Advisory Committee on Equal Opportunities for Women and Men

Opinion on Artificial Intelligence – opportunities and challenges for gender equality

The Opinion of the Advisory Committee does not necessarily reflect the positions of the Member States and does not bind the Member States

18 March 2020

Contents

1. Introduction

Gender equality is a fundamental right and value of the European Union. However, challenges to gender equality and discrimination on the basis of sex remain a reality today.[1] In this context, the increased presence of Artificial Intelligence (AI) in our everyday lives is central to the transformation of our economy and society. While posing challenges for gender equality and non-discrimination, AI also creates opportunities. Due to the convergence of many digital technologies and the influence of those technologies in our professional and private lives, it is high time to reflect specifically on the interplay between AI and gender equality.[2] The new EU Gender Equality Strategy will not only underpin the political commitment to gender equality and support strategic and systematic work for gender equality in the EU and beyond, but will also address the link between AI and gender equality. The President of the European Commission declared that "Digital technologies, especially Artificial Intelligence (AI), are transforming the world at an unprecedented speed. They have changed how we communicate, live and work. They have changed our societies and our economies." and highlighted that "In my first 100 days in office, I will put forward legislation for a coordinated European approach on the human and ethical implications of Artificial Intelligence."[3] It will build on earlier initiatives by the Commission, such as (but not limited to) the coordinated plan on AI[4], the Ethics guidelines for trustworthy AI[5] and the Commission's report "The future of work? Work of the future".[6]

On 19 February 2020, the European Commission presented the "White Paper on Artificial Intelligence: a European approach to excellence and trust", which sets out some policy objectives[7]. Artificial Intelligence is developing fast and it will change and improve many areas of our lives. At the same time, Artificial Intelligence (AI) entails a number of potential risks, such as opaque decision-making, gender-based or other kinds of discrimination. The EU's approach to Artificial Intelligence, based on trust and excellence, will give citizens the confidence to embrace these technologies while encouraging businesses to develop them[8]. The AI White Paper forms part of the broader policy objectives of the Commission and needs to be seen in the framework of the ideas and actions for a digital transformation put forward on 19 February 2020[9]. It presents a European society powered by digital solutions that put people first, opens up new opportunities for businesses, and boosts the development of trustworthy technology to foster an open and democratic society and a vibrant and sustainable economy. The European data strategy and the policy options to ensure the human-centric development of Artificial Intelligence (AI) are the first steps towards achieving these goals.

Furthermore, the December 2019 UN Secretary General's review and appraisal of the implementation of the Beijing Declaration and Platform for Action[10] addresses the digital revolution as one of the major shifts that have taken place since 1995. It recognizes that technological

---

[1] https://rm.coe.int/cm-rec-2019-1-on-preventing-and-combating-sexism/168094d894

[2] For a general overview, see for example: Kaplan, Jerry. Artificial Intelligence. 2016. What Everyone Needs to Know;

[3] https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

[4] https://ec.europa.eu/digital-single-market/en/news/coordinated-plan-artificial-intelligence

[5] https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

[6] https://ec.europa.eu/epsc/publications/other-publications/future-work-work-future_en

[7] https://ec.europa.eu/commission/presscorner/detail/en/FS_20_282

[8] See the factsheet on the topic: https://ec.europa.eu/commission/presscorner/detail/en/FS_20_282

[9] https://ec.europa.eu/commission/presscorner/detail/en/IP_20_273

[10] Review and appraisal of the implementation of the Beijing Declaration and Platform for Action and the outcomes of the twenty-third special session of the General Assembly Report of the Secretary-General, CSW64: https://undocs.org/E/CN.6/2020/3

advances, including AI, automation and robotics, are having profound effects on gender equality and women's rights in all spheres of life.

