

## Feedback to the Consultation on the *White Paper on Artificial Intelligence (AI)*

### Focusing on **AI & gender: An EU law perspective**

by Prof. Dr. Gloria González Fuster<sup>1</sup>

1. **Gender** is a recurrent theme in discussions around the regulatory framework for Artificial Intelligence (AI), as a variety of problematic gender-related issues have surfaced in relation to AI information systems. Many examples of **gender-based discrimination** have notably reached the media in the past years.<sup>2</sup> **Gender bias** has been decried, for instance, in facial recognition systems deployed in different contexts.<sup>3</sup> AI systems and technologies supporting AI, like many other contemporary technologies, have also been criticised for conveying and normalising **gender stereotyping**.<sup>4</sup> The collection and widespread **processing of data about gender** has made visible particularly acute problems for transgender communities.<sup>5</sup> Identified problems have sometimes been directly connected to the **lack of gender diversity** in the industries behind AI,<sup>6</sup> as well as more generally among AI researchers. Issues related to bias have been framed as rooted in the **design of algorithms**,<sup>7</sup> but also in the **training data**s used for developing the systems, and more generally connected to their deployment and use.<sup>9</sup>
2. **Taking into account gender** in discussions around AI regulation requires acknowledging the potential **disparate impact of AI** on individuals on the basis of their gender, and foreseeing steps to **prevent** such disparate impact. This necessary also obliges to take into account how gender intersects with other factors determining such disparate impact. There is widespread knowledge, in this sense, on how poverty<sup>10</sup> and race and ethnic factors can aggravate the differentiated impact of practices based on the automated processing of personal data.<sup>11</sup> In addition, taking into account gender also requires devoting attention to **how gender is defined**

---

<sup>1</sup> Gloria González Fuster is a Research Professor at the Faculty of Law & Criminology of the Vrije Universiteit Brussel (VUB), and Co-Director of VUB's Law, Science, Technology & Society (LSTS) Research Group. She is an academic member, appointed in her personal capacity, at the European Commission's Multistakeholder expert group to support the application of Regulation (EU) 2016/679 (E03537).

<sup>2</sup> See, for instance: Rachel Metz (2016), '[Why Microsoft Accidentally Unleashed a Neo-Nazi Sexbot](#)', *MIT Technology Review*, 24 March 2016; Guadalupe González (2018), '[How Amazon Accidentally Invented a Sexist Hiring Algorithm](#)', *Inc*, 10 October 2018; Karen Hao (2019), '[Facebook's ad-serving algorithm discriminates by gender and race](#)', *MIT Technology Review*, 5 April 2019; Jonathan Cohn (2019), '[Google's algorithms discriminate against women and people of colour](#)', *The Conversation*, 24 April 2019.

<sup>3</sup> See, for instance: Fight for the Future, '[Backlash forces UCLA to abandon plans for facial recognition surveillance on campus](#)', *Medium*, 19 February 2020.

<sup>4</sup> UNESCO (2019), *I'd blush if I could: Closing gender divides in digital skills through education*.

<sup>5</sup> See notably, on this issue and other a variety of other issues on gender and contemporary data practices: Anja Kovacs (2020), '[When our bodies become data, where does that leave us?](#)', *Medium*, 28 May 2020.

<sup>6</sup> Itself a problem with a long history, cf. for instance; Marie Hicks (2017), *Programmed inequality: How Britain discarded women technologists and lost its edge in computing*. MIT Press.

<sup>7</sup> See, for instance: Sara Wachter-Boettcher (2017), *Technically Wrong: Sexist apps, biased algorithms, and other threats of toxic tech*, W. W. Norton & Company.

<sup>8</sup> Generally, on data and bias, see: Caroline Criado Perez, *Invisible Women: Exposing data bias in a world designed for men*. Random House, 2019.

<sup>9</sup> See, for instance: Cathy O'Neil, *Weapons of Math Destruction: How big data increases inequality and threatens democracy*, Crown, p. 113 and ff.

