European Commission

The EU Mutual Learning Programme in Gender Equality

Artificial Intelligence and Gender Biases in Recruitment and Selection Processes

12-13 November 2020

Comments paper - Portugal

ösb Consulting
Ein Unternehmen der ÖSB Gruppe.

Justice

This publication is supported by the European Union Rights, Equality and Citizenship Programme (2014-2020).

This programme is implemented by the European Commission and shall contribute to the further development of an area where equality and the rights of persons, as enshrined in the Treaty, the Charter and international human rights conventions, are promoted and protected.

For more information see: http://ec.europa.eu/justice/grants1/programmes-2014-2020/rec/index_en.htm

Artificial Intelligence and Gender Biases in Recruitment and Selection Processes. Unpacking the (in)visible agenda in Portugal

Rosalina Pisco Costa

University of Évora

1. Introduction

1.1 Background

Despite the widespread entry of the topics related to artificial intelligence (hereafter AI) in the political discourse and media, in Portugal, specific issues regarding gender bias in recruitment and selection processes seem to be apparently out of the public agenda. Nonetheless, as Portugal's labour market is highly segregated along the classification gender, inequalities persist and there is resistance to change gender bias and stereotypes (Silva, 2018; Torres, 2018). This might mean that, despite being visibly out of the public agenda, invisible gender bias in recruitment and selection processes would certainly be pervasive in AI applications. Therefore, unpacking this issue is currently both a challenging and a relevant topic to address.

The paper by Miriam Kullmann is a concise yet focused analysis of the important debate regarding gender biases in selection and recruitment. Underpinning both its technical and social determinants, Kullmann's discussion paper provides detailed information on AI's potential and applicability in recruitment processes while addressing potential sources of gender and other biases. As pointed, one of the major problems related to the functionality and "fairness" of algorithms has to do with the quality of data. The example of Amazon's machine learning algorithm, that became known for having favoured male candidates because it had been trained on male-dominated resumes over a period of ten years, and, consequently, had outnumbered the female applications[1], is strongly consistent with the idea that "as algorithms often emulate training databases on which they are built, input that is biased will result in outcome that is biased" (Kullmann, 2020, p. 2).

In 2015, in Portugal, eight professions employed more than 50% of women (Torres, 2018). The labour market is characterised by an under representation of women in ICT and STEM and the under representation of men in health, education and welfare

[1] Insights Team, Overcoming AI's Challenge In Hiring: Avoid Human Bias, *Forbes*, Issue 3 (November 29, 2018).

sectors (Silva, 2018). In spite of the formal requirements of equality, it is well-known that gender influences interactions, perceptions and social expectations, as well as power relations, opportunities and, ultimately, the living conditions of men and women. Because curricula constitutes material artefacts of a specific society and culture, including written and visual language, algorithms based on specific corpus will reproduce inequalities and constitute an obstacle in concluding a fair selection and overcoming gender bias and stereotypes, underrepresentation and segregation in education and labour market.

1.2 Approaching the topic of algorithmic discrimination in Portugal

In Portugal, the topic of algorithmic discrimination starts to be discussed. The Portuguese media refers to the general situation of existing gender biases in selection and recruitment processes, using general considerations yet giving no concrete examples. In July 2019, Marta Santos, EY (Ernst & Young) Associate Partner, People Advisory Services wrote that it is necessary to ensure that Artificial Intelligence does not replicate the scenario of gender inequalities for the future and it is necessary to remember that “unlike humans, algorithms cannot consciously combat the prejudices that are instilled in them. Thus, conscious influence must come from us to ensure that technology is a driving force in a world that is moving towards gender parity”[2]. In August 2019, the lawyer Isabel da Silva Mendes signed an opinion article titled “Quo vadis: inequalities, women and artificial intelligence”, in which she claimed for the need of having a woman as Chair of the Portuguese Bar Association[3]. At a time when artificial intelligence is entering the justice, “there is a revolution that is slow to take effect”, she says. In courts, artificial intelligence promises to reduce repetitive and time-consuming tasks, speed up legal proceedings, facilitate the work of judges and lawyers, and make judicial decisions more reliable and fair. In this “new” context, lawyers face an increasing challenge of "framing the development, use of tools and applications of artificial intelligence to guarantee the fundamental rights and freedoms of citizens" and, at the same time, to prepare themselves “to be able to detect and correct errors in algorithms.” To this end, she advocates that “we still have to fight for equality, fight for the elimination of differences and try to equal the value of everyone, without any distinction”, while it is “evident that where there is less diversity, there is less creativity, innovation, profitability and even profit“.