AI offers significant opportunities to improve our lives. However, considering that every decision a human being makes in life is prone to biases, it has been widely acknowledged that biases play a major role in the area of gender equality and non-discrimination specifically.[11] Equally, when AI is used, it can mirror human biases and even reinforce them, either because the bias in the training data is programmed into the algorithm or because biased people simply create biased algorithms. Therefore, one of the main goals is to identify potential gender biases of AI already present in technologies and in consumer products, which might become mainstreamed. Recent books highlight the issue in relation to gender equality and even speak of "exposing data bias in a world designed for men"[12]. In a nutshell, the challenge lies in identifying potential risks to achieving gender equality and exploring corresponding solutions. The challenge lies not so much in the occurrence of completely new forms of discrimination or problems for gender equality, but rather in the fact that the decision making process is automated, hidden and placed further away from citizens, who benefit from or are impacted by those decisions.

It is important to have a human centric approach. Algorithms can effectively support people in making decisions and indicate the possibility of avoiding stereotypical thinking. We need to ensure that human beings with relevant training on biases and non-discrimination are involved in every stage of development and operation of algorithms.

In the framework of the ongoing work of the European Commission on AI[13], including the recent publication of the "White Paper on Artificial Intelligence: a European approach to excellence and trust"[14] and the public consultation from 19 February until 31 May 2020[15] , this opinion will explore the impact of AI on gender equality with an emphasis on specific areas (education, STEM, recruitment, data and monitoring and accountability). The opinion will therefore look at the challenges and opportunities posed by AI, and show concrete examples of the application of this technology and highlight existing best practices. One of the aims of the opinion besides pointing to certain avenues for addressing some of the identified problems is also to stimulate the debate on the impacts of AI systems on gender equality and framework of protection against discrimination.

Both AI and the fight to achieve gender equality remain areas where a lot of work remains to be done.[16] Some best practices on how to deal with any problems and opportunities may become outdated and new ones may emerge over time. Thus, member states should exchange these best practices on a regular basis to inspire others even if some practices might be tailor-made for a given context in a given member state.

---

[11] See for example in general, Bohnet, Iris. *What Works : Gender Equality by Design.* 2016

[12] See Criado-Perez, Caroline. Invisible Women : Exposing Data Bias in a World Designed for Men. 2019.

[13] See notably the inclusion in the European Commission's work programme: "This will cover the best possible use of the potential of digital data and the development and uptake of artificial intelligence that respects our European values and fundamental rights.", https://ec.europa.eu/commission/presscorner/detail/en/IP_20_124 ; https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe

[14] https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[15] https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-european-approach-excellence-and-trust

[16] See for example, the most recent annual report of the European Commission 2019, the report on the Colloquium of fundamental rights or EIGE's gender equality index measuring the progress on gender equality in different fields. Latest report from October 2019.

This opinion deals mostly with the questions of how AI might address the risk of perpetuating gender inequalities and risk for discrimination, how to mitigate these risks with policy actions, and how AI can contribute to reducing gender inequalities. Bearing in mind that the problems surrounding AI and gender equality are numerous, the focus of this opinion is on the four areas highlighted below. Aside from these four areas, there are other issues related to AI, such as legal responsibility for AI systems, which are also of great importance but cannot all be addressed in this opinion.

2. Context

Definition[17]

> *"Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."*

What is the impact of AI on gender equality?

We need to connect the digital space, including AI, to the real world as a continuum and a mirror. Therefore, tackling gender inequalities in the real world is crucial to avoid perpetuation of gender stereotypes and inequalities in the digital world. If the data collected and used for the development of AI is gender biased[18], the outcome/results will be biased as well[19]. In addition, the algorithm itself, which has been coded by humans, may reflect biases of the developers. Rather than being neutral, human assumptions and the use of selected data lead to biased decisions[20].

