REPUBLIC OF TÜRKİYE

ANKARA UNIVERSITY

INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF PUBLIC LAW

GENDER BIAS OF ARTIFICIAL INTELLIGENCE

Master's Thesis

Öznur HANCI

Ankara, 2024

REPUBLIC OF TÜRKİYE

ANKARA UNIVERSITY

INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF PUBLIC LAW

GENDER BIAS OF ARTIFICIAL INTELLIGENCE

Master's Thesis

Öznur HANCI

Thesis Supervisor

Prof. Dr. Gülriz UYGUR

Ankara, 2024

REPUBLIC OF TÜRKİYE

ANKARA UNIVERSITY

INSTITUTE OF SOCIAL SCIENCES

DEPARTMENT OF PUBLIC LAW

GENDER BIAS OF ARTIFICIAL INTELLIGENCE

Master's Thesis

Öznur HANCI

Thesis Supervisor

Prof. Dr. Gülriz UYGUR

THESIS JURY

Name and Surname
1-Prof. Dr. Gülriz UYGUR
2-Assoc. Prof. Nadire ÖZDEMİR
3-Asst. Prof. Olcay KARACAN

Thesis Defense Date: 16.04.2024

ACKNOWLEDGEMENTS

I would like to extend my sincere thanks to my thesis supervisor, Prof. Dr. Gülriz UYGUR for her encouragement regarding this novel and challenging subject. Her academic vision and guidance have been a huge contribution to the very core of this thesis.

This endeavor would not have been possible without all the special people around me who inspired me throughout my journey. The understanding and wit of my friends kept me motivated along the way.

Lastly, I am eternally grateful for the unwavering support and trust of my beloved family.

To the great minds and kind souls burdened with the noble task of creating a better, fairer world for all.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **I**

TABLE OF CONTENTS ..... **II**

ABBREVIATIONS ..... **V**

INTRODUCTION **1**

SECTION ONE: ADDRESSING THE GENDER BIAS OF AI **5**

I. OVERVIEW OF GENDER BIAS IN AI APPLICATIONS **5**

A. THE MEANING OF BIAS 5

1. Epistemic Injustice and Gender Bias 7

2. Social Constructs and Gender Bias 15

B. THE GENDER BIAS: WOMEN'S DISADVANTAGE 17

C. WOMEN AND OTHER MARGINALIZED GROUPS 20

D. BASIC PRINCIPLES OF ARTIFICIAL INTELLIGENCE 23

E. DISCUSSION POINTS AROUND MACHINE LEARNING 26

1. The Problem of Consensus and Diagnosis 26

2. Insufficiency of Legal Protections in Terms of Problem Resolution 27

3. The Assumption that There Will Be No Bias or Discrimination in Scientific Fields 29

II. THE EMERGENCE OF BIASES IN TECHNOLOGY **33**

A. DISPROPORTIONATE GENDER PARTICIPATION IN THE INDUSTRY 33

1. Structural Injustice 34

2. Male Domination in Computer Science 38

B. THE DATA PROBLEM 41

1. Lack of Literacy and Comprehension 41
2. False or Biased Data 43
a. The Internet and User-Generated Data 44
b. Data from Certain Processes 48

SECTION TWO: IMPLICATIONS OF THE GENDER BIAS OF AI 51

I. CONSEQUENCES FOR SOCIETY AT LARGE 51

A. ECONOMIC IMPLICATIONS 51
1. Gendered Labor Market Effects 52
a. Impact on Employment Opportunities for Women 53
b. Wage Disparities and Gender Pay Gap 57
2. The AI Impact: Marketing and Boundaries 59
a. Effect on Marketing and Consumer Behavior 59
b. Limitations in Product Development and Innovation 60
B. SOCILOGICAL IMPLICATIONS 62
1. Social Norms and Attitudes 62
a. Reinforcement of Societal Biases and Prejudices 62
b. Influences on Gender Roles and Perceptions 64
2. Exacerbating Inequality and Exclusion 69
a. Amplification of Biases in Decision-Making Processes 69
b. Differential Access to AI-Driven Services and Technologies 71

II. QUESTIONS WITHIN THE LEGAL LANDSCAPE 73

A. ETHICAL USE OF AI SYSTEMS 74
B. ACCOUNTABILITY FOR AI-GENERATED OUTCOMES 77
C. DEVELOPING REGULATIONS SPECIFICALLY FOR AI 79

CONCLUSION 83

REFERENCES 87

ABSTRACT 113

ÖZET 114

ABBREVIATIONS

AI : Artificial Intelligence

CAGR : Compound Annual Growth Rate

ET AL. : Et Alia (and others)

EU : European Union

GDP : Gross Domestic Product

NGO : Non-Governmental Organization

p. : Page

pp. : Pages

R&D : Research and Development

STEM : Science, Technology, Engineering and Math

USD : United States Dollars

INTRODUCTION

"Globally, studies show that women in the labour force are paid less, hold fewer senior positions and participate less in science, technology, engineering and mathematics (STEM) fields."[1]

It is no revelation that women are underrepresented in the science world. They face biases at every aspect of lives, in their homes, at their work, even at their own safest places. They know the glass ceiling exists for them, so they put in extra effort to overcome challenges their male peers do not even know exist.[2] Yet, even though our civilization and science advance, these advancements bring extra complications for gender inequality. Women face newer, bolder faces of biases. What is considered objective or scientific becomes another tool that knowingly or unknowingly applies biases on women. This new, shiny, and exciting tool now is artificial intelligence.