Inside academia, the topic is being increasingly researched. Nunes (2019) discusses Artificial Intelligence in Human Resources Management (HRM), specifically the skills needed by managers to create value in face of digitization and recent developments such as Artificial Intelligence (AI), Big Data and Data Analysis. The work concludes for more value creation through AI and automation, with an ideal combination of

[2] Marta Santos, Será a Inteligência Artificial livre de preconceitos de género?, *Sapo* (July 15, 2019).
[3] Isabel da Silva Mendes, Quo vadis: desigualdades, feminino e inteligência artificial, *Dinheiro Vivo* (August 29, 2019).

technical and social skills, to solve possible gaps and focus on strategic tasks. Individuals in the sample evaluate themselves as having little knowledge to deal with new tools. However, “they do not show motivation to develop new skills, what means that the support from top management to promote mind-set changes, communication and the development of new knowledge will be essential for a balance between soft and hard skills” (Nunes, 2019, p. iii).

A study by Costa (2019) titled “Conversations with ELIZA: on gender and artificial intelligence” is particularly interesting as it unveils the process of assigning digital assistants human-like traits or features, namely a tendency for their feminisation. Following such ideas, the study explored the relationship between gender and artificial intelligence, seeking to understand how gender relates to AI, why femininity seems to be often present in AI and which gender roles or stereotypes are reinforced in this process. An analysis of Alexa, Cortana and Siri was developed, in order to examine how gender is displayed by these entities and complementing this discussion; the project Conversations with ELIZA focuses on the exploration of femininity in AI through the creation of four chatbots integrated into a web-based platform. By the end, “the study aimed to raise awareness on the implications of assigning gender to digital assistants who integrate our daily lives” and “to incite reflection on the social and cultural conventions that inform the conception of artificial intelligence, and how digital assistants end up reflecting these conventions back to their users” (Costa, 2019, p. 5).

Finally in this context, in 2020 a study was developed aimed to build a model to identify the benefits and functionalities of the use of AI, (of the type Artificial Neuronal Network (ANN) multiple layers feedforward), in the process of recruitment and selection (R&S) of professionals for the position of sales technician, as a support system for decision making, to assist the Human Resources (HR) professional (Jatobá, 2020). This quantitative exploratory research used descriptive, inferential statistical techniques and the RNA model. Secondary data (14 technical attributes) were used in 800 CVs of individuals who applied for the position, from 2014 to 2018. Among other interesting conclusion, it was found that “there is a difference in the average grades attributed to the CV by sex and that there is an association between the grade attributed to the CV and the length of experience” (Jatobá, 2020, p. v).

Kullmann summarises five key-domains to take into account in any policy that addresses the question how to counter potential risks of discrimination based on gender through AI: awareness-raising, legal issues, risk assessment, auditing for algorithms, and measure outcomes. As presented hereafter, despite being referred in various contexts, there is no evidence that specific discussion is being raised in Portugal as far as gender biases in selection and recruitment are concerned, as for Portuguese companies, there are at least no publicly available studies about specific evidence of discriminating effects of AI related to that. Consistently, Portugal is never given as an example in the activities included in the AI Guide “Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence. A good practice guide” (Equinet, 2020).

2. Policy debate

2.1 Background

Both at the European and National level, the risk of gender stereotypes is clearly assumed at the policy debate. The Strategic Objective 1 of the Council of Europe Gender Equality Strategy 2018-2023 is to Prevent and Combat Gender Stereotypes and Sexism (Council of Europe, 2018). At the national level, in 2018, Portugal approved The National Strategy for Equality and Non Discrimination (ENIND) 2018-2030 – Portugal + Equal (Portugal + Igual), aiming to fight against inequalities, discrimination, educational and professional segregation based on gender differences and stereotypes. The National Strategy closely follows the 2030 Agenda principle of "No one left behind", as included in the 5th objective of the Sustainable Development Aim – Reach Gender Equality and Empower all Women and Girls.

ENIND sustains in three National Action Plans, each defining strategic and specific objectives related to equality between women and men (IMH); prevention and fight against violence against women, gender and domestic violence (VMVD); and fight against discrimination based on sexual orientation, gender identity and sexual characteristics (OIEC). The strategy is particularly important as it recognises that structural obstacles are creating inequalities for men and women to fully participate in the wider society and to guarantee the conditions for a full and equal participation of women and men in the labour market, including fighting against sexual segregation in professions. For the purposes of this comments paper, it is important to refer to specific axis related to equality between women and men, namely

- Axis 2, which aims to ensure the full participation of women and men in the public and private sphere, promoting equal participation and opportunities in the labour market and professional careers, and
- Axis 3, which refers to a scientific and technological future development, highlighting the aim to promote equal and inclusive participation in the digital process and to use the potential of the technological development to promote equality.