<sup>10</sup> See, for instance: John Gilliom (2001), *Overseers of the Poor: Surveillance, Resistance, and the Limits of Privacy*, University of Chicago Press; Michele E. Gilman (2008), 'Welfare, Privacy, and Feminism', *University of Baltimore Law Forum*, Vol. 39, Number 1, Article 4.

<sup>11</sup> See, for instance: Safiya Umoja Noble (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York University Press.

**and constructed through AI-systems**, as we are nowadays all being increasingly automatically classified by algorithmic systems, sometimes on unclear grounds, and without our knowledge.

3. The significance and variety of gender issues at stake in AI imply that they cannot be addressed by a single type of responses. The problems identified have multiple roots and manifestations, requiring **different types of interventions**, some of which might need to be legislative.
4. The European Commission (EC)'s **White Paper on AI** explicitly recognises **the importance of gender for the future of AI regulation**. The White Paper's very opening paragraph mentions, in this sense, that **gender-based discrimination** is one of the potential risks triggered by AI.<sup>12</sup> It also references two academic papers illustrating that AI can display **gender bias**.<sup>13</sup> The most specific intervention that the White Paper puts on the table, in relation to gender, are possible '*obligations to use data sets that are sufficiently representative, especially to ensure that all relevant dimensions of gender, ethnicity and other possible grounds of prohibited discrimination are appropriately reflected in those data sets*'<sup>14</sup> – that is, requirements to tackle specifically the problem of diversity in data, including training data.
5. The White Paper provides three main pointers as to how the EC is mapping more broadly existing problems and possible solutions to the gender-related challenges of AI, concretely: the EU Gender Equality Strategy 2020-2025; an Opinion by the European Commission's Advisory Committee on Equal Opportunities for Women and Men; and a report of the European Network of Equality Bodies (Equinet). An analysis of these documents reveals that the **EC policy approach to gender-related issues connected to the AI** is at the moment **particularly imprecise**, not to say superficial:
  - a. Regarding gender and AI, the **EU Gender Equality Strategy 2020-2025** (COM(2020) 152 final) basically states that, in relation to AI in general, the EC has published a White Paper on AI setting out the European approach to AI '*grounded in EU values and fundamental rights, including non-discrimination and gender equality*'. Thus, as the White Paper refers to the EU Gender Equality Strategy for issues on gender and AI, which just refers back to the White Paper regarding AI matters, what is generated is a full circle of circular cross-references without real substance; and
  - b. Regarding the Opinion on AI of the **European Commission's Advisory Committee on Equal Opportunities for Women and Men**,<sup>15</sup> and the report of the **European Network of Equality Bodies** (Equinet),<sup>§</sup> they both coincide in granting special attention to the work of the **High-Level Expert Group on AI** as a relevant source for guidance on these matters, failing however to fully acknowledge that the work of the High-Level Expert Group on AI does not deal with the legal component of 'trustworthy AI'.<sup>16</sup> Thus, what is precipitated here are a series of cross-references that move the reflection on AI and gender from the *White Paper on AI* to 'specialised' bodies with more knowledge on gender, which however refer to a 'specialised' body working on AI, which decided to exclude the analysis of EU law from its reflections, thus marking

---

<sup>12</sup> COM(2020) 65 final, p. 1.

<sup>13</sup> Ibid., p. 11.

<sup>14</sup> Ibid., p. 19.

<sup>15</sup> Advisory Committee on Equal Opportunities for Women and Men (2020), *Opinion on Artificial Intelligence – opportunities and challenges for gender equality*, 18 March 2020.

<sup>§</sup> Robin Allen QC and Dee Masters (2020), *Regulating AI: the new role for Equality Bodies: Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence*, An Equinet Publication.