Recently, artificial intelligence has become an important part of our daily lives that influences various aspects of our existence. Algorithms decide what we see online, adjust advertisements to our personalities, and even assess our resumes for job opportunities. However, the ever-increasing use of AI applications has triggered some concerns regarding

---

[1] UNESCO, "Does Artificial Intelligence advance gender equality?", 8 Mar 2023, https://www.unesco.org/en/articles/does-artificial-intelligence-advance-gender-equality, Accessed 1 Dec 2023.

[2] İbrahim Yıldız et al., "A Case Study On Glass Ceiling Syndrome of Female Employees in the Information Technology Sector", *Atatürk İletişim Dergisi,* vol. 16, 2018, pp. 99-112, https://doi.org/10.32952/atauniiletisim.489417, Accessed 12 Dec 2023.

their ethical use.[3] The reliance on AI systems raises deep questions about fairness and equality, especially as instances demonstrate how AI fails to produce neutral outcomes or, worse, exhibits discriminatory behavior. The biases of AI go hand in hand with human biases due to the vast amount of data AI models possess[4] and the seemingly objective approach they are thought to possess. These concerns have sparked the urgent need to explore and address the complex issue of gender bias within artificial intelligence.

Several instances have underlined the inherent biases in AI systems, especially those that concern gender.[5] AI applications, predominantly powered by machine learning algorithms, are designed to recognize patterns, and make decisions based on vast datasets.[6] However, these algorithms often inherit societal biases present in the data they are trained on. Consequently, this existence of bias within AI systems has led to instances where gender-based discrimination is observed; including but not limited to hiring processes, loan approvals, or even translation tools that assign different gender pronouns to different job occupations.

---

[3] UNESCO, "Artificial Intelligence: examples of ethical dilemmas", 21 Apr 2023, https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases, Accessed 30 Nov 2023.

[4] LANGLOIS, "Ethics in artificial intelligence: Discriminatory biases", 12 Sep 2023, https://langlois.ca/ethics-in-artificial-intelligence-discriminatory-biases/#:~:text=Discriminatory%20AI%20has%20many%20underlying%20causes&text=The%20biases%20found%20in%20AI,not%20neutral%2C%20fair%20or%20equitable, Accessed 1 Dec 2023.

[5] International Women's Day, "Gender and AI: Addressing bias in artificial intelligence", https://www.internationalwomensday.com/Missions/14458/Gender-and-AI-Addressing-bias-in-artificial-intelligence, Accessed 1 Dec 2023.

[6] IBM, "What is machine learning?", https://www.ibm.com/topics/machine-learning, Accessed 1 Dec 2023.

The main purpose of this thesis is to show that deep-seated biases of societal consciousness infiltrate technological systems, and the reliance of the field of artificial intelligence on science conceals gender bias. This issue not only exacerbates existing gender inequality stemming from epistemic and structural injustices but also opens a new battleground for gender equality due to the lack of discourse on the subject and the absence of legal protections focusing on it. Consequently, women find themselves disadvantaged in what appears to be an objective field and are unable to benefit from protections. This thesis aims to bridge the gap between the technical qualities of AI and the social dimensions that contribute to gender bias. By shedding light on how AI systems learn and perpetuate biases, alongside the human inputs and societal contexts that influence these systems, this thesis explores how insufficient current strategies are to support fairer, more inclusive AI technologies that promote equality and fairness for all individuals.

Questions which are aimed at cementing the foundation of this study were formed and the main evaluation of the thesis is built upon these questions. The questions are as follows: What is bias and its relation to injustice? Why do women face more bias than other groups and minorities disadvantaged by artificial intelligence? How can artificial intelligence be taken out of the scientific realm and made the subject of ethical discussions? Why is there a problem of diagnosis and consensus in this field? How can social norms and behaviors, structural injustices, and the capabilities approach be included in discussions about artificial intelligence ethics? What are the consequences of gender bias of AI for society at large? Where do legal regulations fall short in addressing the gender bias in the field of artificial intelligence and how can they succeed?

This thesis consists of two main sections. The first section is concerned with addressing the gender bias of AI, aiming to build the theoretical foundation of the thesis by focusing on

the reasons why the gender bias exists, as well as the factors concealing and exacerbating it. The second section focuses on exposing the implications of gender bias of AI from the perspective of economic and sociological, and then legal disciplines; underlining the broad effects of the gender bias problem to build the conceptual foundation of the thesis.