Due to these biases, AI can affect gender equality in a number of fields. For example, in the field of women's socio-economic independence and leadership, when AI systems are used to select candidates for possible vacancies, some algorithms may favour or even only select candidates of one sex, for example men, due to a historic imbalance in hiring practices reflected in the data set used for the AI.[21] This could also have an effect in the healthcare sector. At present, there is a gender data gap when it comes to connecting health issues and diagnoses to symptoms. The symptoms for men vary from those of women. The AI-systems, such as symptomatic machines, rely on the current data and may as such be biased and lead to misdiagnosis.

Some publicly available AI applications are predominantly gendered. Digital assistants have been found to remain impassive in the face of sexist insults or even consider sexist remarks as compliments ("I'd blush if I could" as a response to "You're a slut"). This may seem harmless, but in

---

[17] https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

18 Gender bias can be understood as "Prejudiced actions or thoughts based on the gender-based perception that women are not equal to men in rights and dignity", see EIGE Glossary, https://eige.europa.eu/thesaurus/terms/1155

[19] See for example on the gender data gap in general, Kraft-Buchman, C. and Arian, R. (2019). The Deadly Data Gap: Gender and Data. [online] Geneva: Women at the Table. Available at: http://bit.ly/DeadlyDataGenderGap

[20] See for example Meredith Broussard, Artificial UnIntelligence, MIT 2019 who explains that AI is not always getting it right just as human beings.

[21] https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G ; https://hbr.org/2019/05/addressing-the-biases-plaguing-algorithms

reality, it perpetuates sexist behaviours.[22] Here it is important to identify patterns in the algorithm that lead to such sexist behaviours and to find solutions to stop this, for example by the development of counter algorithms.

Harmful behaviour, sexist remarks, and violence against women, such as instances of (sexist) hate speech or the use of lifelike "deepfake" videos and images"[23] also occur online, disproportionately affecting young women and girls, women journalists, politicians, public figures and women human rights defenders[24]. As a result of such online hate speech/abuse, women are more likely to restrict themselves from engaging online[25][26]. AI can be used to detect harmful online content against women and girls, but needs to be monitored and evaluated by humans to understand the context when analysing harmful online content. Here AI can play a role and assist in order to create an online world that is more respectful and embraces a gender equal society[27].

3. Recommendations

Before making some specific recommendations by area, this opinion would like to recommend in general:
- Reflect and analyse at the EU[28] and Member State level whether the existing legislative and institutional framework is equipped to deal with arising legal questions related to AI and gender equality and non-discrimination or whether there is any need for reform within the general framework for promotion and protection of human rights.[29]
- Guidance in the form of mission statements, charters and/or guidelines addressed to developers of AI that could state in essence for example that "AI should be neutral, avoiding any bias, in particular regarding gender and all other grounds of discrimination under current EU law."[30]
- Continuous monitoring of new and existing algorithms can be beneficial to detecting discrimination.
- Explore at the Member State level, potentially with the support of national equality bodies, policy measures, such as, sharing of good practices, training and education to raise citizens' awareness and provide information on how algorithms operate and how to detect biases and their impact in daily lives from a gender perspective.
- Promote the participation of diverse groups of citizens in the development of AI, and in this way, enhance the user-friendliness, transparency and non-discriminatory use of AI.

