# Feminist perspectives to artificial intelligence: Comparing the policy frames of the European Union and Spain

Ariana Guevara-Gómez, Lucía O. de Zárate-Alcarazo and J. Ignacio Criado*
*Department of Political Science and International Relations, Universidad Autónoma de Madrid, Madrid, Spain*

**Abstract.** Artificial Intelligence (AI) is a disruptive technology that has gained interest among scholars, politicians, public servants, and citizens. In the debates on its advantages and risks, issues related to gender have arisen. In some cases, AI approaches depict a tool to promote gender equality, and in others, a contribution to perpetuating discrimination and biases. We develop a theoretical and analytical framework, combining the literature on technological frames and gender theory to better understand the gender perspective of the nature, strategy, and use of AI in two institutional contexts. Our research question is: *What are the assumptions, expectations and knowledge of the European Union institutions and Spanish government on AI regarding gender?* Methodologically, we conducted a document analysis of 23 official documents about AI issued by the European Union (EU) and Spain to understand how they frame the gender perspective in their discourses. According to our analysis, despite both the EU and Spain have developed gender-sensitive AI policy frames, doubts remain about the definitions of key terms and the practical implementation of their discourses.

Keywords: Feminism, gender theory, artificial intelligence, public policy, European Union, technological frames, public administration, Spain

**Key points for practitioners:**
  – Gender perspective is included as explicit and implicit references about the nature, the strategy and the use of AI in European and Spanish official documents.
  – Explicit references to gender perspective in AI strategies are important to build on strategies and policies for an inclusive and just AI.
  – Governments are encouraged to develop specific regulations and plans to apply the gender perspective to AI initiatives.
  – Future scholarly studies need to incorporate gender perspectives to gauge the meanings of implementing AI in our societies.

## 1. Introduction

Artificial Intelligence (AI) has emerged in the public sector as a groundbreaking technology. Due to the potentialities and capabilities associated with this disruptive technology, important changes are expected to take place in governments and public administrations all around the globe (Margetts & Dorobantu, 2019). Among the benefits and advantages of implementing AI in the public sector we can

*Corresponding author: J. Ignacio Criado, Department of Political Science and International Relations, Universidad Autónoma de Madrid, Madrid, Spain. E-mail: ignacio.criado@uam.es.

count, among others: optimization and personalization of public services, better and fairer decision-making processes, improvements during public policy implementation and evaluation processes, etc. (Kuziemski & Misuraca, 2020; Valle-Cruz et al., 2020). However, the use of AI in the public sector also arises both technical and non-technical problems. Algorithmic opacity, unfair bias, outcomes leading to discrimination, lack of explicability of algorithms, data quality, etc. (Mittelstadt et al., 2016) are some of the most important problems public policymakers and technologists still need to solve in order to make sure that AI implementation in the public sector is ethical and benefits us all.

If these problems are not properly solved, AI implementation will increase injustice, unfairness and inequality, as well as discrimination towards already excluded members of society due to gender, race, religion and sexual orientation reasons (European Commission, 2020). AI risks leading to gender bias, gender discrimination, gender exclusion, increase of inequality, gender digital divide, have been sufficiently proven in empirical studies and analysis about digital assistants (Costa & Ribas, 2019), word embeddings in search engines (Bolukbasi et al., 2016), the use of algorithms for public policy purposes (Peña & Varon, 2020). Thus, acknowledgement from public institutions and governments is the first step to unravel them. This could eventually lead to the construction and embodiment of an equalitarian, fair and inclusive framework, that is, to a feminist approach to AI policies that promotes and builds on the fundamental rights and values of the European Union (EU) (European Commission, 2000).

Thus, departing from a feminist socio-cognitive approach, the aim of this paper is to examine and analyze the policy framework that the EU has designed for the implementation of AI. At the same time, we investigate to what extent Spain has followed the European guidelines, and which one of them, if any, has settled more adequate values, strategies and policies to reach gender equality. We find the case of Spain especially noteworthy for two key reasons: First, due to its capacity of influence on AI implementation in Latin America, where AI is increasingly gaining momentum. Second, due to its fairly good position (8th) in the European Gender Equality Index (European Institute for Gender Equality, 2020). This gives the Spanish government the chance to remark its commitment to gender equality through the implementation of gender-sensitive public policies. To achieve this goal, we have classified assumptions, expectations and implicit knowledge on AI regarding gender in different dimensions that have been settled for the study of technology.

The remainder of the paper is divided in the following sections. In Section 2 we present the theoretical framework of our study. That is an intersectional approach to technology from socio-cognitive and feminist theories. Section 3 covers the analytical framework in which we present the three technological dimensions that we will be developed in this work and the codification we have used to identify references to gender equality. In Section 4 the results of our study are displayed. Then, in Section 5 we discuss the main findings from the previous section. Finally, we outline the conclusions and limitations of our work, and the future lines of study.

## 2. Theoretical framework

Socio-cognitive and behavioral approaches to public policies, administrations and governmental issues are on trend. Departing from theories of bounded rationality (Simon, 1957; Tversky & Kahneman, 1981; Simon, 1997; Gigerenzer, & Gaissmaier, 2011), these approaches state that people act on the bases of their interpretative schemes or frames about the world (Giddens, 1984; Orlikowski & Gash, 1994; Eden, 1992). By frame we mean "*definitions of organizational reality that serve as vehicles for understanding and action*" (Gioia, 1986, p. 50) that an individual holds in his or her mind. As frames give us a holistic knowledge about the world, they also include technology. Frames on technology are

usually known as technological frames, that is, "*subset of members organizational frames that concern the assumptions, expectations, and knowledge they use to understand technology in organizations*" (Orlikowski & Gash, 1994, p. 178) and depending on their content, people will act in a certain way or another. Thus, understanding people's and organization's technological frames is a key aspect to be able to anticipate the direction of their actions.

According to Orlikowski's socio-cognitive approach to technology, technological frames have three different dimensions. These dimensions are: (1) Nature of technology, (2) Technology strategy, and (3) Use of technology (Orlikowski & Gash, 1994). Each dimension includes a different subset of assumptions, expectations and knowledge about what the technology is, why is it used and how is it used. However, it is important to note that people from the same culture are more likely to share large portions of their frames as these are strongly based on cultural and social norms and traditions (Simon, 1993). The same thing happens in regard to technological frames. Thus, technological frames shared by European citizens and institutions will only shed light on the assumptions, expectations and knowledge on technology underlying European cultural norms.

In the socio-cognitive approach to technology, there is an element that should be taken into account: the shared understanding of gender roles in a society. Following Ridgeway (2009, p. 145), we assume that "*gender is a primary cultural frame for organizing social relations*", which means that social and institutional structures are influenced by individual conceptions of gender roles and identities. Thus, it is expected that gender conceptions and stereotypes would permeate the assumptions, expectations and knowledge on technology.

Here, the co-construction of gender and technology, inspired by social constructivism and, specifically, by sociotechnical perspectives of technology (Lagesen, 2015; Faulkner, 2001; Berg & Lie, 1995; Wajcman, 2009), unfolds two perspectives: gender in technology and the gender of technology (Faulkner, 2001). The first concept refers to the mutual influence of gender and technology, meaning that the design of technology reflects and, at the same time, reinforces gender roles and stereotypes (Faulkner, 2001). Some scholars have demonstrated empirically this argument, especially for the case of AI technologies. For example, facial recognition algorithms show more inaccuracies in the identification of black women, leading to conclude that there are gender and race biases in the design of these technologies (Buolamwini & Gebru, 2018; UNESCO, 2020). In fact, some big tech companies have been involved in the design of algorithms with the potential to support discriminatory practices (Harwell & Dou, 2020). Also, the analysis of different AI applications (such as digital assistants, word embedding in search engines, and digital platforms) reveals the reproduction and perpetuation of gender stereotypes through the design of algorithms (Costa & Ribas, 2019; Bolukbasi et al., 2016; Peña & Varon, 2020). In these cases, the training of algorithms with non-representative data contributes to perpetuate discrimination (Ntoutsi et al., 2019).

