimination Against Women (adopted 6 October 1999, entered into force 22 December 2000) A/RES/54/4.

<sup>46</sup> *Communication No 18/2000* (n 44) para 3.5.

<sup>47</sup> *Communication No 18/2000* (n 44) para 8.4 and 8.8. For a further analysis of this case see also Simone Cusack and Alexandra S. H. Timmer, ‘Gender Stereotyping in Rape Cases: The CEDAW Committee’s Decision in Vertido v The Philippines’, (2011) 11(2) Human Rights Law Review, 329–342.

<sup>48</sup> Office of the High Commissioner for Human Rights Commissioned Report, ‘Gender Stereotyping as a Human Rights Violation’ (October 2013), 2.

OHCHR is therefore telling, demonstrating that it is a priority concern across the various UN human rights treaty bodies (including the Convention on the Rights of the Child, and the Convention on the Rights of Persons with Disabilities). The OHCHR report on gender stereotypes defines stereotypes thus:

A gender stereotype is a generalised view or preconception about attributes, or characteristics that are or ought to be possessed by women and men or the roles that are or should be performed by men and women.

It further notes that ‘there is a growing consensus that gender stereotyping poses a significant, yet largely unaddressed, challenge to the recognition, exercise and enjoyment of women’s human rights’, and that gender stereotyping is harmful and discriminatory ‘when it limits women’s or men’s capacity to develop their personal abilities, pursue their professional careers and make choices about their lives and life plans’. It notes too that ‘both hostile/negative or seemingly benign stereotypes can be harmful’.<sup>49</sup>

The provisions of CEDAW provide a broad canvass on which to assess the indirect discrimination at play in the use of gender stereotypes in the design and marketing of VPAs. The obligations placed on states parties under CEDAW require that states parties undertake due diligence in protecting women from discrimination – both direct and indirect – at the hands of private actors, such as Apple and Amazon. It should not be a caveat for such companies that the products they produce are only of commercial value, and thus their consumers *chose* to buy into the gender stereotyping they implicitly offer, but instead it must be recognised that VPAs constitute a ubiquitous form of AI technology, and that their presence is increasing across various digital platforms, including at banks (Royal Bank of Scotland has rolled out at VPA called Cora at its branches), in cars (noted above), and in workplaces.

Indeed, the UN Guiding Principles provides further guidance to private actors on their human rights responsibilities, and specifically notes that private actors must take into account

---

<sup>49</sup> Office of the High Commissioner for Human Rights, ‘Gender Stereotypes and Stereotyping and Women’s Rights’, (September 2014)  
<[https://www.ohchr.org/documents/issues/women/wrgs/onepagers/gender\\_stereotyping.pdf](https://www.ohchr.org/documents/issues/women/wrgs/onepagers/gender_stereotyping.pdf)> accessed 21 May 2019.

the potential human rights impact of their products and activities.<sup>50</sup> The UN Guiding Principles sought to address the gap in international law pertaining to the role of private actors by developing a framework for both States parties and private companies to promote respect for human rights. The UN Guiding Principles were unanimously endorsed by the Human Rights Council in 2011, and established the first global standard relating to the role of states and private companies with respect to business and human rights, set around three basic pillars: 1) The state duty to protect against the violation of human rights by private actors; 2) the corporate responsibility to respect human rights; and 3) the duty of both states and private companies to provide access to remedies where human rights occur. While the OHCHR issued a report on the right to privacy in the digital age, which calls for state parties to take the necessary steps to mitigate against the negative effects of surveillance and Big Data technologies and which has broad relevance for our concerns here with regard to VPAs,<sup>51</sup> the UN Guiding Principles themselves are silent on issues of women and the role of businesses in perpetuating discrimination against women. An additional concern for the UN guiding Principles is that, to date, they have not been widely implemented in practice, likely because they constitute a soft law mechanism in international law, meaning that their provisions are not binding on state parties or private companies.

Despite these setbacks, the UN Guiding Principles still provides important tools that, if properly embedded, can serve as useful mechanisms through which to address adverse gender stereotyping and discrimination, such as may be discernible in the gendering of AI-driven VPAs. Such mechanisms include Principle 18 which provides that ‘business enterprises should identify and assess any actual or potential adverse human rights impacts with which they may be involved either through their own activities or as a result of their business relationships’. Principle 18 further delineates that:

While processes for assessing human rights impacts can be incorporated within other processes such as risk assessments or environmental and social impact assessments, they should include all internationally recognized human rights as

---

<sup>50</sup> United Nations Human Rights Council, ‘Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework’ (adopted 16 June 2011) HR/PUB/11/04.