---

[22] https://en.unesco.org/Id-blush-if-I-could
23https://www.washingtonpost.com/technology/2018/12/30/fake-porn-videos-are-being-weaponized-harass-humiliate-women-everybody-is-potential-target/
[24] https://rm.coe.int/prems-055519-gbr-2573-cmrec-2019-1-web-a5/168093e08c
[25]https://www.humanrights.dk/sites/humanrights.dk/files/media/dokumenter/udgivelser/equal_treatment_2017/hate_speech_in_the_public_online_debate_eng_2017.pdf
[26] https://eige.europa.eu/publications/gender-equality-and-youth-opportunities-and-risks-digitalisation
[27] See in general, Kraft-Buchman, C. and Arian, R. (2019). Affirmative Action for Algorithms - Artificial Intelligence, Automated Decision-Making & Gender - We Shape Our Tools, and Thereafter our Tools Shape Us. [online] Geneva: Women At the Table. Available at: http://bit.ly/AI_ADM_Gender
[28] *The new EU Gender Equality Strategy 2020-2025 will underpin the political commitment to gender equality and support strategic and systematic work for gender equality in the EU and beyond, which will also address the link between AI on gender equality.*
[29] Unboxing Artificial Intelligence: 10 steps to protect Human Rights" (Page 9, third paragraph*). Member states should establish a legislative framework for independent and effective oversight over the human rights compliance of the development, deployment and use of AI systems by public authorities and private entities."*. The High-Level Expert Group on AI (AI HLEG) formulated concrete recommendations to the European institutions and Member States in its Policy and Investment Recommendations for Trustworthy AI. The first of them is to empower and protect humans and society and, among others, states that ensuring trustworthy AI requires an appropriate governance and regulatory framework fully aligned with fundamental rights.
[30] AI, The future of work, work of the future, p. 131

- Capacity-building and the provision of adequate resources[31] for equality bodies is fundamental to ensure that they are better equipped with knowledge and tools to detect and address discrimination in the context of AI. [32]
- Strengthen and disseminate the work carried out by civil society and groups that defend and promote equality in AI.

1. Gender stereotypes in education

Gender stereotypes hinder peoples' freedom to develop themselves to their full potential. The use of AI can aggravate this issue by building systems based on gender biases and stereotypes or by developing algorithms that mirror these stereotypes, which can result in discriminatory practices. The digital systems used by educational institutions in the public and private spheres must not discriminate on the basis of sex or other grounds and have to promote gender equality in accordance with EU minimum standards on non-discrimination. In addition, the rapid development of AI creates an increased need for lifelong learning.

Therefore, the committee recommends member states to:
- Raise awareness on the issue of gender stereotypes in education and encourage more women and girls to take up STEM education, possibly also through the training of teachers where necessary.
- Consider giving incentives to non-governmental parties to help schools encourage female students to choose STEM-related fields by i.e. raising awareness on career opportunities and highlighting role models.
- Draw attention to solutions in the field regarding the broader struggle for equality between women and men. Training on stereotypes and their consequences should be incorporated in the curricula.
- Based on the GDPR, encourage that ethical councils, and procurement officials in educational institutions, which have the mandate to monitor privacy and ethical aspects of the use of data in education, also pay attention to gender biases, stereotypes and norms in the real world and the digital sphere.
- Carry out further studies on the impact of AI on gender equality and non-discrimination, and look into ways to ensure a multi-disciplinary approach to AI research where humanities, social sciences, STEM-fields/technological studies and gender research are incorporated with the aim of preventing gender biases finding their way into algorithms.
- Encourage the use of EU programmes such as Horizon Europe, Digital Europe and Erasmus+ to promote multi-disciplinary research, pilot projects, experiments, development of tools including training, for the identification of gender biases in AI, and awareness raising campaigns for the general public.
- Implement an intersectional perspective concerning the interplay of AI and gender.

2. Science, Technology, Engineering and Mathematics (STEM) sector

Despite the demand for ICT and digital professionals with technical backgrounds, as well as the positive growing trend of the sector as a whole, the gender gap in the technology industry persists.

---

[31] See COMMISSION RECOMMENDATION of 22.6.2018 on standards for equality bodies, https://ec.europa.eu/info/sites/info/files/2_en_act_part1_v4.pdf
[32] The European Network of Equality Bodies (Equinet) will publish a report (by Robin Allen and Dee Masters) on "Regulating AI: the new role for Equality Bodies – Meeting the new challenges to equality and non-discrimination from increased digitalisation and the use of AI", expected early 2020."