Portugal strategy for artificial intelligence is included in the Portuguese initiative established in 2017, "National Digital Competences Initiative e.2030, Portugal INCoDe.2030", an integrated public policy to enhance and foster digital competences[4]. To address such target, INCoDE.2030 initiative has proposed a wide range of measures involving various governmental areas. These measures should be implemented alongside the private sector, academia and civil society initiatives with similar aims. The measures are structured around five main action lines: inclusion, education, qualification, specialisation and research. As for now, an action plan has been drawn up for consideration by the European Commission and some projects are

[4] https://www.incode2030.gov.pt/en

under development (INCoDE2030, 2018). Specifically, the elaboration of an Action Plan titled *Closing the gender gap in digital technologies. "No one left behind"* (Silva, 2018) is a crucial policy instrument, as it offers a plan aiming to contribute to the increase of the number of girls and women in STEM and ICT fields. Tackling structural constrains and inequalities is particularly important, and this instrument ultimately aims at generating inclusive opportunities for girls and women to participate and to engage in significant education and work activities.

In February 2020, the Minister for Modernisation of the State and Public Administration, Alexandra Leitão, stated that "the use of artificial intelligence in the State sector is fundamental to increase the speed and transparency of processes in public institutions"[5]. Nonetheless, the majority (54%) of public sector entities consider that they have limited Artificial Intelligence (AI) structures and half of them say that they have not yet implemented any, as shows the study by EY and Microsoft ("Artificial Intelligence in the Public Sector", held in 12 European countries, with the participation of 213 organisations, 23 of which were Portuguese[6]). The authors analysed three areas of the public sector - Public Administration, Health and Transport - and concluded that public sector institutions are in the initial phase of implementing AI and that most of the technologies used by public companies are limited AI solutions that allow to increase efficiency and quality of service. The health sector has the highest AI adoption rate. Moreover, in the framework of INCoDe.2030 (Axis 5 - Research), in 2018, the Portuguese agency for funding of research and technology (FCT) launched a research Programme in Data Science and Artificial Intelligence in Public Administration, to support new R&D projects with partnerships between public administration and scientific institutions. Yet, none of the 32 research projects selected for funding in the 2018 and 2019 editions and 4 pilot projects specifically address gender issues (FCT, 2020). In "Gender Research 4 COVID-19", an initiative promoted by FCT aimed at granting special support directed to research projects on the impact of the health emergency caused by COVID-19 on gender inequalities and violence against women and domestic violence, one of the 16 projects selected for funding specifically intended to use AI to study stereotyped language in public communication directed to prevent and combat COVID-19[7].

2.2 Approaching the AI policy debate in Portugal

In Portugal, a wider public debate on the effects of AI regarding equality and discrimination perspectives and the need for regulations and legal instruments to combat algorithmic discrimination is missing.

Recently, specific and episodic discussions arose in the Portuguese media, regarding the legal status and use of algorithms. They were focusing on the question whether

[5] Portuguese Republic, «Recurso à inteligência artificial no setor do Estado é fundamental», *News* (February 12, 2020).
[6] Lusa, Metade do setor público não tem estruturas de inteligência artificial, *TSF* (October 7, 2020).
[7] FCT, Resultado do apoio "Gender Research 4 COVID-19", *FCT Notícias* (July 1, 2020).

algorithms should be made public by law, on the issue of inexistence of regulatory entities to audit them[8], as well as on the (i)legal use algorithms to coordinate prices, even in cases when the decisions are made by algorithms, without human intervention[9].

It is essential to point out that the Decree-Law nr. 47/2019 (2019-04-11)[10] created an early warning mechanism (MAP), regarding the economic and financial situation of companies, which consists of a procedure for providing economic and financial information to members of the management bodies of companies based in Portugal, on an annual basis. This is a mechanism, which supports the decision-making and business management. It is built on quantitative and trend analysis, using economic and financial indicators based on information for each company, complemented by a qualitative assessment of the company's financial and economic situation, resulting from an algorithm operated by IAPMEI - Agency for Competitiveness and Innovation, I. P.

Additionally, in a Notice of 11/2019, the Competition Authority in Portugal (AdC) warned the companies operating in Portugal about the fact that they are responsible for the algorithms they use and that the use of these tools for coordination of prices, or weakening of the competition in the market, is incompatible with the Competition Law[11]. The alert followed the analysis that the AdC has developed in the area of digital ecosystems, big data and algorithms, which gave rise to an Issues Paper (AdC, 2019). The paper refers to the fact that in digital markets, instead of the competition being "just a click away", it may be the exclusion that is "just a click away", since the incumbent platforms can adopt exclusion strategies for competitors based on exploitation of consumer behavioural biases.