In some cases, researchers and practitioners have stated the importance of having women representation within the design teams and institutions to reduce gender stereotypes in technologies (Avila et al., 2018; Collett & Dillon, 2019; Leavy, 2018; Criado et al., 2021), while others state that this claim is not uncontested and should be analyzed with care (Sørensen, 1992; Faulkner, 2000). Despite we do not have consistent evidence, conceptions about gender roles and gender equality, despite the sex of the individual, will have indeed an influence in the interpretation of data, the design of algorithms, and the relationship between humans and machines (Lagesen, 2015; Ferrando, 2014; Gil-Juárez et al., 2018), meaning that technology is not gender-neutral (Nass, 1997; Gil-Juárez et al., 2018).

The second concept, gender of technology, refers to the gender associations in the use of artifacts. This encompasses assuming they have feminine or masculine characteristics according to the gender roles

defined in a society (Faulkner, 2001). This dimension is related to the previous one, as "*features designed into artifacts tailored specifically for women or men users tend to reflect and reinforce gender stereotypes, which in turn, play in to design choices*" (Faulkner, 2001, p. 84). This means that designer's choices influence the use of certain artifacts: i.e., reproductive technologies specially designed for women, or devices to facilitate household chores created with women in mind and based on stereotypes perpetuated in society (Faulkner, 2001). Certain technological devices could also be considered as symbols of masculinity, but it would depend on the specific contexts (Lie, 1995).

Moreover, AI poses new questions: do artifacts have gender? Examples of this idea could be found in designs such as voice assistants and chatbots. The majority of voice assistants are defined as females, not only through their voices, names and avatars, but also through the background stories developed by the companies to support these characters (West et al., 2019). Most chatbots are also classified as female (Feine et al., 2019). For both types, there is evidence suggesting that gender stereotypes are mutually shaped by the relationship between the AI artifact and the user: the answers given by the machine could reinforce the stereotypes and prejudices that already exist (West et al., 2019; McDonell & Baxter, 2019). Thus, developers should also take this dimension into account to address gender issues in AI. Considering these dimensions, gender in technology and gender of technology, we assume that both technological and gender frames will influence political and technical decisions on the use of technologies, and specifically AI, in public sector.

## 3. Analytical framework

In this section, we introduce the analytical framework of the study. We first state our research question that is based on the literature on socio-cognitive approaches to technology and gender theory: *What are the assumptions, expectations and knowledge of the European Union institutions and Spanish government on AI regarding gender?* In order to respond to our research question, we will build on Orlikowski's theory on technological frames, according to the three dimensions acknowledged in her work: the nature, the strategy, and the use of technology. To identify each dimension, we might answer "what is AI?" (nature of AI), "why is AI used?" (strategy of AI) and "how is AI used?" (use of AI), according to the EU institutions and the Spanish government. At the same time, we should identify the gender perspectives in these dimensions. This approach is complementary to the perspective on framing AI policies recently developed by Ulnicane et al. (2020).

Following a recent report of UNESCO (2020) and research on ethical principles in AI strategies (Jobin et al., 2019; Fjeld at al., 2020), we include explicit and implicit references to gender in AI, and we build our analytical framework combining these elements with Orlikowski's dimensions of technological frames (see Table 1). We assume that gender issues could arise in the discourses on AI, whether in explicit forms (i.e., using the words 'women', 'gender' or 'feminist') or implicit constructs (i.e., using words such as 'justice', 'fairness', 'equity', 'diversity' or 'non-discrimination'). At the same time, these expressions would be related to the nature, the strategies, and the use of AI.

In general terms, the *explicit references* to gender are related to a binary approach of this concept: in this case, we assume that gender refers to the categories 'women' and 'men', socially shaped and culturally defined (Martínez-Bascuñán, 2015). As defined by the United Nations Population Fund (UNFPA) (2005), "*gender refers to the economic, social and cultural attributes and opportunities associated with being male or female*". We are aware that this approach is contested, and as Butler (2007) explains, gender is performative, which means it could be transformed and take different forms beyond the limits of heterosexuality. In a first attempt to understand the interplay of gender and AI conceptions, we consider

Table 1
Analytical dimensions of gender and technological frames in AI

| Dimensions of technological frames | Gender dimensions in technology | Explicit references to gender | Implicit references to gender |
| --- | --- | --- | --- |
| Nature of AI<br>What is AI? Definitions and approaches. Attributes or Traits issues. | – Gender in technology<br>– Gender of technology | – AI with a gender perspective<br>– Feminist AI | – AI is designed by humans<br>– Co-construction of AI<br>– Fair, just and inclusive AI<br>– Human-centric AI |
| Strategies of AI<br>Why is the organization adopting AI? What are the motivations and objectives to do it? Strategic issues. | – Gender in technology<br>– Gender of technology | – AI to increase gender equality and balance (equality between women and men) in society<br>– Gender equality as a criterion to assess the success of AI<br>– AI to promote women empowerment | – AI to promote: justice, fairness, inclusion, equity, equality, diversity, empowerment, representativeness, shared benefits, non-bias and non-discrimination |
| Use of AI<br>How will AI be used? What are the projects and specific plans for implementation? Practical issues. | – Gender in technology | – Reduction of gender digital divide<br>– Training and education to increase women's use of AI<br>– Increase the number of women in AI design teams<br>– Increase the number of women in public institutions responsible for the implementation of AI<br>– Include the gender perspective in the design of algorithms<br>– Include gender representative data in the design of algorithms | – Reduction of digital divide<br>– Training and education to increase the individuals' use of AI, with special attention to inclusion, non-discrimination and fairness.<br>– Training and education of public servants for a fair use of AI<br>– Promote a fair and equal access to AI<br>– Use of representative and quality data in the design of AI<br>– Promote individuals and public servants' empowerment in the use of AI |
| | – Gender of technology | – Avoiding gender stereotypes in the characterization of AI<br>– Promoting gender equality through the characterization of AI | – Avoiding stereotypes in the characterization of AI<br>– Promoting equality, fairness, inclusion, non-discrimination through the characterization of AI |

Sources: Own elaboration based on Jobin et al. (2019), Fjeld at al. (2020), Orlikowski & Gash (1994), UNESCO (2020).

adequate to start following the mainstream focus in feminist theory (Dietz, 2003), acknowledging its limitations and potentialities.

In this line, we understand the term *gender equality* in its wider sense: as a matter of redistribution of resources and as a political issue of recognition and representation (Squires, 2007). Linked to this concept is the term *gender equity*, defined as "*the process of being fair to women and men*", which means to develop "*strategies and measures (. . .) to compensate for women's historical and social disadvantages*" (UNFPA, 2005). As UNFPA states, gender equity would lead to gender equality. Thus, we assume that when a discourse includes the terms *gender equality and balance, and gender equity* it could be related to the promotion of women's access to AI or to the inclusion of women in the design and implementation of policies on AI. In association with these concepts is the *digital gender divide*, related to the girls and women's access to the Internet and new technologies. In this case, we are not just thinking of the physical

access to digital devices, Internet and, specifically, AI, but also to the ability and skills for using these tools and producing outcomes from them (Sáinz et al., 2020; Lutz, 2019; Bimber, 2000).

Regarding the *implicit references* to gender, we assume that at least eleven terms are related to the gender perspective in AI. First, the term *equity* refers to the idea that sometimes it is desirable to treat people differently to fulfill their specific needs and achieve justice (Doyal, 2000). In this case, as explained by Doyal (2000), some have argued that women and men should receive different treatments, specifically in health issues. On the other hand, "*the treatment of individuals is inequitable if it is capricious or relates to 'irrelevant' characteristics. Commonly cited characteristics of this sort include race, religion, and gender*" (Culyer, 2001, p. 276). Despite the different perspectives on the relevance or irrelevance of gender issues in these debates, it should be noted that the term equity is commonly associated with gender. Second, the association between the term *equality* and gender is well established: according to the liberalism perspective, being a woman or a man should not influence the treatment received by others (Martínez-Bascuñán, 2011). It should be noted that, as we explained before, both terms *equity* and *equality* are interlinked.

Third, the term *justice* is frequently associated with gender issues (Martínez-Bascuñán, 2011). If we understand that justice "*does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many*" (Rawls, 1997, p. 3), we will agree that this concept could refer implicitly to the injustice against women, although it should be noted that women are not a minority in global society. Moreover, gender equality is perceived as "*a positive ideal and its pursuit is depicted as a core requirement of social justice*" (Squires, 2007, p. 1). Fourth, *fairness* is closely related to justice, and refers to the "*ability to judge without reference to one's feelings or interests*" (Velasquez et al., 1990). It could be stated that gender biases in decision-making are unfair (Camilli, 2005; Friedler et al., 2016), as lead to make decisions based on stereotypes.