When considering advanced digital skills, girls under 24 surpass their male counterparts, while in the other age groups a gender gap negatively affecting women still persists.[33] There are differences among the EU countries regarding the gender gap of digital skills. Despite having similar levels of basic digital skills, women question their own skills more often than men do. This data is consistent with existing literature that shows that women tend to undermine their own capabilities and skills to a greater extent than men.[34]

According to the Global Gender Gap Report 2020 by the World Economic Forum, female workers make up an estimated 26% of workers in data and AI roles globally[35]. Only 17% of ICT specialists in the EU are women[36] even though women-led entrepreneurial start-ups are more likely to be successful. Globally, women make up 12% of AI researchers.[37] They hold 20% of technical positions within major machine-learning companies.[38] They represent 18% of the authors attending leading AI conferences.[39] Moreover, only three (2006, 2008, 2012) of the 68 Turing Award winners, the equivalent of the Nobel Prize for Computer Science, were women. The subset of the population that currently shapes and designs AI systems is thus very narrow, and with a clear underrepresentation of women.

One of the explaining factors is that the "geek culture" that prevails in the AI industry, characterized by a male predominance, might act to some extent as a social censorship of women's access to that industry. This predominantly male environment might facilitate or contribute to a mentality that generates or accepts harassment, thus pushing women to abandon their careers in new technologies. Thus, 56% of female technical staff in the tech industry quit their career at the mid-level point, twice the resignation rate of men. Finally, science faculties subtly favour male CVs, thus practicing gender-biased recruitment decisions.

Despite its positive potential in working life, automation and digitalisation may further compromise women's position in the labour market, because there is still a relatively low number of women in ICT education and the ICT professions, or because of gender bias consciously or unconsciously embedded in algorithms and artificial intelligence, thus increasing the risk of gender digital divide. Here the STEM sector could be an avenue to tackle the gender pay gap by attracting more women in AI-related professions[40].

The committee recommends to:
- Develop, from a young age, positive action measures to attract more women to STEM fields and particularly to AI. Further awareness of the flexibility, working conditions and possibilities for reconciliation of work and family life in this sector is required targeting girls and women.

---

[33] EC Report (Study on Women in Digital Age, see https://ec.europa.eu/digital-single-market/en/news/increase-gender-gap-digital-sector-study-women-digital-age) which shows that having a third-level education (university degree) in ICT-related studies increases the probability of employment for men by 2-3%; for women with the same degree however, *the probability of being employed decreases by 1-2%[33]*

[34] Study on Women in Digital Age, see https://ec.europa.eu/digital-single-market/en/news/increase-gender-gap-digital-sector-study-women-digital-age

[35] Global Gender Gap Report 2020, World Economic Forum, http://www3.weforum.org/docs/WEF_GGGR_2020.pdf

[36] https://eige.europa.eu/publications/work-life-balance/eu-policies-on-work-life-balance/women-in-ict

[37] Tom Simonite, "AI Is the Future—But Where Are the Women?," *Wired*, August 17, 2018.

[38] https://www.vox.com/2017/8/7/16108122/major-tech-companies-silicon-valley-diversity-women-tech-engineer

[39] https://www.elementai.com/news/2019/2019-global-ai-talent-report

[40] AI, The future of work, work of the future, p. 103

- Incorporate humanities, ethics and gender equality research into STEM fields of study and into AI vocational and academic programmes to raise awareness, and more importantly to guarantee a multidisciplinary approach.
- Ensure regular training is provided to experts and stakeholders in the AI industry on gender equality, gender mainstreaming and stereotypes to ensure that the lessons learned remain relevant in the field. Such courses should also raise awareness on the importance of equality, fairness, accountability (where possible and relevant through auditing), transparency and robustness of an AI system, alongside other elements.
- Address sexism in the particularly male dominated work environment of AI and programming, and invite Member States, stakeholders and employers to reflect on how to prevent sexism and introduce tougher consequences against perpetrators.
- Ensure regular training as well as the development of positive algorithms and algorithms capable of addressing existing digital gender biases and prejudice.
- Raise awareness among stakeholders on the need to have a more gender balanced AI and technology sector by attracting and incentivising more women to work in this field at all levels; thus making use of the full human potential, i.e. by organising expert meetings, transparency of policy and quantitative research into the needs of the labour market.
- Raise awareness, among men in the technology and AI sector, on gender equality and the benefits of equal sharing of caring responsibilities and a world without sexual harassment and violence against women. Involving men in this manner is key to bring about positive change.