3. Recommendations

Based on the above-described context, key recommendations on how to address the potential risk of (gender) discrimination of algorithms in recruitment processes and how to raise awareness of the issue of gender bias in algorithms processes in Portugal include the following:

- at the employers' level: transparency through audit. As many companies operating in the HRM advertise themselves as offering AI services for recruitment and

[8] Nuno David, Devem os algoritmos ser públicos por Lei?, *Público* (November 15, 2018).
[9] Karla Pequenino, Autoridade da Concorrência avisa que algoritmos de preços podem infringir a lei, *Público* (July 2, 2019).
[10] Decreto-Lei n.º 47/2019, Diário da República n.º 72/2019, Série I de 2019-04-11 https://dre.pt/home/-/dre/122074191/details/maximized
[11] AdC, Comunicado 11/2019 AdC alerta que utilização de algoritmos para coordenar os preços no mercado é incompatível com Lei da Concorrência, Autoridade da Concorrência (July 1, 2019).

selection, it will be interesting to follow such services and to audit them for transparency procedures and results;

- at the employees' level: training in designing non-biased curricula. Because stereotypes are rooted in the ways of thinking, they often manifest themselves in the ways we write and, inevitably, take on expression in the curriculum, namely in the adjectives and verbs used to describe past actions, experiences and skills. CV workshops can be developed with an inclusive, non-sexist language in order to create awareness around the importance of non-biased curricula;
- at the developers' level: to promote multidisciplinary and inclusive teams. Teams of developers are usually gender biased. Moreover, developers usually have a background in STEM-related fields (science, technology, engineering and mathematics), and a learning trajectory in higher education, which is far from the issues explored by the social sciences. Working in multidisciplinary and inclusive teams is key in promoting greater awareness for gender bias in real-world applications.

4. References

AdC (2019). *Ecossistemas digitais, Big Data e Algoritmos Issues Paper*. s.l. AdC. file:///C:/Users/rosal/Desktop/Issues%20Paper_%20Ecossistemas%20Digitais%20Big%20Data%20Algoritmos.pdf

Costa, P. (2019). *Conversations with ELIZA: sobre género e inteligência artificial.* Lisboa: Universidade de Lisboa, Faculdade de Belas Artes. http://hdl.handle.net/10451/37542

Council of Europe (2018). *Council of Europe Gender Equality Strategy 2018-2023*. Brussels: Council of Europe.

Equinet. (2020). *Meeting the new challenges to equality and non-discrimination from increased digitisation and the use of Artificial Intelligence. A good practice guide.* https://equineteurope.org/wp-content/uploads/2020/06/ai_guide_digital.pdf

FCT (2020). *Research in Data Science and Artificial Intelligence applied to Public Administration*. . https://www.fct.pt/media/docs/Brochura_ResearchinDataScienceandAIappliedtoPA.pdf

INCoDE 2030 (2018). *Progress Report. Summary of the ongoing and planned activities under the INCoDE.2030 initiative presented at the 2nd conference of the permanent forum for digital competences*. https://www.incode2030.gov.pt/sites/default/files/incode.2030_progress_report_12122018.pdf

Jatobá, M.N. (2020). *Inteligência artificial no recrutamento & seleção: inovação e seus impactos para a gestão de recursos humanos*. Bragança: Instituto Politécnico de Bragança. http://hdl.handle.net/10198/21805

Kullmann, M. (2020). *Gender Biases in Selection and Recruitment Software. Discussion Paper*. The EU Mutual Learning Programme in Gender Equality Artificial Intelligence and Gender Biases in Recruitment and Selection Processes. The Netherlands, 12-13 November 2020.

Nunes, B. (2019). *Inteligência Artificial na GRH: Competências necessárias aos gestores para a criação de valor.* Lisboa: Universidade de Lisboa, ISEG. http://hdl.handle.net/10400.5/18699

Silva, S.M. (2018). *Closing the gender gap in digital technologies. "No one left behind" | 2030 Agenda Action Plan Proposal*. Porto: University of Porto. https://www.incode2030.gov.pt/sites/default/files/closing-the-gender-gap-in-digital-technologies-versao-9-marco.pdf

Torres, A. (Coord.). (2018). *Igualdade de género ao longo da vida: Portugal no contexto europeu* [Lifelong gender equality: Portugal in the European context]. Lisboa: Fundação Francisco Manuel dos Santos. https://www.ffms.pt/FileDownload/12e83bd9-2346-4686-a8ba-94e54a5684b6/igualdade-de-genero-ao-longo-da-vida-resumo-do-estudo