Fifth, *inclusion* should account for characteristics as race and gender: women and black, among others, have been excluded historically from the public sphere, based on essentialist ideas about the concept of *others* (Lister, 2017). Sixth, the notion of *non-discrimination* is related to the concept of equity: in some circumstances, discrimination has negative connotations, but in others, it has positive associations, as it could be used to affirmative actions in favor of certain groups (Martínez-Bascuñán, 2011; Hellman, 2008). Here, we understand that *non-discrimination* alludes to the negative connotation of discrimination and refers to the avoidance of "*less favorable treatment because of certain traits, such as their race, age, gender, or religion*" (Moreau, 2010, p. 143). Seventh, *non-bias* is related to the reproduction of stereotypes (gender included) and it has been studied in the field of technology and AI (Caliskan et al., 2017; May et al., 2019; Costa & Ribas, 2019; Bolukbasi et al., 2016).

Eighth, *diversity* and *representativeness* are also two terms related to gender and women inclusion in organizations (van Knippenberg & Schippers, 2007; Ahmed & Swan, 2006), and data (Nowakowski et al., 2016). Nineth, the word *empowerment*, as Rowlands states, is "*about bringing people who are outside the decision- making process into it*" (1996, p. 87) and it incorporates the feminist perspective considering the oppression suffered by women. Finally, the term *shared-benefits* is included as one ethical principle of Artificial Intelligence, and it refers to the use of outcomes of AI to "*benefit and empower as many people as possible*" (Fjeld et al., 2020, p. 51). We assume this expression associates with other concepts as inclusion, empowerment, and non-discrimination, which, as we exposed earlier, have gender connotations. All in all, here we assume the following: when a discourse includes explicit allusions to gender, and/or states that the AI should promote equity, equality, justice, fairness, non-discrimination, diversity, citizen's empowerment, and shared benefits, with representative and non-bias data, it should be considered gender sensitive.

## 4. Research methods

In our aim to understand assumptions, expectations and knowledge on AI regarding gender, we have focused on the documents and reports published by the European Union (EU) and the Spanish government. On the one side, the reason for choosing the EU as a case study is its relevance for settling policy frameworks and regulations across different fields, including technology. The EU is recognized as the main settler and legislator of the boundaries and limits for policy-making for all member states (Wallace et al., 2020) and traditionally has shaped their digital agendas using different Europeanization mechanisms (Criado, 2012). For this reason, the European Commission has published a common strategy on AI (European Commission, 2018), that will serve as a guide for the rest of EU member states. This is also the case of gender policies and discourses (Bustelo & Lombardo, 2006; Lombardo & Meier, 2006; Lombardo & Kantola, 2019). Additionally, the EU has taken the lead in promoting an ethical approach to AI compared to other AI giants as China and the US (Cath et al., 2018; Valle-Cruz et al. 2020), and has publicly committed to promote technological EU standards to other regions of the world. In this sense, the EU launched a joint strategy with Latin American and the Caribbean to work on a common future where technological development plays a key role, through projects such as BELLA (European Commission, 2018d, 2019b).

In consequence, the Spanish case has been selected as a study case due to its relevance as a central piece in the creation of agreements, treaties and synergies between Latin America and the European Union (Grugel, 2002). The existence of a shared language and culture is a decisive nexus between Spain and Latin America that has been broadly recognized by most European and non-European institutions (Freres, 2000). In this sense, the Spanish Government has acknowledged on several occasions its role as a mediator between both continents, including its role in the promotion of an ethical and human-centered AI (Ministry of Economic Affairs and Digital Transformation, 2020). Additionally, it is worth noting that the presence of the Spanish language in the configuration and design of AI is of extreme importance to be able to establish a high-quality and representative programming language that works for everyone. Spanish is the second most spoken mother tongue in the world (Cervantes Institute, 2020). Thus, it should be proportionally represented in AI to avoid language, and therefore, cultural biases.

Once both case studies were selected, we proceeded to the document and report selection. In the case of the European Union, we chose the following selection criteria. Since the European Strategy for AI was released in March 2018, relevant documents and reports on AI have been in charge of three different working groups: the AI Watch, the High-Level Expert Groups on AI (HLEG), and the European Commission itself. Thus, in representation of the European Union we have analyzed all documents and reports released by these three groups that are available at their corresponding websites.[1] Then, for the case of Spain, we have selected the three official documents and reports released by the Spanish government in relation to AI and the Secretary of State for Digitalization and Artificial Intelligence appearance at the Spanish National Congress to explain the national budget on AI. This means 4 documents for the case of Spain, 19 for the UE and a total of 23 (see Table 2).

For the analysis of explicit and implicit references to gender in these documents, we followed content analysis and thematic analysis. Using these techniques (Rowen, 2009), we were able to identify the categories included in our analytical framework to gauge the gender perspectives in the European and

---

[1]AI Watch: https://knowledge4policy.ec.europa.eu/navigation-page/key-reports-analysis_en; HLEG: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence; European Commission: https://ec.europa.eu/digital-single-market/en/artificial-intelligence.

Table 2
List of analyzed documents and reports

| Name of the document/report | Issue date | Issuer |
| --- | --- | --- |
| Diary of Sessions of the Congress: Carme Artigas appearance | November 4th, 2020 | Secretary of State for Digitalization and Artificial Intelligence |
| Spanish National Strategy on AI | November 2020 | Spanish Government (Ministry of Economic Affairs and Digital Transformation) |
| Strategy Digital Spain 2025 | July 2020 | Spanish Government (Ministry of Economic Affairs and Digital Transformation) |
| Spanish RDI Strategy in Artificial Intelligence | 2019 | Spanish Government (Ministry of Science, Innovation and Universities) |
| Artificial Intelligence in public services | 2020 | AI Watch |
| Defining Artificial Intelligence | 2020 | AI Watch |
| AI Uptake in Health and Healthcare | 2020 | AI Watch |
| AI Watch 2019 Activity Report | 2020 | AI Watch |
| TES analysis of AI Worldwide Ecosystem in 2009–2018 | 2020 | AI Watch |
| National strategies on Artificial Intelligence: A European perspective in 2019 | 2020 | AI Watch |
| Estimating investments in General Purpose Technologies: The case of AI Investments in Europe | 2020 | AI Watch |
| European enterprise survey on the use of technologies based on artificial intelligence | 2020 | AI Watch |
| Sectoral Considerations on the Policy and Investment Recommendations | July 23rd 2020 | HLEG |
| The assessment list for Trustworthy Artificial Intelligence (ALTAI) | July 17th 2020 | HLEG |
| Policy and Investment Recommendations for Trustworthy AI | June 26th 2019 | HLEG |
| Ethics guidelines for trustworthy AI | April 8th 2019 | HLEG |
| A definition of Artificial Intelligence: main capabilities and scientific disciplines | April 8th 2019 | HLEG |
| Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics | February 19th 2020 | European Commission |
| White Paper on Artificial Intelligence – A European approach to excellence and trust | February 19th 2020 | European Commission |
| Building Trust in Human-Centric Artificial Intelligence | April 8th 2019 | European Commission |
| Coordinated Plan on Artificial Intelligence | December 7th 2018 | European Commission |
| Declaration of Cooperation on AI | April 10th 2018 | European Commission |
| Artificial Intelligence for Europe | March 25th 2018 | European Commission |

Source: Own elaboration.

Spanish official documents with the strategies on AI. The researchers reviewed and manually coded all the documents, using as key words the terms listed as implicit and explicit references to gender (see Table 1). Each author coded independently all the documents, and then conducted a collective double-check of the codification to continue with the analysis.

One of the advantages of this research method is its efficiency, as all the documents were publicly available (Bowen, 2009). Moreover, we followed a general inductive approach to the analysis: The interpretations were made directly from raw data, allowing us to identify the main categories of our analytical framework and "*develop a framework of the underlying structure of experiences or processes*

*that are evident in the raw data*" (Thomas, D., 2006, p. 237). This approach is adequate to understand a novel issue from an exploratory point of view.