3. Recruitment

The Commission's report on "AI - The future of work? Work of the future!" recalls that "the increased uptake of AI in the economy also has important implications for non-discrimination and gender equality [and...] would have to be constantly monitored for infringements in hiring"[41].

The number of companies and administrations using AI in recruitment processes[42] or for other job-related situations, such as unemployment benefits and promotion is constantly rising. The use of AI has notably also attracted national courts' attention, which are exploring to set limits to the use of AI in case human and fundamental rights are violated. [43] Companies use either complex AI systems or AI assisted procedures to review job applicants' CVs, résumés and video interviews (including speech or facial recognition[44] to profile candidates), preselect potential candidates,[45] and to offer a point of contact. Their motive is to tackle the central challenge of human resources of matching applicants to vacancies, as well as enhancing efficiency in the hiring process[46]. It is important to ensure that national and European policies, which encourage the use of AI for automated hiring decisions, do not

---

[41] AI, The future of work, work of the future, p. 102

[42] Equality implications of government decision-making and artificial intelligence, Cloisters, see https://www.cloisters.com/equality-implications-of-government-decision-making-and-artificial-intelligence/
https://verfassungsblog.de/koennen-algorithmen-diskriminieren/#primary_menu_sandwich

[43] Such as for example in the Netherlands, where a court decision showed concerns in relation to human rights violations in the automated surveillance system for detecting welfare fraud: https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules

[44] https://dzone.com/articles/smart-interview-a-new-way-for-recruiting-candidate

[45] One recent example is Amazon that stopped a particular AI because it showed discriminatory behaviour, see: Amazon scraps secret AI recruiting tool that showed bias against women, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

[46] enhancing efficiency is only one aspect; most AI providers advertise through saving costs and manpower, leave time for the essential part of the process (interview) and find more fitting candidates and avoid high costs due to mismatches. Communication AI providers promise to cover the 'black hole' in the application process and others even promise to help you diversify your team.

result in the infringement of non-discrimination laws and/or create inequalities of any another nature. According to research, there is an indication that around 40 percent of the HR functions in international companies are currently using AI-applications, which shows that the technology is widely used and is of great importance to the labour market and recruitment of workers[47]. Many recruitment technologies using AI use a scoring system, evaluating and preselecting applicants based on their attributes in the context of the position's requirements.[48] These systems are based on the objective to make applicants' knowledge, skills and attributes measurable, facilitating their comparability. Amongst others[49], these technologies entail the risk of algorithmic bias and discriminatory practices: bias and discrimination in the choice of input variables, bias and discrimination in the outcome measures and bias and discrimination in the construction of the training procedure. It gives the candidates different scores similar to those used for products online or hotel ratings.

The problem and risk for gender equality is not only related to the fact that these AI tools are used, but also that due to the "black box-nature", decisions often happen in the "hidden" and are difficult to understand for potential victims of discrimination. Thus, problems in terms of gender equality might remain unnoticed or be very difficult to prove. Therefore, it is important to reflect on the transparency and accountability of AI, as well as on ways to ensure that accessibility is in line with constitutional values, such as privacy and intellectual property, to ensure a fair process is possible for citizens.