<sup>16</sup> High-Level Expert Group on Artificial Intelligence (AI HLEG) (2019), *Ethics Guidelines for Trustworthy AI*, p. 2.

another wave of cross-references which basically move the reflections on AI, regulation and gender **into the realm of the ‘ethical’ dimension of AI**, cancelling out a detailed examination of the legal challenges at the concrete crossroads of gender, AI and EU law.

6. In this sense, it is crucial to note that the concrete recommendations emanating from the **High-Level Expert Group on AI** regarding gender have been particularly **limited**, focusing notably on tackling the gender gap in digital skills by improving education<sup>17</sup> - which is certainly an important objective, but cannot be the only response to the many and significant gender-related issues triggered by AI, with serious impacts on EU fundamental rights.
7. Whereas the importance of an active reflection on technology (including AI), gender and human rights is **already recognised in other fora**.<sup>18</sup> EU institutions have until now been lagging behind in this field.
8. This contribution aims at stressing that future EU policies on AI should take gender-related issues seriously, on the basis of the **specific gaps that EU policy and secondary law present in light of EU fundamental rights**. It focuses on discussing two major problems: first, that existing EU data protection law instruments, including the General Data Protection Regulation (GDPR), do not offer by themselves sufficiently precise protection against the processing of personal data leading to **algorithmic gender-based discrimination** (Part A); and, second, that despite the significant contribution of the GDPR to increased transparency around data processing practices, more efforts are necessary to ensure that individuals effectively know when, and how, AI-systems are classifying them into a specific gender (Part B). The contribution concludes with short recommendations.

#### **A. EU data protection law does not counter algorithmic gender-based discrimination**

9. From the perspective of EU law, there is a **major difference between gender-based algorithmic discrimination and other recurrent types of algorithmic discrimination**. While the processing of data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, and trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation are regulated as ‘processing of special categories of personal data’ in EU data protection law instruments, **the processing of data concerning gender is not granted such special protection**.
10. This means that, in EU data protection law, the processing of gender-related data is **denied the special protection** that generally aims at reducing the processing of data connected to factors that might be used for discriminatory purposes,<sup>19</sup> and also left out of the mechanisms that aim at limiting the possibility of **automated decisions**, including **profiling**, which have a significant effect on individual's lives and are based on ‘special’ factors.<sup>20</sup>

---

<sup>17</sup> High-Level Expert Group on Artificial Intelligence (AI HLEG) (2019), *Policy and investment recommendations for trustworthy Artificial Intelligence*, pp. 32-35.

<sup>18</sup> The Council of Europe has, for instance, put in place a Gender Equality Rapporteur in the context of its Ad Hoc Committee on Artificial Intelligence (CAHAI). See also, in this sense, for instance: [Annex 2: The Human Right to Privacy: A Gender Perspective](#) (Gender issues arising in the digital era and their impacts on women, men and individuals of diverse sexual orientations gender identities, gender expressions and sex characteristics’ – Report of Consultation by the SRP Thematic Taskforce ‘Privacy and Personality’), illustrating work carried out by the United Nations (UN) Special Rapporteur on the Right to Privacy in relation to gender, covering also AI.

<sup>19</sup> As established, for instance, by Art. 9 of the GDPR.

<sup>20</sup> As established, for instance, by Art. 22(4) of the GDPR.