The inclusion of real-life instances and news reports in this study is meant to demonstrate that gender bias of artificial intelligence is a justice-threatening problem that is ever-growing with the increased use of AI applications. While the technical complexity and opacity of the field constitutes a barrier for understanding, newly emerging academic studies on the ethics of artificial intelligence have been invaluable sources for the theoretical foundation of the thesis. In today's technologically advanced and fast-paced world, new artificial intelligence applications and their biased assessments are reported frequently. The addressing of legal protections from various jurisdictions is aimed not at conducting a comprehensive and comparative legal analysis, but at underlining the economic gain these measures focus on, failing to tackle ethical issues such as gender bias. Therefore, this thesis is grounded on contributing to the field of ethics of artificial intelligence with a feminist perspective, from both theoretical and conceptual standpoints.

SECTION ONE
ADDRESSING THE GENDER BIAS OF AI

I. OVERVIEW OF GENDER BIAS IN AI APPLICATIONS

This section of the study aims to provide an overview of gender bias in AI applications, exploring what bias and its relation to injustice are, as well as why women are particularly at a disadvantage when other marginalized groups are also subject to AI's bias coupled with the biases in the technology industry. The basic principles of artificial intelligence and discussion points around machine learning are explored in order to take the subject out of the scientific realm and turn it into a subject of ethical discussions. The section continues with the exploration of how biases emerge in the technology industry, underlining the disproportionate gender participation in relation to existing male domination from a feminist point of view as well as the data problem which exacerbates the gender bias of AI with its proliferation. Structural injustice and the capabilities approach are discussed, aiming to shed light on how existing biases came into existence and why they continue.

A. THE MEANING OF BIAS

"*One simple answer is that bias, the undesirably debilitating kind of point of view, exists purely in the eye of the beholder; it is not a kind of point of view, but a point of view seen from a point of view*."[7] This statement by **Eugene Garver** reflects the complex and

[7] Eugene Garver, "Point of View, Bias, and Insight" *Metaphilosophy*, vol. 24, no. 1/2, 1993, pp. 47–60. http://www.jstor.org/stable/24438954, Accessed 9 Dec 2023.

subjective nature of bias, emphasizing that bias is not an inherent quality of a point of view but rather exists as a result of how it is perceived by different individuals or groups. Defining bias is not merely a definition of a word, but an interpretation, which again relies on how the person defining bias reflects their own thinking out to the world. For the sake of this thesis, bias will be interpreted as a point of view where an algorithm or human knowingly or unknowingly makes a decision that puts a segment of society at a clear disadvantage, thereby lowering their chances of achieving a result that they would if the decision had been made purely without any bias. Therefore, this thesis will focus on the scope of gender bias that puts women in predominantly disadvantageous positions while relying heavily on enforcing the existing biases in both our societies and the human psyche.

Biases are also a tangled component of injustice, which the task of addressing the gender bias of AI is aimed at demonstrating. They are inherently accepted and disguised as natural occurrences, which makes it harder to address the injustices they cause, and even acknowledge the existence of biases despite causing destruction of rights.[8] Therefore, a bias is similar to an invisible filter that distorts the looker's perception especially when it comes to minority groups which are already subject to discrimination, of which biases are a part of. They shape legal regulations and media content which are supposedly modern. The way that biases are intricately linked within the nuances of everyday life perpetuates the notion they ought to exist, because if there is a widely practiced and covert bias against women, it is easy for people to overlook it, accept it, not challenge it, and even defend its existence claiming

[8] Gülriz Uygur, "Toplumsal Cinsiyet ve Adalet: Hukuk Adaletsizdir", *Ankara Barosu Dergisi*, no. 4, 2015, pp. 121–132, https://dergipark.org.tr/tr/download/article-file/398472, Accessed 13 Mar 2023.

rightfulness. The struggle against biases for gender equality is not a small task; the subject of the struggle is woven within the fabric of a society that is used to inequality.

It must also be noted that biases should not be considered as an inherent quality of AI. Nothing written herein ought to be interpreted as a suggestion that only AI has biases, or possessing these biases is a quality that one cannot take away from AI without disrupting its entirety. Like all entities capable of making decisions, AI is not made purely out of biases or one-sided thinking, but rather, is compromised with ill-informed decisions causing disadvantages and inequality. One must acknowledge that gender bias is a direct result of AI's complexity, vast datasets, human conditioning, and decisions that contribute to these datasets, hence, a deep-rooted problem which requires a thorough approach. Tacking this deep-rooted problem requires an examination of forms of injustices that give rise to biases, as well as the social constructs which shape how biases are formed. The point of view of decision-makers on the basis of knowledge, perception and stereotypical thinking is what makes outcomes biased. Thus, the following subsections explore how epistemic injustice and social constructs distort the collective point of view and build the gender bias problem, which in turn is placed within AI applications.