<sup>51</sup> Office of the High Commissioner for Human Rights, ‘The Right to Privacy in the Digital Age’ (3 September 2018), A/HRC/39/29.

a reference point, since enterprises may potentially impact virtually any of these rights.<sup>52</sup>

In this way, while the UN Guiding Principles do not directly address issues such as gender discrimination, it is important to note that their implementation must encompass a holistic reading of international human rights law, which includes the right to non-discrimination set out in CEDAW and its interpretation.

While the OHCHR, CEDAW and the work of the Committee has demonstrated a clear prioritisation toward addressing negative gender stereotyping and indirect discrimination, and the UN Guiding Principles provide further guidance to non-state actors on their human rights responsibilities, the broad scope of application of these provisions does not directly bear on some of the more complex issues associated with emerging technologies. The Beijing Declaration and Platform for Action by the UN at the Fourth World Conference on Women in 1995 in fact goes somewhat further than the aforementioned mechanisms in articulating some of the gender concerns relating to emerging technologies.<sup>53</sup> This document represented a global commitment to women's empowerment and gender equality. Amongst an extensive range of principles for achieving substantive gender equality and eradicating discrimination were a number of principles which effectively modernised those set out almost twenty years previously in CEDAW. This includes Article 35 which states:

With advances in computer technology and satellite and cable television, global access to information continues to increase and expand, creating new opportunities for the participation of women in communications and the mass media and for the dissemination of information about women. On the other hand, the global communication networks have been used to spread stereotyped and demeaning images of women for narrow commercial and consumerist purposes.<sup>54</sup>

---

<sup>52</sup> Ibid, Principle 18.

<sup>53</sup> United Nations, 'Beijing Declaration and Platform for Action by the UN at the Fourth World Conference on Women in 1995' <<https://www.refworld.org/docid/3dde04324.html>> accessed 21 May 2019.

<sup>54</sup> Ibid, Article 35.

In response to this concern – which is arguably reiterated in the gendering of AI-driven VPAs – Article 35 calls on the media to promote non-stereotyped portrayals of men and women, and for the equal participation of both sexes in communications and the media.

The idea of promoting women’s participation as a key measure for States parties to take to advance women’s empowerment and gender equality is evident in many of the international and regional agreements aimed at addressing discrimination, or bias (for example, Sustainable Development Goal 5). Another key response aimed at advancements in technology has been to call for science and technology to respond to women’s needs. This was set out at the 55<sup>th</sup> Session of the UN Commission for the Status of Women which took as its priority theme ‘Access and participation of women and girls in education, training and science and technology, including for the promotion of women’s equal access to full employment and decent work’.<sup>55</sup> As part of the agreed conclusions to the 55<sup>th</sup> Session, States parties were called upon to:

Develop gender-sensitive curricula for educational programmes at all levels and take concrete measures to ensure that educational materials portray women and men, youth, girls and boys in positive and non-stereotypical roles, particularly in the teaching of scientific and technological subjects, in order to address the root causes of segregation in working life.<sup>56</sup>

Putting this commitment into action may involve measures such as requiring gender training within large tech companies, and ensuring gender education forms part of the curriculum for subject such as computer science at universities. Addressing the root causes of issues such as the reification of negative gender stereotypes in AI-driven VPAs, as above, is a critical touchstone for promoting AI for good and the ethical use of AI.

In addition, the agreements at the 55<sup>th</sup> Session of the UN Commission on the Status of Women further required that states parties:

---

<sup>55</sup> UN Commission for the Status of Women, ‘Agreed conclusions on access and participation of women and girls in education, training and science and technology, including for the promotion of women’s equal access to full employment and decent work’ (n.d.)  
<[https://www.un.org/womenwatch/daw/csw/csw55/agreed\\_conclusions/AC\\_CSW55\\_E.pdf](https://www.un.org/womenwatch/daw/csw/csw55/agreed_conclusions/AC_CSW55_E.pdf)> accessed 20 May 2019.

<sup>56</sup> Ibid, para v.