11. Thus, even though Article 21 of the EU Charter of Fundamental Rights prohibits discrimination on the basis on race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age and sexual orientation, but also on the basis of ‘sex’ in the sense of **gender**, EU data protection law leaves without special protection the processing of gender related data, one of the factors that is most often used by AI systems to discriminate.
12. This fact needs to be fully acknowledged and tackled by the EU legislator. Indeed, although arguably EU primary and secondary law, including the GDPR, might be regarded as not compatible in their spirit with any unfair data processing based on gender, the truth is that in practice they do not offer **concrete adequate tools** to mitigate the risks attached to gender-based discrimination in our society, considering its widespread nature, and its clear tension with the EU fundamental rights and values. In the absence of such specific protection, it is crucial that **other mechanisms** are put in place, possibly together with a reinforcement and clarification of the **transparency obligations imposed on data controllers and processors as regards to the processing about gender**.
13. An example of the **practical implications** of the legal gap described can be found in the upcoming **European Travel Information and Authorisation System (ETIAS)**,<sup>21</sup> a system based on the profiling of individuals. To reduce possible discriminatory data practices, ETIAS provides a series of rules aiming at reducing the use of risk indicators connected to certain categories of data, such as data revealing a person’s colour, race, ethnic or social origin, political or any other opinion, religion or philosophical belief, trade union membership, membership of a national minority, disability, or sexual orientation, but not gender (together with age). Ultimately, the system leaves open the door to the creation and automated use of ‘risk indicators’ such as ‘*women older than...*’, or ‘*men younger than...*’, connecting different impacts to a classification under these profiles, in a border management context in which data subject rights (and most notably transparency) are additionally particularly fragmented and restricted – making it particularly difficult, if not impossible, for individuals to become aware of how they have been profiled on the basis of their gender, and the implications of such profiling.

## **B. Transparency must effectively encompass the data subject’s right to know which gender they have been attributed by AI-systems, and why**

14. Connected to the previous problem, but nevertheless different, emerges the specific issue of gender classification, and in particular automated gender attribution. In this sense, it needs to be recalled that the **GDPR** constituted a major step in **increasing the transparency of data processing practices**.<sup>22</sup> It not only established the obligation for data controllers and processors to provide detailed information on any personal data processing taking place, but also the obligation to provide such information in a way that meets formal transparency requirements (it must be clear, concise, intelligible), and accompanied on usable information on data subjects’ rights, which shall facilitate the use by data subjects of the right data protection rights.
15. Despite such reinforcement of transparency, many **grey areas remain** in relation to the effective **contribution of the GDPR to algorithmic transparency**. In this sense, there has been much discussion in the literature about what might be called ‘**algorithmic data subject**

---

<sup>21</sup> Regulation (EU) 2018/1240 of the European Parliament and of the Council of 12 September 2018 establishing a European Travel Information and Authorisation System (ETIAS) and amending Regulations (EU) No 1077/2011, (EU) No 515/2014, (EU) 2016/399, (EU) 2016/1624 and (EU) 2017/226, 20J L 236, 19.9.2018, pp. 1–71.

<sup>22</sup> Other EU data protection law instruments present similar provisions, albeit none offers higher standards than the GDPR. For the sake of brevity, the analysis is limited here to the GDPR.

**rights**', that is, the bundle of rights in the hands of data subjects in relation to algorithmic processing of personal data relating to them. These algorithmic data subject rights cover for instance certain entitlements derived from Article 22 of the GDPR in relation to some instances of automated decision-making, but also more generally standard data subject rights (right to be informed, right of access, right to rectify, etc.) insofar as they apply to profiling and automated decision-making. The **deficiencies concerning algorithmic data subject rights**, and most notably their **enforcement**, are particularly visible nowadays in relation to automated gender attribution.