1. Epistemic Injustice and Gender Bias

The relationship of epistemic injustice with the inherent gender bias prominent in AI applications can be based on **Miranda Fricker**'s definition, positioning epistemic injustice as the unjust treatment of individuals in terms of knowledge and understanding, including both testimonial injustice, where credibility of a speaker is unfairly undermined due to social

prejudices, and hermeneutical injustice, where certain experiences or perspectives are marginalized due to gaps in collective understanding.[9]

The injustice here stems from the determination that discrimination, which has to do with knowledge, against someone's capacity as a knower based on prejudices about the speaker, such as gender, social background, and so on can occur. [10] This is an injustice that both affects people's point of view in terms of the ability to know things, as well as the ability to be seen objectively by others. At this point the question arises: How does AI exacerbate the gender bias problem if it seemingly lacks the societal nuances and point of views of humans as social beings?

Fricker argues that understanding injustice in epistemic activities, particularly in relation to social identity and power dynamics is important and underlines the concept of epistemic injustice and its implications for epistemology, claiming that examining negative experiences is crucial for understanding justice.[11] This approach is taken within the context of focusing on the concept of injustice, rather than justice. She argues that focusing on justice creates the perception that justice is the norm and injustice is therefore the "unfortunate aberration", while this puts injustice at a negative space, rendering it subject to less focus and understanding.[12] The focus on injustice is especially crucial for understanding the gender bias of AI, as AI is generally perceived within the scientific domain tied to objectivity, thereby

---

[9] Miranda Fricker, *Epistemik Adaletsizlik - İktidar ve Bilmenin Etiği,* Translated by Kadir Gülen, Fol Kitap, Ankara, 2023.

[10] Morten Fibieger Byskov, "What Makes Epistemic Injustice an "Injustice"?", *Journal of Social Philosophy,* vol. 52, 18 May 2020, pp. 114-131, https://doi.org/10.1111/josp.12348, Accessed 3 Mar 2024.

[11] Fricker, 2023, p. 24.

[12] Fricker, 2023, p. 25.

just by default. The inherent prejudices and cultural exchanges clouded by biases form the vast data sets infested with epistemic injustice, which are fed to AI, contributing to seemingly scientific and objective practices which in turn perpetuate injustice. Per **Fricker**'s argument above, it is therefore important to examine negative experiences, which this thesis focuses on, particularly those that concern judicial rulings affected by epistemic perceptions, to determine whether AI's decision-making processes can be affected by such practices. The study mentioned below explores how epistemic perceptions based on gender affect the outcome of legal proceedings.

As a social issue, epistemic injustice is related to knowledge attributed to social groups, because ideally, gender should have no influence on treating others as equal epistemic agents.[13] An example of this is how epistemic perceptions affect the ruling in sexual assault and violence against women cases, depending on whether women or men are victims or perpetrators as studied by **Horniczak et al**. An interesting finding regarding this is that men are likely to be seen as the least reliable group when they claim to be assaulted by women. **Horniczak et al.** also posit that stating rational and unprejudiced differences in the assessment process should not be considered gender bias, whereas differences on a population level bring forward an inclination that it is so. This is particularly important in assessing AI's decision-making processes and inherent biases, because the assessment naturally requires an examination of the outcome and whether it puts a particular epistemic agent, such as women, at a disadvantage. Several real-life examples in this thesis note the

---

[13] Klaudyna Horniczak et al., "Is Epistemic Status Gender-Biased? Gender As a Predictor of Testimonial Reliability Assessments in Violent Crimes", *International Journal for the Semiotics of Law,* 17 Jan 2024, https://doi.org/10.1007/s11196-023-10055-6, Accessed 21 Feb 2024.

existence of a disadvantageous circumstance for women, which was either created or exacerbated by the use of AI.

While pointing at the epistemic injustice existing within both human and AI-based decision-making processes, it also important to note that not all decisions made by AI are biased, simply because these decisions are not based on statistical generalization of people. For example, the Retrieval-Augmented Generation process relies on the retrieval of information that utilizes the user input to pull information from a data source.[14] If the AI model has sufficient data, such as, the email inbox of a person, it can respond to the user input about what a certain email was about with a definitive answer pulled from that data source. The AI model in this example does not rely on statistics, therefore, does not attempt to make a generalization. However, when AI models form beliefs about individuals based on groups individuals belong to, the need for a discussion about ethics of AI arises and biases inevitably become the subject of this discussion. The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) recidivism algorithm, which is a criminal justice system used in the United States to assess the risk of reoffending after being arrested is an infamous example of AI beliefs.[15] The algorithm was found to be biased against Black arrestees as it assigned higher risk scores for Black defendants and more often predicted them to reoffend when they did not, as opposed to White defendants. False positives created by such systems and their increasing use raise concerns regarding civil liberties and human

---

[14] Rick Merritt, "What Is Retrieval-Augmented Generation, aka RAG?", *NVIDIA Blog,* 15 Nov 2023, https://blogs.nvidia.com/blog/what-is-retrieval-augmented-generation/, Accessed 24 Mar 2024.