Encourage the integration of a gender perspective in the science and technology curricula throughout all stages of education and continuous learning, and the use of gender-based analysis and gender impact assessments in research and development in science and technology, and promote a user driven approach to technology development in order to increase the relevance and usefulness of advancements in science and technology for both women and men.<sup>57</sup>

This kind of gender based analysis and undertaking of gender impact assessments is, we posit, perhaps the most fundamental tool for addressing some of the gender concerns, as highlighted above, in VPAs.

While these provisions are notable and have direct relevance for thinking about how to address the discerned societal harm of gendering AI-driven VPAs, and particularly for encouraging the development of technology in a way that is responsive to women's needs, the critical concern for international human rights law is how to hold the private sector to account for the reproduction of negative gender stereotypes and the social harm this causes in terms of indirect discrimination against women. Within international human rights law states parties are the primary actors responsible for holding private companies to account. Further, given the non-binding nature of the international human rights law set out above, it becomes imperative to read such provisions alongside other international law provisions which are binding, including national laws where international human rights law commitments have been domesticated and put into effect.

## LEVERAGING LOCAL OVERSIGHT AND MONITORING BODIES

Crucial to the adequacy and effectiveness of any legal framework are its oversight and enforcement mechanisms. Within Europe, the importance of having independent, well-funded, and adequately resourced regulators whose enforcement powers are sufficiently robust is recognised under the EU General Data Protection Regulation (GDPR).<sup>58</sup> Since this major

---

<sup>57</sup> *UN Commission for the Status of Women* (n 55) para rr.

<sup>58</sup> 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 [2016] OJ L119/1 (GDPR).



update to EU data protection law came into force in May 2018, Data Protection Authorities (DPAs) across the EU now have equal powers of enforcement under the GDPR enabling them to issue fines of up to 20 million euro or 4% of a major tech company's worldwide annual turnover.<sup>59</sup> Such robust enforcement powers have contributed to the GDPR being described as a 'game changer' and the 'gold standard' for similar laws in other jurisdictions worldwide.<sup>60</sup> In other words, all the high level protections provided for in international laws are toothless unless their enforcement is possible at the national level. DPAs are tasked with the enforcement of the GDPR in each EU Member State and in less than a year since its entry into force these domestic authorities have issued significant sanctions against major tech companies. For instance, the French DPA (CNIL) recently fined Google 50 million euro for an ongoing and serious failure to meet consent requirements for their targeted advertising of users.<sup>61</sup>

As highlighted above, the UN Guiding Principles have not been widely implemented in practice, including notably the requirement by private companies to undertake 'human rights impact assessments' that identify harms that arise from their activities, including indirect indiscriminate. We would posit that one of the key factors that continues to entrench the lack of compliance with these provisions, particularly Principle 18, by state parties and private companies is the absence of any adequate and effective enforcement regime. There is, however, a clear analogy to be drawn here between the lack of enforcement of the UN Guiding Principles and the patchy enforcement of data protection law across the EU for the past two decades. Thus, the European Commission's assessments of the EU data protection legal identified a number of key challenges to be addressed in the reform process that culminated in what is now the GDPR. These findings showed a clear need to remedy the inconsistent compliance with data protection law across EU Member States and a lack of enforcement capacity due to DPAs at the domestic level being insufficiently resourced and lacking independence in their operation and functions.<sup>62</sup> In order to address these key challenges towards ensuring a more effective enforcement regime in future, the GDPR has introduced a number of mechanisms that speak to the governance challenges facing the UN Guiding Principles. For instance, EU Member

---

<sup>59</sup> *GDPR* (n 58) Article 83.

<sup>60</sup> Giovanni Buttarelli, 'The EU GDPR as a clarion call for a new global digital gold standard' (2016) 6(2) *International Data Privacy Law* 77-78.

<sup>61</sup> CNIL, 'The CNIL's restricted committee imposes a financial penalty of 50 million euro against Google LLC' < <https://www.cnil.fr/en/cnils-restricted-committee-imposes-financial-penalty-50-million-euros-against-google-llc> > accessed 21 May 2019.

<sup>62</sup> European Commission, 'First Report on the Implementation of Data Protection Directive (95/46/EC)' COM(2003) 265, 24 February 2004, 18; European Commission, 'Safeguarding Privacy in a Connected World: A European Data Protection Framework for the 21<sup>st</sup> Century' COM(2012) 9 final, 25 January 2012, 7.