16. **Automated gender attribution**<sup>23</sup> might be described as the automated classification of individuals under a gender label. It can take a variety of shapes. It is most often driven by a commercial logic, as gender is regarded as a basic demographic of great marketing value, also in an online context. The pervasiveness of **online automated gender attribution** is significant, and its impact potentially serious.<sup>24</sup> In many of the cases in which data controllers and processors process data about our gender it is not because we have been given the possibility to express or 'select' our gender, but because our gender (or a gender attributed to us, which might not be the right one) was inferred from other pieces of data related to us. In some cases, data controllers appear to actually classify individuals under certain gender categories irrespective of the gender actively endorsed and communicated by individuals.<sup>25</sup>
17. To give two examples, companies such as **Amazon** and **Twitter** automatically classify individuals in ways that do not appear to be fully transparent to them. Amazon, in this sense, attributes gender on the basis of the surname, which might be a relatively straightforward and in principle predictable operation for some surnames, but is less unproblematic for others – in any case, the algorithm used in such operations is not publicly known. Twitter classifies individuals on the basis of their online activity, which opens the question of how is decided which activities correspond to which gender, who decides that, and how can such assumptions, potentially plagued by gender stereotyping, be contested.<sup>26</sup> The gender labelling is notably used by Twitter to create groups of persons allowing advertisers to 'promote content' to them on the basis of their gender.
18. **Automated gender attribution** can also occur in different settings, for instance connected to systems based on facial recognition. There exists already a myriad of technological solutions on the market that claim to be able to 'detect' the gender of individuals.<sup>27</sup> Automated gender attribution is, **in principle, covered by EU data protection law**, at least to the extent that the processing of personal data takes place. It could be, however, that in certain scenarios those processing the data are not in a position to link the data to identified or identifiable individuals, leading to a non-qualification of data as personal data, and the absence of direct applicability of EU data protection law.
19. When EU data protection applies, it is crucial to acknowledge that, in spite of the presumed ambitions of the GDPR and possibly also in tension with the core requirements of Article 8 of

---

<sup>23</sup> Related technologies are sometimes referred to with the term 'Automated Gender Recognition (AGR)'. Embracing such terminology implies however that gender might indeed be re-cognised externally, also automatically. Because the connections between gender, gender identity, gender expression and interpretation of presumed gender markers cannot be accepted without further problematisation, the more accurate term 'automated gender attribution' shall be privileged.

<sup>24</sup> See, in this sense: Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham (2018), 'Gender recognition or gender reductionism? The social implications of embedded gender recognition systems', *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*.

<sup>25</sup> Cf. Rena Bivens (2017), 'The gender binary will not be deprogrammed: Ten years of coding gender on Facebook', *New Media & Society*, 19.6, pp. 880-898.

<sup>26</sup> For a more detailed review of these cases, cf. my Invited Tutorial at FAT\* 2020: 'Gender: What the GDPR does not tell us (But maybe you can?)'; microsite [here](#); video of tutorial [here](#).

<sup>27</sup> On this, see for instance: Os Keyes (2019), '[The Body Instrumental](#)', *Logic: Nature*, Issue 9.

the EU Charter of Fundamental Rights, there is today a ubiquitous **lack of transparency** around these practices. This lack of transparency directly affects the awareness of individuals as to their very existence (the very fact that they are being classified under a certain gender, and possibly being misgendered), but affects also the individuals' ability to question and contest decisions.

20. The contribution of **Data Protection Authorities (DPAs)** to advancing in these matters has been until now **particularly limited**, despite notable exceptions.<sup>28</sup> In this sense, for instance, the European Data Protection Board (EDPB) endorsed Guidelines on automated-decision making and profiling, prepared by the Article 29 Working Party,<sup>29</sup> that fail to throw enough useful any light of the problem at stake. Most worryingly, the EDPB supports through such Guidelines the idea that classifying individuals according to their gender '*without making any predictions or drawing any conclusion about an individual*' does not constitute profiling,<sup>30</sup> a statement that is highly debatable in itself but, critically, is fundamentally misleading, as it does not take into account that currently most of the time **gender attribution itself constitutes profiling**, in the sense that it is not a mere collection of data provided by the data subject but rather indeed the automated use of other personal data to evaluate, analyse or predict certain personal aspects relating to a person (that is, their gender).<sup>31</sup>
21. A proper and enforceable regulatory framework for automated gender attribution is **a general necessity**, as **all individuals** shall have the right to know which gender they have been automatically attributed by private companies and public authorities, but also why (on the basis of which data, and applying which logic), and individuals shall have the right to have such attribution rectified if appropriate. These rights are however even more necessary for the persons being classified under a gender that eventually might be used in the context of **gender-based discrimination**, and negatively affect them.
22. For **transgender** individuals, keeping track of gender-related information about them available to others is of **major importance**.<sup>32</sup> Not only it is fundamental for them to be attributed the right gender: when a gender change has occurred, it can be critical not to make available in certain contexts information about this very fact, as this could lead to them being involuntarily 'outed' in a way that is still linked in many circles to **stigmatisation**, but also potentially to significant life repercussions, as well as **violence**.<sup>33</sup> In this context, so-called 'gender recognition' technologies, especially to the extent that they pretend to perceive gender as such, shall be regarded as **high-risk**.
23. Freedoms and rights regarding gender identity and gender expression are part of the **international human rights** framework to which the EU is committed. In this context, it must be the most advanced international instruments in the field recognise important relevant principles notably in relation to **gender recognition, including legal recognition**, but also