The committee recommends to:
- Raise awareness in the HR sector about the use of AI tools and their functioning and potential dangers in order to prevent gender bias and discriminatory practices. This can be done e.g. by providing articles, books, (tailor made) guidelines and training and providing (online) resources and information about up-to-date AI research[50].
- Ensure transparency in the HR sector concerning the criteria used in the recruitment process.
- Organize awareness-raising campaigns using concrete and representative examples addressed to media professionals as well as to technical and scientific fields to incorporate the gender perspective into AI.
- Form ethical guidelines and certification for the use of data in recruitment and other job-related situations based on the EU Legal and Ethical Guidelines for trustworthy AI;
- Increase the visibility of expert women in the field of AI e.g. by creating appropriate platforms[51], thinking about the portrayal of women in the media, by highlighting role models where possible and other types of campaigns;
- Consider that more female senior managers can be instrumental for companies to develop more inclusive recruitment policies and prevent application of gender-biased AI and discriminatory practices in recruitment processes.

---

[47] https://www.pwc.nl/nl/assets/documents/artificial-intelligence-in-hr-a-no-brainer.pdf
[48] Such as Textkernel5, which pre-sorts and pre-filters written applications (https://www.textkernel.com/hr-software/semantic-matching/#features-link) or Prescreen10 which creates job profiles, career pages, distributes advertisements and filters and manages incoming applications (https://prescreen.io/de/active-recruiting/).
[49] the lack of gender diversity in the AI workforce, compare: https://www.nesta.org.uk/report/gender-diversity-ai/
[50] Very simple advice for employers has been published in a recent article of the Harvard Business Review, https://hbr.org/2017/06/7-practical-ways-to-reduce-bias-in-your-hiring-process
[51] There is already a platform in the form of a website that lists women in the AI sector with different specialisations, that could for example serve as contacting them for conferences, expert advice etc, see https://www.womeninai.co/ , https://www.ibm.com/watson/women-leaders-in-ai/2019-list

4. **Data**

High quality and representative data is the basis for working towards bias-free and non-discriminatory AI. Data must represent social diversity and be comparable, which can be problematic due to potential gender data gaps[52]. Moreover, it is necessary to use data disaggregated by gender/sex and other characteristics such as location, age, class, disability, race ethnicity etc. Since AI is likely to mirror gender biases, it is important to be able to detect possible correlations that do not directly discriminate against women, but indirectly have the same effect. [53] Should the risk of potential discrimination be too high, data that is not strictly necessary for the algorithm to function properly, could be avoided by not including specific references to sex. However, when using self-learning algorithms, which work with correlations, indirect discrimination could occur because of certain data usage; women or men could be disproportionately affected because of behaviours associated with both genders. If, for example, historical data is used in recruitment and in a given company there were only or mostly men working, it will be difficult to exclude certain data because then no data would be available. In such cases, it would probably be too challenging to restrict the data being used, but rather impose positive obligations to ensure non-discriminatory outcomes.

The committee recommends to:
- Ensure transparency in the use of data for AI applications.
- Endorse and support the EESC[54] calls for a code of ethics for the development, application and use of AI to ensure that the principles of a human centric approach, human dignity, integrity, freedom, privacy and cultural gender diversity, as well as fundamental human rights are safeguarded throughout.
- Research possibilities of non-discrimination by design: the idea that algorithms themselves can incorporate equal treatment laws into their decision-making processes to ensure that no direct discrimination is possible and that no decision is based directly on gender/sex. Scientists are still working out how this could be done, but a potential starting point could be the obligation for human intervention as soon as a decision is based on a data point that correlates significantly with gender/sex.
- Work together with national equality bodies and Equinet to look into ways of monitoring the quality of the data used; that is, the use of unbiased data to create algorithms.
- Work with national and international statistical bodies to develop data collection standards and to increase the production of quality gender data and address gender data gaps.