States are now legally required to provide DPAs with ‘complete independence’ and the human, technical, and financial resources necessary to ensure the effective exercise of their powers, including their monitoring and enforcement duties.<sup>63</sup>

Moreover, Article 35 of the GDPR provides<sup>64</sup> that where a type of processing ‘*in particular using new technologies*’ is ‘likely to result in a high risk’ to the rights and freedoms of natural persons, the controller (private companies)<sup>65</sup> shall, ‘*prior to the processing*, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data.’<sup>66</sup> Although earlier EU data protection law placed a general obligation on private companies to notify DPAs of any processing operations likely to present ‘specific risks’ to the rights of data subjects posed by their processing of personal data *prior* to this processing taking place<sup>67</sup>, the EU legislature later recognised that this provision suffered from ‘very patchy compliance’ by controllers to notify DPAs in general and in practice made little, if any, impact on ‘improving the protection of personal data’.<sup>68</sup> Hence, DPIAs are one of the new *ex ante* governance provisions under the GDPR which has the normative (and admittedly ambitious) aim of being an actionable mechanism that ensures accountable compliance by translating the legal standards of EU data protection law into reality. In the early drafting stages of the GDPR, the European Commission envisaged DPIAs as a means to strengthen the individual’s right to data protection by enhancing the accountability of those organisations involved in ‘risky processing’ in order to identify these risks in advance, foresee problems, and bring forward solutions.<sup>69</sup>

---

<sup>63</sup> *GDPR* (n 58) Articles 51-59, Recital 120.

<sup>64</sup> *GDPR* (n 58) Article 35(1) states: ‘Where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data’.

<sup>65</sup> *GDPR* (n 58) Article 4(7) defines a controller as: ‘the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law’.

<sup>66</sup> Emphasis added.

<sup>67</sup> European Parliament and Council Directive 95/46/EC of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31, Article 20.

<sup>68</sup> *GDPR* (n 58) Recital 89.

<sup>69</sup> EC Impact Assessment: Commission Staff Working Paper ‘Impact Assessment Accompanying the document Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) and Directive of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data by competent authorities for the purposes of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and the free movement of such data’, SEC(2012) 72 final, 25 January 2012, i, 43.

As a co-regulatory measure, the DPIA aims to achieve this by entrenching its principles and safeguards through influencing the technical and organisational culture of individuals and institutions.<sup>70</sup> As other commentators have rightly observed<sup>71</sup>, however, this sharing of responsibility raises some legitimate questions regarding the capacity, and in some cases willingness, of industry to both adequately anticipate and address the intangible harms that fall within the scope of data protection law as part of a DPIA. Accordingly, the capturing of hearts and minds, so to speak, will be a key test of the GDPR's effectiveness in practice if DPIAs are to actually lead data controllers to considering in advance the implications of their data processing and design choices. Otherwise, what DPAs intend to be a substantive and ongoing process undertaken by industry may fail to be meaningfully implemented and amount in practice to no more than a reactive 'paper checklist' or 'box-ticking exercise'.<sup>72</sup>

The GDPR also aims to tackle this key challenge of co-regulation through its new *ex post* enforcement and oversight mechanisms. First, the EU legislature seeks to achieve effective implementation of the GDPR through a 'more defined risk-based approach' towards the governance of EU data protection law<sup>73</sup> and the role of controllers vis-à-vis the accountability principle. Secondly, the deterrent power of the new significant fines introduced under the GDPR that businesses may be subject to should they fail to either undertake, or subsequently implement, the requirements they committed to in their DPIAs will be the other more coercive approach used to also entrench a more GDPR-compliant culture. The new accountability principle under the GDPR provides that data controllers, such as manufacturers who determine the design of the hardware and software of VPAs, and how these devices operate and process personal data, are responsible for complying with the GDPR's key data protection principles and must also demonstrate this compliance.<sup>74</sup> It is worth noting here that the GDPR applies to any data processing targeted at providing any goods or services to individuals within the EU,

---

<sup>70</sup> Bronwen Morgan and Karen Yeung (eds), *An Introduction to Law and Regulation* (Cambridge University Press, 2007) 10.

<sup>71</sup> Raphaël Gellert, 'Data protection: a risk regulation?' (2015) 5(1) *International Data Privacy Law* 3, 15-16; Damian Clifford and Jef Ausloos, 'Data Protection and the Role of Fairness' (2018) *Yearbook of European Law* 1, 54.