---

<sup>28</sup> Worth mentioning contributions from DPAs to gender-related issues, although not specifically related to AI, include the launch of a 'priority channel' of assistance by the Agencia Española de Protección de Datos (AEPD) in case of online dissemination of sexually-related content (cf. [here](#)), and the commitment of the State Data Protection Commissioner of Land Schleswig-Holstein to feminist data protection (cf. [here](#)).

<sup>29</sup> Article 29 Working Party, *Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, Adopted on 3 October 2017, as last Revised and Adopted on 6 February 2018 WP251rev.01.

<sup>30</sup> Ibid., p. 7.

<sup>31</sup> A detailed overview of the limitations of DPAs affecting the consideration of gender-related issues among DPAs cannot be provided here; a useful analysis, with a global scope, can be found in: Elizabeth Coombs and Kara McKee (2019) '[The 'missing women' in data protection reporting](#)', *IAPP*, 13 June 2019.

<sup>32</sup> On the importance of gender information management and some of the related challenges for transgender individuals, see for instance: Ali Shahaab et. al, 'Managing Gender Change Information on Immutable Blockchain in Context of GDPR', *The JBBA*, Volume 3, Issue 1, 2020, [https://doi.org/10.31585/jbba-3-1-\(3\)2020](https://doi.org/10.31585/jbba-3-1-(3)2020).

<sup>33</sup> On the persistent challenges faced by transgendered people in the EU, see, for instance: EU Fundamental Rights Agency (FRA), *A long way to go for LGBTI equality*, 2020.



regarding the **enjoyment of human rights in relation to communication and information technologies**.<sup>34</sup>

### C. Concluding remarks

24. In light of the above, the European Commission should in any future AI policy:

- a. Strengthen its commitment to gender equality and **the protection of all fundamental rights of individuals regardless of their gender**, making sure that policy measures around gender and AI do not merely lead to vague allusions to ethical considerations, but are fully and firmly anchored in a commitment to EU fundamental rights for all, and deliver appropriate legal safeguards.
- b. Acknowledge that **EU data protection law currently does not *per se* provide a satisfactory answer to the gender-related challenges of AI**, most notably in terms of gender-based discrimination, and transparency of gender profiling. Existing data protection instruments such as the GDPR might need to be **supported by specific guidance**, but also accompanied by new legislative measures, dealing for instance also with AI-enabled gender attribution in the cases in which data protection laws might not directly apply.
- c. Promote **gender mainstreaming and awareness of gender-related issues among DPAs**, and first and foremost the EDPB and the European Data Protection Supervisor (EDPS) as advisor to the EC in data protection matters, as well as among any other possible authorities concerned in the future with the enforcement of AI regulation.
- d. Devote special attention to **automated gender attribution**, which is a widespread practice that affects the entire population, potentially sustaining gender stereotyping and enabling gender-based discrimination, but also as a **high-risk practice** that directly and dramatically touches upon the fundamental rights of **transgender and gender-diverse individuals** – as it can affect their right to privacy and data protection, their right to non-discrimination, but also other rights, such as their right to life.

Brussels, 14 June 2020  
Gloria González Fuster

---

<sup>34</sup> See notably the Yogyakarta Principles ‘Plus 10’, adopted on 10 November 2017, Geneva, in particular principles 31 and 36.