5. **Monitoring and accountability**

The AI industry is constantly evolving. In order for it to adhere to the ethical guidelines for trustworthy AI and European and (inter)national human rights, monitoring is of vital importance. To ensure compliance with both national and European human rights, (such as the principle of non-discrimination on the basis of sex) monitoring should play an important role. Especially in times where the implications of AI are not fully understood and analysed, it is important to monitor new

---

[52] Kraft-Buchman, C. and Arian, R. (2019). The Deadly Data Gap: Gender and Data. [online] Geneva: Women at the Table. Available at: http://bit.ly/DeadlyDataGenderGap

[53] Seck and Maskey. 2019. Use data to take guesswork out of policymaking. UN Women. https://data.unwomen.org/features/use-data-take-guesswork-out-policymaking

[2] See for example Caroline Criado-Perez book, "Invisible Women: Data Bias in a World Designed for Men"

[54] https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence

developments both positive (how to use it as a means to progress in the area of gender equality) and negative (when and where AI-systems discriminate or amplify existing gender inequalities).

It is also important to note that discrimination or amplification of inequalities may find its basis in all stages of the lifetime cycle of AI systems. For example, a bias may be present in the training data itself, it may be coded into the algorithm, or it may develop because of new input data. Therefore, monitoring should apply to all stages of the lifecycle of AI-systems.

An interesting set of ideas can be found in the "Villani plan" drafted in France. In terms of institutional reform, the creation of an Advisory Committee on Ethics for Digital Technologies and Artificial Intelligence is one of the commitments of this plan. It will (amongst other tasks) carry out audits of algorithms and databases[55]. In fact, a panel of experts can use AI to detect sexism. Those experts would then have the possibility, for example, to test the equity of a recruitment site in relation to potential gender biases or discriminatory behaviour by submitting a large number of CVs of women and men with the same skills and career path, and then observe the requests for interviews granted, or the proposed average wages[56].

The Committee recommends to:
- In line with the recommendations of the expert report on work and AI by the European Commission, to study and analyse possible mechanisms at EU or national level that could be utilised for supervising AI related problems in relation to gender equality and discrimination. This could be done in collaboration with equality bodies.
- Assess whether the AI industry should carry out gender impact assessments during all stages, to analyse whether specific groups of women and men are directly or indirectly affected during all stages of the algorithmic decision-making.
- Explore whether governments should look into how systems, which are already used, could be tested for (indirect) discriminatory effects. Where possible, transparency of the used data or the assumptions in the algorithm could help during this monitoring phase. Also, national governments could look into ways of strengthening their laws in this area as well as supporting existing investigating and enforcement agencies, including equality bodies, to deal with AI in their respective fields;
- National equality bodies and other actors could explore ways to apply existing legislation in a way that could help create suitable jurisprudence on AI;
- Monitor AI systems, in compliance with national systems of monitoring and prosecution, to encourage that they meet the European Commission's Trustworthy AI principles and requirements[57];
- Require regular reporting and evaluation of developments in this field by regulators, developers and operators, in compliance with national systems of monitoring and prosecution;
- Ensure that agencies and organisations, particularly equality bodies, tasked with monitoring AI-compliance are empowered to the extent that they are able to look into the possible

---

[55] Cédric Villani, *Donner un sens à l'intelligence artificielle : Pour une stratégie nationale et européenne* (2018)
[56] In Belgium, des tests pratiques et les *mystery calls* have been incorporated into anti-discrimination law some time ago, see https://www.dekamer.be/FLWB/PDF/54/1132/54K1132001.pdf ;
A recent colloquium was organised in France, *Les tests de discrimination: pratiques et perspectives, Actes du Colloque du 11 décembre 2009* : ec.europa.eu%2Fmigrant-integration%2Findex.cfm%3Faction%3Dmedia.download%26uuid%3D2A05A23C-01D7-7B44-418ABA229B10F001&usg=AOvVaw0vdhefEZVnLbMX-bsSBxdM
[57] https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

infringements of (gender) equality law both from a legal point of view, but also the *de facto* understanding of how discrimination in this field works.
- In light of the new EU Gender Equality Strategy, use gender mainstreaming as a tool and strategy to achieve gender equality in all fields.