[15] Prathamesh Patalay, "COMPAS : Unfair Algorithm?", *Medium*, 22 Nov 2023, https://medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a, Accessed 24 Mar 2024.

rights causes, such as gender equality.[16] Similar algorithms which are capable of profiling and stereotyping are subject to debates in ethics because they often create "ethical-epistemic dilemmas" where what should be believed based on knowledge clashes with what should be believed based on ethics.[17] It could be argued that instead of descriptive questions about whether specific beliefs meet various evaluative criteria like truth, or choosing to leave the assessments of AI out of discussion if they indeed turn out to be true even though they are based on statistical generalizations, unlike the COMPAS example above, focusing on what humans and AI should believe could be a novel approach to tackle the epistemic injustice and gender bias problem within AI. The group or groups individuals belong to should not get in the way of them being treated as equal epistemic agents, but some groups are subject to more epistemic injustice due to a lack of inclusion issue when it comes to cultural diversity and shaping norms that govern social interaction.

Chinala T. Okolo posits that explaining users how machine learning models work and make predictions, often promoted with the idea of an "Explainable AI (XAI)" could help AI become more relevant in real-life contexts.[18] The growing use of AI systems within everyday tasks creates the need for understanding these systems both from the developers' and users' perspective. This is particularly important for users to trust AI systems, yet existing biases and the general lack of technical know-how complicate the matter as many people

---

[16] Martina Sardelli, "Epistemic Injustice in the Age of AI", *Aporia*, vol. 22, 17 Aug 2022, pp. 44-53, https://ojs.st-andrews.ac.uk/index.php/aporia/article/view/2455/1871, Accessed 24 Mar 2024.

[17] Winnie Ma, Vincent Valton, "Toward an Ethics of AI Belief" *ArXiv*, 18 Sep 2023, p. 5, https://doi.org/10.48550/arXiv.2304.14577, Accessed 18 Mar 2024.

[18] Chinasa T. Okolo, "Towards a Praxis for Intercultural Ethics in Explainable AI", *ArXiv*, 25 Apr 2023, p. 1, https://doi.org/10.48550/arXiv.2304.11861, Accessed 18 Mar 2024.

even in today's increasingly tech-savvy world cannot access AI technologies, partially due to structural injustices which are discussed further down below. The inefficacy of XAI in terms of explaining the working mechanism of AI to users can be attributed to the lack of cultural diversity and inclusion, which is why **Okolo** believes researchers should incorporate an intercultural ethics approach into AI development.[19] This way, ethical standards that are sensitive to cultural norms can be applied with the negotiation and inclusion of people from various cultures and especially low-income domains which do not have an equal ease of access to AI technologies compared to the developed parts of the world. When some groups of individuals experience bias against them, the opposing groups usually benefit from this bias, which is why this unequal power dynamic continues at the advantage of the benefiting groups due to the discriminated groups' lack of understanding of and access to the systems discriminating against them. Forming beliefs about certain groups of individuals that get in the way of them being treated as equal epistemic agents could be repressed to a certain degree if disadvantaged groups understood the systems that acted biased towards them.

The lack of diversity and inclusion in AI systems is one of the factors contributing to the epistemic injustice and gender bias problem of digital environments in general, but **Catriona Gray** notes that these environments are capable of creating novel and distinctive epistemic injustices as well.[20] She refers to **Symons** and **Alvarado** to shed light on the epistemic opacity and complexity of AI systems, basically showing their complexity as the reason for hermeneutical and testimonial epistemic injustices arising through AI

---

[19] Okolo, 2023, p. 2.

[20] Catriona Gray, "Testimonial Injustice in Governmental AI Systems", *KI-Realitäten,* 2023, pp. 67-92, https://doi.org/10.1515/9783839466605-004, Accessed 18 Mar 2024.

technologies.[21] **Gray** explores three evidential strategies -obviation, diminishment, and impugnment- which reshape the significance of testimony, along with the implications for epistemic justice arising from their implementation. Through a proposal of a set of corrective virtues to address instances of epistemic injustice, **Gray** notes that epistemic injustices are inseparable from unjust social structures.[22] It is therefore a conclusion that the ability to cultivate epistemic virtues as collective and organizational values means that these can be extended to governing AI's decision-making.[23] The public use of AI alters power dynamics between institutions and individuals, positioning people as subjects which can be read by the state in a way that silences and oppresses them.[24] This way the concept of democratic citizenship and the place of testimony in public policy are reshaped[25], giving rise to novel and distinctive epistemic injustices. Similarly, the Natural Language Processing (NLP) process, which is at the intersection of linguistics, computer science and AI leads to the speech of marginalized groups being systemically misunderstood and silenced due to their reliance on human language on the internet which is full of biased and sexist comments.[26] This occurrence is another indicator that problems of hermeneutical injustice are present in algorithms which learn how to be biased from humans, but also cause new epistemic

---

[21] Gray, 2023, p. 71.

[22] Gray, 2023, p. 80.

[23] Gray, 2023, p. 81.

[24] Gray, 2023, p. 82.

[25] Gray, 2023, p. 85.

[26] Karen Frost-Arnold, "Epistemic Injustice and AI Ethics: Learning from Google's Treatment of Timnit Gebru", *Feminism, Social Justice, and AI*, https://feminism-social-justice-ai.org/epistemic-injustice-and-ai-ethics-learning-from-googles-treatment-of-timnit-gebru/, Accessed 20 Mar 2024.

injustices. Computer scientist Timnit Gebru's firing from Google for exposing the culture of epistemic injustice within the AI industry and the attacks on her credibility[27] is the crowning touch that shows why gender bias must be addressed.