## 5. Results

In this section we present the main results of our analysis. First, we analyze the explicit and implicit references to gender, in both European and Spanish documents, regarding the nature of AI. Second, we present the results for the dimension related to the AI strategy. Third, we explain the gender references in the written discourses about the use of AI. The following structure is due to the fact that the nature, strategy and use of technology are the three dimensions in technology frames identified by Orlikowski. Thus, through the analysis of these three dimensions we will explore assumptions, expectations, and knowledge on AI. In Table 3, a concise approach to the results is presented.

### 5.1. Gender approach in the nature of AI

This dimension encompasses the assumptions, expectations and knowledge on AI about the nature of this technology. The aim of this section is to answer the question, what is AI? To achieve this goal, we present the definitions and theoretical approaches that the European Union and Spain have used for AI. Although since the first moment the UE has emphasized the need for a human-centric and inclusive AI for Europe (European Commission, 2018a; 2018c), the first definition of AI that appeared in the document *"Artificial Intelligence for Europe"* (European Commission, 2018) did not make explicit the humanly shaped condition of this technology. That definition was maintained until 2019 when the EU, assisted by the HLEG, evolved towards a clearer socio-technical perspective by acknowledging that *"Artificial intelligence systems are software (and possibly also hardware) systems designed by humans"* (HLEG, 2019a, p. 6). Consequently, the EU has classified AI as a new technology based on the fundamental rights and values of the EU (European Commission, 2018a; 2018b; 2018c; 2019a; 2020a, 2020b; HLEG, 2019b; 2029d; Samoili et al., 2020). The necessity of an AI based on Europe's fundamental rights and values has called for a lawful (that respects the law), ethical (respectful of human autonomy, preventive of harm, fair and explicable) and robust (that does not cause unintentional harm) AI that distinguishes European AI from the rest of the world (HLEG, 2019b, European Commission, 2020a).

In regard to Spain, the southern European country has clearly followed the EU frame on the nature of AI. All the documents and reports that have been analyzed emphasize on shared values between the EU and Spain, as well as their common goal to develop a human-centric AI: *"It is also a shared commitment with our European partners (to help) the EU to become a leader in the deployment of an inclusive, ethic, trustworthy and economically efficient AI"* (Ministry of Economic Affairs and Digital Transformation, 2020, p. 4). It is also worth noting that two documents (the Spanish RDI Strategy in Artificial Intelligence and the Spanish National Strategy for Artificial Intelligence) make explicit references to socio-technical approaches of technology: *"The success of AI will depend on how people and machines work together to provide better services – transparent, reasonable and ethical – to potential users, in a world where we will be increasingly demanding in terms of the quality of services provided"* (Ministry of Science, Innovation and Universities, 2019, p. 40).

In none of the analyzed documents explicit references to gender have been found in regard to AI definitions. Therefore, it is not possible to gauge if European and Spanish institutions considered the gender dimensions for technology (gender in technology and gender of technology), explained in the theoretical section, to define AI in their strategies. However, considering the ideas expressed in the

Table 3
Implicit and explicit references to gender in European and Spanish documents

| Issuer | Dimensions of technological frames | Explicit references to gender | Implicit references to gender |
|---|---|---|---|
| European Union | Nature | None | – Human-centric and inclusive AI for Europe<br>– AI designed by humans<br>– AI based in Europe's fundamental rights and values |
| | Strategy | – Using AI to promote gender balance | – Using AI to improve human welfare and freedom<br>– Using AI to achieve a fair society and foster equality<br>– AI as a tool to empower humans and societies |
| | Use | – Policies to build data and infrastructure for AI considering gender dimensions<br>– Policies to generate appropriate skills for AI considering gender sensitivity dimensions | – Policies to establish an appropriate governance and regulatory framework to identify unfair bias and discrimination in AI |
| Spain | Nature | None | – Human-centric AI, in line with the European values<br>– Sociotechnical approach to AI |
| | Strategy | – Using AI to reduce gender gap<br>– Digital transformation to promote gender equality | – Using AI and other technologies to promote equality, justice, inclusion and diversity<br>– Using AI and other technologies to protect people at risk of exclusion<br>– Using AI and other technologies to reduce inequalities<br>– Using AI and other technologies to support the protection of human rights |
| | Use | – Policies to design algorithms that avoid gender bias and discrimination<br>– Policies to develop training programs to increase digital skills among women.<br>– Plans to promote gender equality in research teams and companies working on AI | – Policies to build data and infrastructure for AI considering the principle of non-discrimination and inclusion, and the reduction of biases. Use of representative data in AI.<br>– Policies to train different communities in digital skills, including vulnerable groups<br>– Governance and regulatory framework: programs to make the public services more inclusive and accessible. Charter of Digital Rights |

Source: Own elaboration based on the official documents from the EU and Spain listed in Table 2.

official documents, we assume the following: if these levels of government denote AI as a co-constructed technology, it is more likely they will prioritize a human-centric perspective and, therefore, European values and human rights, including gender equality. It should be noted that these general considerations do not guarantee the real implementation of policies following these criteria, but it is a general framework acting as a starting point for the design and implementation of strategies with a gender perspective.

## 5.2. Gender approach in the AI strategy

Regarding the dimension on the AI strategy, we included all references to motivations underlying the

implementation of AI in the public sector. We have also focused on goals and objectives that guide the introduction of AI. It is important to note that we only present those motivations and goals that either explicitly or implicitly aim to promote gender equality and balance. In this regard, the EU frames AI as a technology that can improve human welfare and freedom as well as it "*can help to facilitate the achievement of the UN's Sustainable Development Goals, such as promoting gender balance*" (HLEG, 2019b, p. 4). AI also is seen as a technology that "*can contribute to achieving a fair society, by helping to increase citizens' health and well-being in ways that foster equality in the distribution of economic, social and political opportunity*" (HLEG, 2019b, p. 9). In the same line, the UE envisions AI as a tool to empower humans and societies (European Commission, 2018a, p. 16; HLEG, 2019c, p. 10).

In the case of Spain, three of the four documents analyzed include explicit references to gender. They mention the use of AI to reduce gender discrimination and gender gaps, and to promote gender equality: "*There is a great opportunity to use AI as an element to transform the economy and the society, including the performance of public services and transparency of public administrations, as well as addressing major social challenges such as the gender gap, the digital divide or the ecological transition*" (Ministry of Economic Affairs and Digital Transformation, 2020a, p. 11). Moreover, in her appearance at the Spanish National Congress, the Secretary of State for Digitization and Artificial Intelligence, pointed out that the digital transformation should help the process of recovery from the 2020 COVID-19 pandemics, which should take into account the promotion of gender equality (Congress of Deputies of Spain, 2020). Concerning the implicit references, the documents mention the use of AI and other technologies to promote equality and protection of people at risk of exclusion (Ministry of Science, Innovation and Universities, 2019); to promote justice, inclusion and diversity; to reduce inequalities (Ministry of Economic Affairs and Digital Transformation, 2020b), and to support the protection of human rights and foster the inclusion and social welfare (Ministry of Economic Affairs and Digital Transformation, 2020a).

Regarding the gender dimensions for technology, the strategies proposed by European and Spanish institutions are in line with both gender *in* technology and gender *of* technology. As we explained in the section devoted to the nature of AI, in this case there are general statements about the inclusion of gender perspectives, explicitly and implicitly. Hence, we understand that these assumptions would permeate both the design processes (gender in technology) and the symbolism associated with gender roles in AI (gender of technology). Therefore, this strategic dimension of AI is more aligned with gender perspectives, encompassing clearer statements in both European and Spanish official documents.

### 5.3. Gender approach in the use of AI

In the dimension on the use of AI we focused on answering the question of how AI is going to be used. Thus, in this section we have highlighted the specific measures and policies that the European Union and Spain plan to implement in order to achieve the goals and objectives that they settled for AI. In this regard, the European Union has planned a large list of actions that will help to implement a trustworthy AI. Within the measures considered we can find:

Policies to build data and infrastructure for AI: In regard to gender equality, the EU explicitly calls for high-quality data, measures and "*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*" (European Commission, 2020a, p. 19).

Policies to generate appropriate skills and education for AI: In this sense the EU plans to promote measures to reduce the gender gap in STEM professions such as, setting "*incentives to offer gender*

*sensitivity trainings for STEM educators"* (HLEG, 2019d, p. 32) or *"provide sustained substantial incentives and funds for initiatives that focus on closing the "self-efficacy gender gap" in primary and secondary education systems"* (HLEG, 2019d, p. 33), etc.