<sup>72</sup> Roger Clarke, 'Privacy Impact Assessment: It's Origin and Development' (2009) 25(2) *Computer Law and Security Review* 123, 124; Bert-Jaap Koops, 'The trouble with European data protection law' (2014) 4(4) *International Data Privacy Law* 250, 255; Reuben Binns, 'Data protection impact assessments: a meta-regulatory approach' (2017) 7(1) *International Data Privacy Law* 22, 26.

<sup>73</sup> Gellert (n 71); Orla Lynskey, *The Foundations of EU Data Protection Law* (Oxford University Press, 2015) 82-84.

<sup>74</sup> *GDPR* (n 58) Article 5(2).

such as the collection of data and profiling of users behaviour from gendered AI VPAs by companies based in the U.S.<sup>75</sup>

The effectiveness of both these approaches will largely turn on the effectiveness of the guidance, supervision, and (now rather extensive under the GDPR) enforcement powers exercised by DPAs<sup>76</sup> and the European Data Protection Board (EDPB).<sup>77</sup> The latter replaces the Article 29 Working Party and will play an important role in the future development of the GDPR through its guidelines, recommendations, and best practice documents that are aimed at harmonising the pan-EU approach by DPAs to oversight and enforcement. Hence, it is welcome that the GDPR places a clear obligation on EU Member States to ensure that these bodies – regarded as the ‘guardians’ of EU data protection law by the Court of Justice of the EU<sup>78</sup> – are adequately equipped with the resources necessary both to provide effective supervision and guidelines to controllers, and, if required, the capacity to effectively exercise their enforcement powers.<sup>79</sup>

We would posit that there are several useful lessons outlined from the recent reform of EU data protection law that may guide how current gaps in implementation and enforcement of the UN Guiding Principles should be addressed. First, it is argued that any local enforcement bodies tasked with monitoring the UN Guiding Principles should be given similar powers and (critically) resources to enhance the implementation and enforcement as provided to DPAs under the GDPR. Secondly, there is greater scope for more consistent implementation and consistency of the UN Guiding Principles in the area of human rights impact assessments if the local bodies responsible for its enforcement also co-operated with their respective national DPAs. Much could be learned from this collaboration by both institutions with respect to addressing how to better ensure compliance with human rights obligations in this specific area of emerging technology by private companies. Finally, there is also scope for the provisions of the UN Guiding Principles, namely Principle 18, to inform the DPIA process being undertaken by private companies with respect to the issue of indirect discrimination and the social harm to women posed by AI VPAs.

---

<sup>75</sup> *GDPR* (n 58) Article 3(2).

<sup>76</sup> *GDPR* (n 58) Article 58.

<sup>77</sup> *GDPR* (n 58) Article 68.

<sup>78</sup> C-518/07, *European Commission v. Federal Republic of Germany*, EU:C:2010:125, para 23 (GC); C-614/10, *European Commission v. Republic of Austria*, EU:C:2012:631, para 52 (GC); C-288/12, *European Commission v. Hungary* EU:C:2014:237, para 53 (GC).

<sup>79</sup> *GDPR* (n 58) Article 52(4).

## CONCLUSIONS

Addressing gender stereotypes and the resultant indirect discrimination toward women of VPAs requires wholesale reform within the industry out of which these devices are produced. While addressing the lack of diversity in the tech industry is a fundamental step in the right direction, companies must also recognize and act upon their broader societal obligations, including their human rights obligations to ensure that violations do not occur within their products or value chains. The recently published report from AI Now on ‘Discriminating Systems: Gender, Race and Power in AI’ includes within their recommendations two pertinent steps that companies should take:

Rigorous testing should be required across the lifecycle of AI systems in sensitive domains. Pre-release trials, independent auditing, and ongoing monitoring are necessary to test for bias, discrimination, and other harms.

The field of research on bias and fairness needs to go beyond technical debiasing to include a wider social analysis of how AI is used in context. This necessitates including a wider range of disciplinary expertise.

This article has sought to respond to the call for interdisciplinary approaches to addressing discrimination and bias in AI systems by drawing on both critical gender and legal analysis methods. In addition, we have discussed how impact assessments, undertaken prior to the launch of AI products provides a crucial mechanism for enforcing international human rights norms and standards, and addressing the negative social impact of AI technologies. Lastly, while the efforts of companies to conduct human rights due diligence is central to addressing the concerns we have highlighted with respect to gender stereotyping and indirect discrimination against women in AI VPAs, it is important that states – as the primary duty holder with respect to human rights – uphold their human rights obligations and monitor the protection, promotion and respect for human rights by private actors too.