The extent of epistemic injustice is not just confined to bias in emerging technologies. Epistemic injustice also plays a role in the collective trauma experienced by women as they are subject to daily media reports on domestic violence and femicide. [28] Each case is unique, and an overgeneralization may not be the most accurate way to address these violence problems at large. Yet, the mutual aspect in all these situations is that the victims can be classified altogether as an epistemic agent, which highlights gender differences at a population level. Through the constant and widespread representation of epistemic injustice, biases strengthen their hold in the collective conscious and morals. As discussed above, tackling bias is a tricky task because it has to do with the covert and invisible judgment that conceals itself as natural, whether in relation to inherent qualities or long-practiced social rituals. One could argue that epistemic injustice has more to do with the lack of knowledge rather than its existence. Even further, this lack of knowledge infiltrates into societal and institutional structures in ways that shape the conditions and challenges faced by individuals, leading to structural injustices. Surely, identifying an epistemic agent, and acknowledging

---

[27] John Harris, "'There was all sorts of toxic behaviour': Timnit Gebru on her sacking by Google, AI's dangers and big tech's biases", *The Guardian,* 22 May 2023, https://www.theguardian.com/lifeandstyle/2023/may/22/there-was-all-sorts-of-toxic-behaviour-timnit-gebru-on-her-sacking-by-google-ais-dangers-and-big-techs-biases, Accessed 26 Mar 2024.

[28] Baris Cayli Messina, "Breaking the silence on femicide: How women challenge epistemic injustice and male violence", *The British Journal of Sociology,* vol. 73, 10 Jul 2022, pp. 859-884. https://doi.org/10.1111/1468-4446.12968, Accessed 8 Mar 2024.

someone as a woman affects the assessment process in ways that put the subject in disadvantageous conditions, but it is the lack of awareness about the thought process that contributes to the ultimate injustice. This lack of awareness could easily be classified as a bias, but there are also causes to the lack of awareness as the way of collective thinking and point of view is heavily based on social constructs which regulate power dynamics and social interactions.

2. Social Constructs and Gender Bias

Social construct refers to a conception that has been created and accepted by people in a society.[29] An example could be social classes; as they are not biological but rather, born out of widely practiced cultural ideas and acceptances. The distinction between the biological, or natural per se, and the socially constructed is important in terms of addressing how biases are formed in parallel with the social constructs that pave the way. The difference between sex and gender therefore is a natural subject or discussions surrounding gender bias. The usual definition is that sex is a biological category which is divided into female and male, while gender refers to social constructs which assign identities such as girls, boys and women to individuals based on roles and behaviors.[30] These social constructs play a role in the creation and exacerbation of biases in two ways: They create borders on what is considered socially acceptable for a specific role or category, deeming concepts outside these borders socially

---

[29] Merriam-Webster, "social construct", https://www.merriam-webster.com/dictionary/social%20construct, Accessed 8 Mar 2024.

[30] Council of Europe, "Sex and gender", https://www.coe.int/en/web/gender-matters/sex-and-gender, Accessed 8 Mar 2024.

unacceptable, therefore subject to negative outlook and bias; and they firmly constitute epistemic agents within said borders which then are assessed unequally by virtue of existence.

A common misconception regarding social constructs is that they are not constructs, since they are either based on natural occurrences, such as some gender roles and stereotypes being based on the average physical traits and strength of females and males. Another example could be how humans perceive time. Surely, from the perspective of humans on earth, there is a sunset and sunrise every day. The Earth revolves around the Sun. There is a conception of progression, as seasons change, living beings get older and the Sun disappears and reappears each day. Yet, it is civilized humans who divided this ever-continuing progression into time periods, such as, years, weeks, and hours. These time periods are based on natural occurrences, but also socially constructed. They may vary depending on cultures, or even not exist in a hypothetical Earth which has the same of everything our Earth does except this. The same is applicable for gender and its impact on social exchanges. A hypothetical, identical earth may have mothers as biologically females give birth to offspring, but they may not have the same conception of motherhood, Mother's Day celebrations and holiness around the subject as our earth does.[31] Likewise, women tend to be generally of smaller build compared to men[32], yet this does not mean that they should be assessed with the biased acceptance of weakness and have access to only household tasks in terms of job

---

[31] Abigail Thorn, "Social Constructs (or, 'What is A Woman, Really?')", *YouTube*, Uploaded by Philosophy Tube, 9 Jul 2021, https://youtu.be/koud7hgGyQ8, Accessed 11 Mar 2024.

[32] WorldData, "Average height and weight by country", https://www.worlddata.info/average-bodyheight.php, Accessed 10 Mar 2024.

opportunities. In addition, biases are likely to stick around for a while despite changing times and cultural climates. **Judith Butler** claims that what it means to be a woman within social definitions change from time to time, and it needs to be so.[33] This can be interpreted both as a pathway to more inclusion, and a natural evolutionary response for struggle against inequality. However, the adaptability of biases and stereotypes manifests itself as coming into existence in new practices, such as laws, technologies, cultural resets, and media.