Policies that help establishing an appropriate governance and regulatory framework: In this section we found especially relevant the development of auditing mechanisms for AI systems that *"should allow public enforcement authorities as well as independent third-party auditors to identify potentially illegal outcomes or harmful consequences generated by AI systems, such as unfair bias or discrimination"* (HLEG, 2019d, p. 41).

In the case of Spain, its documents and reports include actions and plans that can be distributed in the same categories set by the EU: Regarding data and infrastructure for AI, the Spanish government has called for the design of inclusive and equitable AI and the improvement of the digital infrastructure considering the principle of non-discrimination and inclusion. Mentions to plans that aim to the reduction of biases and the use of representative data in the design of AI are also recurrent in Spanish reports and documents on AI (Ministry of Science, Innovation and Universities, 2019; Ministry of Economic Affairs and Digital Transformation, 2020a). Explicitly, Spain aims to design algorithms that avoid gender bias: *"It is a condition in the development of technologies and applications of AI linked to this RDI Strategy to avoid the negative bias and prejudices of our society, such as gender, race, or other forms of discrimination, which must be avoided by decision support systems"* (Ministry of Science, Innovation and Universities, 2019, p. 40).

A new Data Office, and the associate Chief Data Officer (CDO), has been settled by the Spanish government to guarantee the adequate use of government data. This Data Office and CDO will be responsible for designing strategies to manage data and ensure security as well as big data and AI governance (Official State Gazette, 2020). There is also a plan to create a new project named Big Data for Social Good, aiming at developing initiatives with open data, citizen-generated data and government to citizen transactions (Ministry of Economic Affairs and Digital Transformation, 2020a). Despite official documents do not explicitly mention the term gender, the attention to an *"Artificial Intelligence inclusive, ethical, transparent, which promotes equal opportunities"* (Official State Gazette, 2020) suggests the potential to include a gender sensitive perspective in these policies.

In the case of skills and education for AI, Spain aims to train different communities of individuals in digital skills (Ministry of Science, Innovation and Universities, 2019), including vulnerable groups (Ministry of Economic Affairs and Digital Transformation, 2020b), and people at risk of exclusion (Ministry of Economic Affairs and Digital Transformation, 2020a). For this reason, the government has plans to develop training programs to increase digital skills among women, and initiatives to promote gender equality in research teams and companies working on AI. The plan Digital Spain 2025 *"aims at meeting the demand of experts in digital technologies, including cutting-edge technology, such as data analysis, Artificial Intelligence or cybersecurity. Special attention will be paid to the gender composition of these specialists"* (Ministry of Economic Affairs and Digital Transformation, 2020b, p. 30). Additionally, we found references to projects fostering the creation of digital start-ups led by women (Ministry of Economic Affairs and Digital Transformation, 2020a; Congress of Deputies of Spain, 2020). Finally, in reference to governance and regulatory framework, Spain has stated its intention to promote inclusive access to AI (Ministry of Science, Innovation and Universities, 2019) through programs that transform public services and make them more inclusive and accessible. Also, there are references to the design and implementation of the Charter of Digital Rights, aiming at *"guaranteeing the protection of individual and collective digital rights of the citizens, both in the national and European context, inspiring the*

*development of a humanist framework at a global level, to contribute to close the gaps (digital, gender, etc.)*" (Ministry of Economic Affairs and Digital Transformation, 2020a, p. 70).

In the analyzed documents, both in the European and Spanish cases, there are only references to gender in technology regarding the use of AI: most measures are related to the use of representative datasets in the design of algorithms, policies to improve digital skills among women, and the increase participation of women in design teams and institutions. Even if these measures are also related to gender of technology, as design processes influence the symbolic representations of AI and subsequent use of artifacts, we have not found any direct reference to the avoidance of stereotypes in the characterization of gender roles in AI products such as chatbots or voice assistants. Nonetheless, these documents compose the general framework that public and private developers will follow during the design of specific devices that could be gendered themselves.

## 6. Discussion

In this section we highlight the main contributions from this study and elaborate on some guidelines for policymakers and practitioners who want to implement an ethical, inclusive, and gender-sensitive AI. First, considering the terms found in the official documents, it is important to note that AI framing in the EU denotes a socio-technical approach, reflecting somehow constructivist perspective, as explained in the theoretical framework and the section devoted to the nature of AI. Following socio-technical and feminist approaches to technology (Lagesen, 2015; Faulkner, 2001; Berg & Lie, 1995; Wajcman, 2009), the assumption that technology is designed by and for humans is essential to understand the potentialities of technology to transform society, including the reduction of gender gaps. This way of framing AI seems to be also adopted in Spain, closely following European guidelines. Thus, in both cases, the studied documents and reports have several and reiterated mentions to the humanly designed nature of AI, its reality as a co-constructed technology and the necessity of a human-centric approach to it. In this sense, references to a value-based AI have been commonly identified, where value based always refers to the values and rights displayed in the Charter of Fundamental Rights of the European Union (European Commission, 2000).

Explicit references to equality between men and women are included in Charter III that addresses equality. Article 21, on non-discrimination, states that *"Any discrimination based on any ground such as sex, 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 or sexual orientation shall be prohibited"*. Additionally, article 23, on equality between men and women, affirms that *"Equality between men and women must be ensured in all areas, including employment, work and pay. The principle of equality shall not prevent the maintenance or adoption of measures providing for specific advantages in favour of the under-represented sex."*. It is worth noting that in both cases the EU seems to refer to sex as the reason of exclusion and discrimination, thus implying that the definitory feature to classified someone as a woman is sex. On the contrary, in all documents and reports on AI examined there is no single reference to sex, but always to gender.

This might show an evolution in the predominant views of regulators, legislators, politicians and European citizens, as a whole, on gender equity, a condition to achieve gender equality. The introduction of gender as the main discriminatory element might be due to the auge and increasing political value of feminism in the European political scenario. However, lately, a strong debate between traditional forms of feminism and queer theory might call for a more detailed approach to gender issues in technology (Jagose, 1996). In this sense, it is important for the EU to define whether the measures and policies target

to increase equality between men and women will be based on sex, as the reason to associate gender roles that emphasize discrimination, or gender alone with independence of sex.

Additionally, despite the existence of a gender-sensitive approach to AI, in none of the examined documents the word feminism or feminist AI is mentioned. Even if gender issues are in the public agenda and are deemed as essential to achieve equality within the framework of the Agenda 2030 and the Sustainable Development Goals (United Nations, 2015), the avoidance of using the word feminism might reflect the intention of the EU and Spain to not politicize AI, as still today this term still raises strong responses among European citizens and is usually related to certain political affiliations. This would make sense in the case of the EU as the European Institute for Gender Equality defines feminism as a *"political stance and commitment to change the political position of women and promote gender equality, based on the thesis that women are subjugated because of their gendered body, i.e., sex."* On the contrary, the Royal Academy for the Spanish Language defines feminism as *"the principle of equal rights between men and women"*. One could think that this difference in approaches to feminism would be reflected on the AI strategies and reports of both institutions; however, in this case it seems obvious that Spain is closely following the European guidance. Thus, deepen on the politics of Europeanization (Featherstone & Radaelli, 2003).

Second, both cases include explicit references to gender. While the EU talks about directing AI to fulfill the UN's Sustainable Development Goals, including gender equality; Spain visualizes AI as a tool that can contribute to address major social challenges such as gender equality. In the case of the EU, the motivations and goals that underlie the implementation of AI are in accordance with the abovementioned fundamentals values and rights of the institution. In the case of Spain, it is important to point out that apart from following the guide of EU on AI issues, its public sector has been reflecting on the relationship between gender and technology prior to the publication of the already mentioned official documents (e.g., the Ministry of Economic Affairs and Digital Transformation published in March 2019 a white paper on women and technology). In addition to the explicit references to gender, both cases include multiple references to the promotion of justice, equality, inclusion and diversity through technology and, specifically, through AI. Although it should be noted that the documents and reports do not include the definitions of these terms, meaning that it is not possible to understand their views on complex concepts such as "justice" or "equality". As we previously explained, the reference to these elements should be considered as implicit references to gender, based on the academic and legal literature on the topic, and serve as the foundations to elaborate more on the gender perspective in AI.

Third, we have listed three groups of policies promoted by the EU and Spain that are aimed to reduce gender discrimination. Here, we highlight (1) policies to build data and infrastructure for AI, (2) policies to generate appropriate skills and education for AI, and (3) policies that help establishing an appropriate governance and regulatory framework. In this sense, we have reasons to believe that the EU and the member states are taking seriously the risks derived from an unethical use of AI. Successful cases such as *Tengai*, an AI system implemented in Sweden for internal management purposes in the general public sector, that has been adopted, among other reasons, to improve recruitment services by reducing unbiased selection processes, are still available (AI Watch, 2020a). On the contrary, unsuccessful cases, such as, *Syri*, in the Netherlands, or an unemployment profiling system in Poland, have been canceled. *Syri* has been referred to as "*discriminatory towards the poor and vulnerable citizens*" (AI Watch, 2020a, p. 46). In the case of the unemployment profiling system in Poland, *"many unemployed persons have complained through administrative courts, claiming the categorisation to be unjust"* (Misaruca et al., 2020).

However, there are also important reasons to be sceptic about the real commitment of the Spanish government with a gender-sensitive AI that works to reduce gender inequality and discrimination. Nowadays, one of the most popular AI systems that are significantly being adopted by public administrations

are intelligent virtual assistants, also known as Chatbots. In fact, in Spain, due to the national lockdown caused by the COVID-19 pandemic in 2020, many public administrations started using chatbots to assist and provide citizens with medical information and solve possible doubts derived from the health crisis. In this sense, it has been observed that most of these virtual assistants are characterized as women, due either to its name, its voice, or the language used that is commonly associated with women, in line with other chatbots developed all over the world (Feine et al., 2019). As we noted in our theoretical section on gender of technology, this approach to AI only reinforces already existing gender stereotypes that work against gender equality and women empowerment. Although gendered chatbots are not a novelty (Marino, 2006; McDonnell & Baxter, 2019; UNESCO, 2019), due to the formal commitment of the EU and the Spanish government on the promotion of gender equality through AI, one could have waited for a different outcome.

Examples like these are not an exception in the EU. This proves that strategies on AI alone, even if elaborate from a constructivist approach and an ethical perspective, are not enough. The EU needs to deepen on an openly feminist approach and policy framework for AI. This would call for increasing the number of explicit references to gender related issues or prioritize gender design issues in EU funding projects. Moreover, as we explained in the results section, most references to gender in official documents are related to the gender in technology dimension, while the gender of technology dimension is neglected, at least explicitly. Even if these dimensions are interrelated, it would be important to take into account all the edges of the issue to ensure that the relations between gender and AI are properly addressed. On the other hand, while implicit references are undoubtedly important, it is more likely that a more explicit approached to a feminist AI will be more successful answering these problems. As citizens and policy-makers acknowledge the real biases of AI associated to gender, the chances of taking actions to tackle these issues through will also increase.

Additionally, it would be greatly beneficial to women if the EU elaborated on new regulation that would target the use of AI from a feminist perspective in order to avoid cases as the ones mentioned above. As the EU plays a key role as regulatory framework, Spain and the rest of the European countries would be largely benefit by it. In fact, the European Network of Quality Bodies (Equinet), in a co-funded project by the EU, released in the year 2020 a report named "Regulating AI: the new role for Equality Bodies. Meeting the new challenges to equality and non-discrimination from increased digitalization and the use of AI" (Allen & Masters, 2020), that could be used as a guide. After all, good intentions and words might be the first step, but in order to implement a gender-sensitive AI and make real change in women's lives, both the EU and Spain need to keep working on an ethical AI that goes beyond governmental formalities.

## 7. Conclusions

After our systematic analysis of official documents from the European Union (EU) and Spain, we found that, in broad terms, there is a political commitment to the promotion of a gender-sensitive approach to AI. The framework established by the EU provides the basis for the Spanish national strategy on AI, meaning that there is a common ground in both cases. Despite the fact that explicit references to gender are absent from the definitions of AI, the allusions to a human-centric, ethical and inclusive AI lay on the foundations to incorporate a gender perspective that would be more explicit in the strategies and use of AI. Our analysis suggests, however, that there are doubts regarding the definitions of 'gender', the perspectives on justice, inclusion, ethical and other implicit concepts, and, especially, the real commitment with the gender agenda on the technological uses, considering some practical developments in the EU and Spain. Thus, we conclude that some possible steps to reduce gender inequality through AI would be

to assume a more openly feminist approach to AI, to increase the number of explicit references to gender in documents and reports, to incorporate all the dimensions to study gender and technology (i.e., gender in technology and gender of technology), and to build on new regulation that targeted issues related to AI and gender.

Also, this study presents some limitations. First, we conducted an exploratory documental analysis, following a general inductive approach. Despite the advantages of this technique, especially to understand a novel issue as AI, we acknowledge its limitations to gauge all the elements influencing the political discourse and public policy agenda-setting, as in other recent studies (Ulnicane et al., 2020). We also lack of stronger empirical data that would be extremely beneficial for this work and that will be part of the future developments of this approach. Moreover, it would be adequate to include other cases in Europe and Latin America, to better understand the influence of context on their discourses. Finally, we are aware that the theoretical approach of this article, based on a binary definition of gender, has its own limitations. Thus, it would be adequate to develop future research considering the perspectives of the theory of performativity (Butler, 2007) and the queer theory (Jagose, 1996).

All in all, we have presented an exploratory study on the approaches of the EU and Spain to gender and AI from a feminist perspective. We do have intended a better understanding of the current state of the art in this matter that could be useful to develop a more complex analytical framework to evaluate gender equality in AI in the near future. It would also be interesting to complement these findings with the attention to political actors and public servants involved in the definition of the strategies in different international contexts, to gain more insights on their gender perspectives and the technological frames regarding AI. Finally, we acknowledge this work as a first step into a broader study on gender and AI in the EU. This is extremely important as disruptive technologies play a key role in the present and future of our societies. Therefore, it is essential to advance in the study of their ethical and social implications, including their potentialities and risks for gender equality.

## Acknowledgments

## References

Allen, R., & Masters, D. (2020). Regulating for an equal AI: A new role for equality bodies Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence. Equinet. https://equineteurope.org/wp-content/uploads/2020/06/ai_report_digital.pdf.

Ahmed, S., & Swan, E. (2006). Doing Diversity. *Policy Futures in Education]*, *4*(2), 96-100. doi: 10.2304/pfie.2006.4.2.96.

Avila, R., Brandusescu, A., Ortiz, J., & Thakur, T. (2018). *Artificial Intelligence: open questions about gender inclusion*. http://webfoundation.org/docs/2018/06/AI-Gender.pdf.

Berg, A.-J., & Lie, M. (1995). Feminism and Constructivism: Do Artifacts Have Gender? *Science, Technology, & Human Values*, *20*(3), 332-351. doi: 10.1177/016224399502000304.

Bimber, B. (2000). Measuring the gender gap on the Internet. *Social Science Quarterly*, *81*(3), 868-876.

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). *Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*. 30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain. https://papers.nips.cc/paper/2016/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf.

Bowen, G. (2009). Document Analysis as Qualitative Research Method. *Qualitative Research Journal*, *9*(2), 27-40. doi: 10.3316/QRJ0902027.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1-15.

Bustelo, M. & Lombardo, E. (2006). 'Los 'marcos interpretativos' de las políticas de igualdad en Europa: conciliación, violencia y desigualdad de género en la política' *Revista Española de Ciencia Política* (14), 117-140.

Butler, J. (1990). *Gender trouble. Feminism and the subversion of identity*. Routledge.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183-186. doi: 10.1126/science.aal4230.

Camilli, G. (2005). Test fairness. In R. Brennan (Ed.), *Educational Measurement*. American Council on Education/Praeger, pp. 221-256.

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018) 'Artificial Intelligence and the "Good Society": the US, EU, and UK approach', *Science and Engineering Ethics*, *24*(2), 505-528.

Cervantes Institute. (2020). Yearbook of Spanish in the World. https://cvc.cervantes.es/lengua/espanol_lengua_viva/pdf/espanol_lengua_viva_2020.pdf.

Collett, C., & Dillon, S. (2019). *AI and gender: Four Proposals for Future Research*. doi: 10.17863/CAM.41459.