In reality, concepts which are thought to be based on nature, biology, and generally science are taken from their origin and given abstract meanings in a social context, which is why what a woman is subject to discussions, as well as biases based on positive or negative qualities a woman is thought to possess. Woman is a social construct which has a myriad of connotations surrounding it, naturally entangled in subjectivity, which is how some of the biases in relation to gender are formed and the point of view is distorted.

B. THE GENDER BIAS: WOMEN'S DISADVANTAGE

There are many ways of being a woman. Although there are claims at conflict regarding the topic, this study will not focus on who is a woman or what identities qualify as women. To gain a deeper insight into the gender bias in AI, and view it from a broader lens, one ought to acknowledge that a woman is someone who identifies as a woman[34], regardless of their body features, the cultural stereotypes they conform to, and the gender they were

---

[33] Jules Gleeson, "Judith Butler: 'We need to rethink the category of woman'", *The Guardian,* 7 Sep 2021, https://www.theguardian.com/lifeandstyle/2021/sep/07/judith-butler-interview-gender, Accessed 8 Mar 2024.

[34] LGBT Foundation, "What Is A Woman?", 8 Aug 2023, https://lgbt.foundation/whatisawoman, Accessed 30 Nov 2023.

assigned at birth. In summary, both cisgender and transgender woman are broadly referred to as women in this study. Nonetheless, this acknowledgement does not come at the price of undermining intersectional feminism[35], but rather, aims to enforce the approach that the gender bias should be viewed from a broader lens in order to comprehensively identify the problems surrounding the matter, as discussed above.

Women are and have been underrepresented in science and digital fields. Only a third of researchers in science are women.[36] This study aims not to focus on the core reasons for these disproportionate statistics, however it is clear that the more women are represented in science and digital fields, the more their contributions will help AI become an inclusive tool.

It is also worth noting that AI shows many biases towards different social segments and groups. Not only are already-discriminated groups subject to AI's biases, but also the core functions of AI seem to favor the dominant race, gender, and sexual orientation. The main reason why this thesis will focus on women as the subject of AI's bias is that women are arguably the largest group to be discriminated against. They are subject to sexual abuse, the glass ceiling, female genital mutilation, and various forms of harassment and discrimination which manifest themselves as numerous instances happening all over the world. All the headlines below are reported from different countries, wealth groups and cultures.

---

[35] IWDA, "What Does Intersectional Feminism Actually Mean?", 11 May 2018, https://iwda.org.au/what-does-intersectional-feminism-actually-mean/, Accessed 30 Nov 2023.

[36] UNDP, "Women and Girls in science and technology: Bridging the gender gap", 10 Feb 2022, https://www.undp.org/ghana/blog/women-and-girls-science-and-technology-bridging-gender-gap?gad_source=1&gclid=CjwKCAiAjrarBhAWEiwA2qWdCMTCi5Cf66EDleK_qbRiRivhynG-2QNIa1fF6tgtmM7Qn3yafv70XRoCquQQAvD_BwE, Accessed 1 Dec 2023.

"Discrimination in the form of sexual harassment has been in the headlines for weeks now, but new poll results being released by NPR show that other forms of discrimination against women are also pervasive in American society."[37]

"This analysis argues that gender quotas have been ineffective and that several aspects of Nordic social policies have negatively affected women's career progress and even contributed to a glass ceiling."[38]

"More than 200 million girls and women alive today have undergone female genital mutilation (FGM) in 30 countries in Africa, the Middle East and Asia where FGM is practiced."[39]

It is therefore a logical conclusion that the widespread discrimination against women, among all minority groups, manifests itself as a form of bias in AI applications. To refer to the subtitle above, regarding what bias is, the bias of AI reflects the thinking of the world it is designed in, and is therefore biased against women, just like the humans and societies that contribute to its existence.

---

[37] Joe Neel, "Poll: Discrimination Against Women Is Common Across Races, Ethnicities, Identities", *National Public Radio*, 11 Dec 2017, https://www.npr.org/2017/12/11/569251267/poll-discrimination-against-women-is-common-across-races-ethnicities-identities, Accessed 2 Dec 2023.

[38] Nima Sanandaji, "The Nordic Glass Ceiling", *CATO Institute,* 8 Mar 2018, https://www.cato.org/policy-analysis/nordic-glass-ceiling, Accessed 1 Dec 2023.

[39] World Health Organization, "Female genital mutilation", https://www.who.int/news-room/fact-sheets/detail/female-genital-mutilation, Accessed 2 Dec 2023.

C. WOMEN AND OTHER MARGINALIZED GROUPS

As discussed above, when discrimination is mentioned, unlike other forms of discrimination, the prominence of gender-based discrimination stands out. This systemic and historical discrimination against half of humanity in terms of sexism is evident throughout history and across nearly every society. This discrimination occurs in artificial intelligence technology and consequently in its active execution through machine learning, due to the transferring of datasets of human memory to algorithms. However, it is crucial to note that machines do not generate or process data on their own; one must emphasize the direct impact of the human factor on the functioning of artificial intelligence.