Costa, P., & Ribas, L. (2019). AI becomes her: Discussing gender and artificial intelligence. *Technoetic Arts: A Journal of Speculative Research*, *17*(1/2), 171-193. doi: 10.1386/tear_00014_1.

Criado, J. I. (2012). Interoperability of e-Government for Building Intergovernmental Integration in the European Union. *Social Science Computer Review*, *30*(1), 37-60.

Criado, J. I., Sandoval-Almazan, R., Valle-Cruz, D., & Ruvalcaba-Gómez, E. A. (2021). Chief information officers' perceptions about artificial intelligence. *First Monday*, *26*(1).

Culyer, A. J. (2001). Equity – Some theory and its policy implications. *Journal of Medical Ethics*, *27*(4), 275-283. doi: 10.1136/jme.27.4.275.

De Nigris, S., Craglia, M., Nepelski, D., Hradec, J., Gomez-Gonzales, E., Gomez Gutierrez, E., Vazquez-Prada Baillet, M., Righi, R., De Prato, G., Lopez Cobo, M., Samoili, S., & Cardona, M. (2020). AI Watch: AI Uptake in Health and Healthcare, 2020, EUR 30478 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-26936-6 (online), doi: 10.2760/948860 (online), JRC122675.

Delipetrev, B., Tsinaraki, C., Nepelski, D., Gomez Gutierrez, E., Martinez Plumed, F., Misuraca, G., De Prato, G., Fullerton, K.T., Craglia, M., Duch Brown, N., Nativi, S., & Van Roy, V. (2020). AI Watch 2019 Activity Report, Desruelle, P. editor(s), EUR 30254 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-19515-3 (online),978-92-76-20649-1 (ePub), doi: 10.2760/007745 (online),10.2760/422011 (ePub), JRC121011.

Dietz, M. G. (2003). Current controversies in feminist theory. *Annual Review of Political Science*, *6*(1), 399-431. doi: 10.1146/annurev.polisci.6.121901.085635.

Doyal, L. (2000). Gender equity in health: Debates and dilemmas. *Social Science & Medicine*, *51*(6), 931-939. doi: 10.1016/S0277-9536(00)00072-1.

Eden, C. (1992). On the nature of cognitive maps. *Jorunal of Management Studies*, *29*(3), 261-265.

European Commission. (2020a). White Paper on Artificial Intelligence – A European approach to excellence and trust. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

European Commission. (2020b). Report on the safety and liability implications of Artificial. https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1593079180383&uri=CELEX:52020DC0064.

European Commission. (2020c). European enterprise survey on the use of technologies based on artificial intelligence. https://ec.europa.eu/digital-single-market/en/news/european-enterprise-survey-use-technologies-based-artificial-intelligence.

European Commission. (2019a). Building Trust in Human-Centric Artificial Intelligence. https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence.

European Commission. (2019b). Joint Communication to the European Parliament and the Council European Union, Latin America and the Caribbean: joining forces for a common future. https://eeas.europa.eu/sites/eeas/files/joint_communication_to_the_european_parliament_and_the_council_european_union_latin_america_and_the_caribbean_-_joining_forces_for_a_common_future.pdf.

European Commission. (2018a). Artificial Intelligence for Europe. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN.

European Commission. (2018b). Declaration of Cooperation. https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence.

European Commission. (2018c). Coordinated Plan on Artificial Intelligence. https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence

European Commission. (2018d). 2nd Workshop on Digital Cooperation between the European Union and Latin America & the Caribbean. https://ec.europa.eu/digital-single-market/en/news/2nd-workshop-digital-cooperation-between-european-union-and-latin-america-caribbean.

European Commission. (2000). Charter of Fundamental Rights of the European Union. Official Journal of the European Union, C 364/1.

European Institute for Gender Equality. (2020). Gender Equality Index 2020. https://eige.europa.eu/gender-equality-index/2020.

Faulkner, W. (2000). The Power and the Pleasure? A Research Agenda for "Making Gender Stick" to Engineers. *Science, Technology, & Human Values*, *25*(1), 87-119. doi: 10.1177/016224390002500104.

Faulkner, W. (2001). The technology question in feminism: A view from feminist technology studies. *Women's Studies International Forum*, *24*(1), 79-95. doi: 10.1016/S0277-5395(00)00166-7.

Featherstone, K., & Radaelli, C. M. (Eds.). (2003). The politics of Europeanization. OUP Oxford.

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2020). Gender Bias in Chatbot Design. In Følstad, A., Araujo, T., Papadopoulos, S., Lai-Chong Law, E., Granmo, O., Luger, E., Bae Brandtzaeg, P. (Eds.) *Chatbot Research and Design* (11970:79-93). Lecture Notes in Computer Science. Cham: Springer International Publishing, doi: 10.1007/978-3-030-39540-7_6.

Ferrando, F. (2014). Is the post-human a post-woman? Cyborgs, robots, artificial intelligence and the futures of gender: a case study. *European Journal of Futures Research*, *2*(1), 43. doi: 10.1007/s40309-014-0043-8.

Fjeld, J., Achten, N., Hilligoss, H., Nagy, A., & Srikumar, M. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3518482.

Freres, C. (2000). The European Union as a global "civilian power": development cooperation in EU-Latin American relations. *Journal of Inter-American Studies and World Affairs*, 63-85.

Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2016). On the (im)possibility of fairness. ArXiv:1609.07236 [Cs, Stat]. http://arxiv.org/abs/1609.07236.

Giddens, A. (1984). The Constitution of Society: Outline of the Theory of Structure. Berkeley CA: University of California Press.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making, *Annual Review of Psychology*, *62*, 451-482.

Gil-Juárez, A., Feliu, J., & Vitores, A. (2018). Mutable technology, immutable gender: Qualifying the "co-construction of gender and technology" approach. *Women's Studies International Forum*, *66*, 56-62. doi: 10.1016/j.wsif.2017.11.014.

Gioia, D. A. (1986). Symbols, scripts, and sensemaking Creating meaning in the organizational experience. In The Thinking Organization. Jossey-Bass, San Francisco, Calif., 49-74.

Grugel, J. (2002). Spain, the European Union and Latin America: Governance and Identity in the Making of "New" Inter-Regionalism. *Real Instituto Elcano* http://www.realinstitutoelcano.org/wps/portal/rielcano_en/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_in/zonas_in/dt9-2002.

Harwell, D., & Dou, E. (2020, December 8th). *Huawei tested AI software that could recognize Uighur minorities and alert police, report says*. The Washington Post. https://www.washingtonpost.com/technology/2020/12/08/huawei-tested-ai-software-that-could-recognize-uighur-minorities-alert-police-report-says/.

Hellman, D. (2008). *When is Discrimination Wrong?* Harvard University Press.

HLEG. (2019a). A definition of Artificial Intelligence: main capabilities and scientific disciplines. https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines.

HLEG. (2019b). Ethics guidelines for trustworthy AI. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.

HLEG. (2019c). Policy and Investment Recommendations for Trustworthy AI. https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence.

HLEG. (2019d). The assessment list for Trustworthy Artificial Intelligence (ALTAI). https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment.

HLEG. (2019e). Sectoral Considerations on the Policy and Investment. https://futurium.ec.europa.eu/en/european-ai-alliance/document/ai-hleg-sectoral-considerations-policy-and-investment-recommendations-trustworthy-ai.

Jagose, A. (1996). *Queer theory: An introduction*. NYU Press.

Jobin, A., Ienca, M., & Vayena, E. (2019). Artificial Intelligence: The global landscape of ethics guidelines. *Nature Machine Intelligence*, *1*, 389-399.

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 10197. doi: 10.1016/j.telpol.2020.101976.

Lagesen, V. (2015). *Gender and Technology: From exclusion to inclusion?* In J. Wright (Ed.), International Encyclopedia of the Social & Behavioral Sciences. Elsevier, pp. 723-728.

Leavy, S. (2018). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering – GE '18*, 14-16. doi: 10.1145/3195570.3195580.

Lie, M. (1995). Technology and masculinity: the case of the computer. *The European Journal of Women's Studies*, *2*, 379-394. doi: 10.1177/135050689500200306.

Lister, R. (2017). *Citizenship: Feminist Perspectives*. Macmillan International Higher Education.

Lombardo, E., & Kantola, J. (2019). European Integration and Disintegration: Feminist Perspectives on Inequalities and Social Justice. *JCMS: Journal of Common Market Studies*, *57*(S1), 62-76.

Lombardo, E., & Meier, P. (2006). Gender mainstreaming in the EU: Incorporating a feminist reading? *European Journal of Women's Studies*, *13*(2), 151-166.

Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. *Human Behavior and Emerging Technologies*, *1*(2), 141-148. doi: 10.1002/hbe2.140.

Margetts, H., Dorobantu, C. (2019). 'Rethink government with AI', *Nature*. Nature Publishing Group, 163–165. doi: 10.1038/d41586-019-01099-5.

Marino, M. C. (2006). *I, chatbot: the gender and race performativity of conversational agents*. University of California, Riverside.

Martínez Bascuñán, M. (2011). 'Ha quedado obsoleta la política de la diferencia': Una exploración y propuesta. *Política y Sociedad*, *48*(3), 603-619. doi: 10.5209/rev_POSO.2011.v48.n3.36437.

Martínez-Bascuñán, M. (2015). Simone de Beauvoir y la teoría feministas contemporánea: Una revisión crítica. *Revista Jurídica*, *31*, 331-348. https://repositorio.uam.es/handle/10486/673872.

May, C., Wang, A., Bordia, S., Bowman, S. R., & Rudinger, R. (2019). On Measuring Social Biases in Sentence Encoders. ArXiv:1903.10561 [Cs]. http://arxiv.org/abs/1903.10561.

McDonnell, M., & Baxter, D. (2019). Chatbots and gender stereotyping. *Interacting with Computers*, *31*(2), 116-121.

Ministry of Economic Affairs and Digital Transformation. (2019). Libro blanco de las mujeres en el ámbito tecnológico. https://www.mineco.gob.es/stfls/mineco/ministerio/ficheros/libreria/LibroBlancoFINAL.pdf.

Ministry of Economic Affairs and Digital Transformation. (2020a). Spanish National Strategy on Artificial Intelligence. https://www.lamoncloa.gob.es/presidente/actividades/Documents/2020/ENIA2B.pdf.

Ministry of Economic Affairs and Digital Transformation. (2020b). Digital Spain 2025. https://portal.mineco.gob.es/Recursos Articulo/mineco/prensa/ficheros/noticias/2018/Agenda_Digital_2025.pdf.

Misuraca, G., Van Noordt, C. (2020). AI Watch – Artificial Intelligence in public services, EUR 30255 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-19540-5 (online), doi: 10.2760/039619 (online), JRC120399.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1-21.

Moreau, S. (2020). What Is Discrimination? *Philosophy & Public Affairs*, *38*(2), 143-179.

Nepelski, D., & Sobolewski, M. (2020). Estimating investments in General Purpose Technologies. The case of AI Investments in Europe, EUR 30072 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-10233-5, doi: 10.2760/506947

Nowakowski, A. C. H., Sumerau, J. E., & Mathers, L. A. B. (2016). None of the above: Strategies for Inclusive Teaching with "Representative" Data. *Teaching Sociology*, *44*(2), 96-105. doi: 10.1177/0092055X15622669.

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.E., Ruggieri, S. et al. (2020). Bias in Data-driven Artificial Intelligence Systems – An Introductory Survey. *WIREs Data Mining and Knowledge Discovery 10*(3), 1-14. doi: 10.1002/widm.1356.

Official State Gazette. (2020). Orden ETD/803/2020, de 31 de julio, por la que se crea la División Oficina del Dato y la División de Planificación y Ejecución de Programas en la Secretaría de Estado de Digitalización e Inteligencia Artificial. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-10008.

Orlikowski, W. J., & Gash, D. C. (1994). Technological frames: making sense of information technology in organizations. *ACM Transactions on Information Systems (TOIS)*, *12*(2), 174-207.

Peña, P., & Varon, J. (2020, September 10th). *Decolonising AI: A transfeminist approach to data and social justice*. Medium. https://medium.com/codingrights/decolonising-ai-a-transfeminist-approach-to-data-and-social-justice-a5e52ac72a96.

Rawls, J. (1999). *A theory of justice* (Rev. ed). Belknap Press of Harvard University Press.

Ridgeway, C. L. (2009). Framed Before We Know It: How Gender Shapes Social Relations. *Gender & Society*, *23*(2), 145-160. doi: 10.1177/0891243208330313

Rowland, J. (1996). Empowerment examined. In M. B. Anderson (Ed.), Development and social diversity. Oxfam.

Sáinz, M., Arroyo, L., & Castaño, C. (2020). *Mujeres y digitalización. De las brechas a los algoritmos*. Instituto de la Mujer y para la Igualdad de Oportunidades. https://www.inmujer.gob.es/diseno/novedades/M_MUJERES_Y_DIGITALIZACION_DE_LAS_BRECHAS_A_LOS_ALGORITMOS_04.pdf.

Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*, EUR 30117 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-17045-7, doi: 10.2760/382730.

Samoili, S., Righi, R., Cardona, M., López Cobo, M., Vázquez-Prada Baillet, M., & De Prato, G. (2020). TES analysis of AI Worldwide Ecosystem in 2009-2018, EUR 30109 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-16661-0, doi: 10.2760/85212.

Simon, H. (1957). Models of Man, New York: John Wiley

Simon, H. A. (1997). *Administrative Behavior: A study of Decision-Making Processes in Administrative Organizations*. Fourth Edi. The Free Press.

Simon, H. A. (1993). Decision Making: Rational, Nonrational, and Irrational, *Educational Administration Quarterly*, *29*(3), 392-411. doi: 10.1177/0013161X93029003009.

Sørensen, K. H. (1992). Towards a Feminized Technology? Gendered Values in the Construction of Technology. *Social Studies of Science*, *22*(1), 5-31. doi: 10.1177/0306312792022001001.

Squires, J. (2007). *The New Politics of Gender Equality*. Macmillan International Higher Education.

Thomas, D. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data, *American Journal of Evaluation*, *27*(2), 237-246. doi: 10.1177/1098214005283748.

Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice, *Science*, *211*(4481), 453-458.

Ulnicane, I., Knight, W., Leach, T., Stahl, B.C., & Wanjiku, W.G. (2020). Framing governance for a contested emerging technology insights from AI policy. *Policy and Society*, 1-20. doi: 10.1080/14494035.2020.1855800.

UNESCO. (2019). I'd blush if I could. Closing gender divides in digital skills through education. https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=85.

UNESCO. (2020). *Artificial Intelligence and Gender Equality*. https://unesdoc.unesco.org/ark:/48223/pf0000374174.

United Nations Population Fund (2005). *Frequently asked questions about gender equality*. https://www.unfpa.org/resources/frequently-asked-questions-about-gender-equality#:∼:text=What%20is%20the%20difference%20between,fair%20to%20women%20and%20men.&text=Equity%20leads%20to%20equality.

United Nations (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*. https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

Valle-Cruz, D., Criado, J. I., Sandoval-Almazán, R., & Ruvalcaba-Gomez, E. A. (2020). Assessing the public policy-cycle framework in the age of artificial intelligence: From agenda-setting to policy evaluation. *Government Information Quarterly* *37*(4), 101509.

Van Knippenberg, D., & Schippers, M. C. (2007). Work Group Diversity. *Annual Review of Psychology*, *58*(1), 515-541. doi: 10.1146/annurev.psych.58.110405.085546.

Van Roy, V., AI Watch – National strategies on Artificial Intelligence: A European perspective in 2019, EUR 30102 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-16409-8 (online), doi: 10.2760/602843 (online), JRC119974.

Velasquez, M., Andre, C., Shanks, T., & Meyer, M. (1990). *Justice and Fairness*. https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/justice-and-fairness/.

Wajcman, J. (2009). Feminist theories of technology. *Cambridge Journal of Economics*, *34*(1), 143-152. doi: 10.1093/cje/ben057.

Wallace, H., Pollack, M. A., Roederer-Rynning, C., & Young, A. R. (Eds.). (2020). *Policy-making in the European Union*. Oxford University Press, USA.

West, M., Kraut, R., & Ei Chew, H. (2019). *I'd blush if I could. Closing gender divides in digital skills through education*. https://en.unesco.org/Id-blush-if-I-could.