The significance of gender-based discrimination in machine learning, or broadly, artificial intelligence lies in the repetitive occurrence of one-sided and discriminatory errors by machines unless intervened by humans in datasets or algorithms.[40] As long as the human factor and its biased effects remain unaddressed, machines will continue to perform operations that bring forth discrimination issues. However, the elimination of biased effects from the human factor will only be possible by accurately diagnosing the problems, determining, and implementing solutions. Examining the historical struggle against gender discrimination, it becomes evident how challenging it is to overcome this widespread and longstanding bias within human communities. Nevertheless, due to machine learning's direct operation and reliance on datasets, control mechanisms on data seems to provide a relatively straightforward solution to this problem.

---

[40] Eda Eryılmaz, "The Sexism Problem of AI", *Medium*, 25 Sep 2021, https://aninjusticemag.com/the-sexism-problem-of-ai-a7594a64f7f2, Accessed 17 Jan 2022.

Certainly, marginalized groups exposed to discrimination in artificial intelligence are not limited to women alone. In the process of machine learning, where only data from white individuals is used, the incident of an erroneous identification of a black man in the state of Michigan, United States, being wrongfully detained due to a facial recognition system's failure to recognize black faces, has raised issues of racism in artificial intelligence.[41] Facial recognition systems built through machine learning primarily recognize white male faces. These systems exhibit an error rate of 0.8% when identifying light-skinned male faces, while the error rate can reach 35.5% in recognizing dark-skinned female faces, attributed to the inadequacy and bias in the provided data.[42] The AI systems reflect the social dynamics in which certain groups of individuals are regarded as the default, whereas the rest are simply biproducts, or worse, anomalies.

Similarly, Automated Gender Recognition (AGR) processes infer gender data about a person based on their legal name, the shape of their facial bones, whether or not they wear makeup, therefore undermining gender expression and squeezing gender indentity into non-inclusive binary terms.[43] Based on queer theory, **Dawn McAra-Hunter** argues that heteronormativity and cisnormativity refer to the assumptions that heterosexuality and

---

[41] Kashmir Hill, "Wrongfully Accused by an Algorithm", *The New York Times,* 24 Jun 2020, https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html, Accessed 18 Jan 2022.

[42] Timnit Gebru, "Race and Gender: Data-Driven Claims Race and Gender Perpetuate the Negative Biases of the Day", *The Oxford Handbook of Ethics of AI* (Compiled by: Markus D. Dubber, Frank Pasquale, Sunit Das), Oxford University Press, New York, 2020, p. 258.

[43] Daniel Leufer, "Computers are binary, people are not: how AI systems undermine LGBTQ identity", *Access Now,* 13 Jan 2024, https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/, Accessed 28 Mar 2024.

cisgenderism are the norms, which AI systems often default to.[44] She refers to **Foucault**'s analysis of power dynamics and exploration of the history of sexuality, positing that "subjectivity is not an isolated concept in philosophy; rather, it is shaped and influenced by knowledge and power". The concentrated and enacted power structures within society that discriminate against LGBTQ+ individuals diffuse into the AI applications trained and used by the same power structures, opening up new discrimination areas especially via algorithms systems that are meant to predict social outcomes.

Artificial intelligence, through incomplete, inadequate, or unsupervised datasets in machine learning, causes discrimination in terms of gender, race, and socioeconomic status. Moreover, categorizing the resulting discriminations into a limited number of groups is not always feasible. However, as indicated earlier, the emergence of gender-based discrimination in machine learning through the biased use of datasets, considering that the existing dataset should naturally consist of half of humanity, i.e., women, is both ironic and specific. Therefore, the issue of sexism in artificial intelligence stands out among the discrimination problems that need to be addressed due to the size of the group it puts at a disadvantage**,** but AI must be taken out of the scientific realm and its basic principles must be understood for such an attempt.

---

[44] Dawn McAra-Hunter, "How AI hype impacts the LGBTQ + community", *AI Ethics,* 14 Feb 2024, https://doi.org/10.1007/s43681-024-00423-8, Accessed 28 Mar 2024.

D. BASIC PRINCIPLES OF ARTIFICIAL INTELLIGENCE

Simply put, artificial intelligence as the general concept or technology refers to computer programs which aim to understand and interpret human intelligence based on datasets they have.[45] It is worth noting that, despite necessarily touching pain points regarding gender inequality and bias within the field, the focus of this thesis is this general concept, rather than the industry or discipline.[46] For the purpose of exploring and exposing the gender bias in the general concept of artificial intelligence, there are instances where one must deem it essential to blend the technology and the discipline together, due to the underlying inequality present at multiple stages and fragments of AI.

One of the most famous applications of the aforementioned definition of AI is the Turing test. Proposed by mathematician **Alan Turing**, the Turing test aims to determine whether a computer can "think", make decisions indistinguishable from those of a human being. The basic premise of this test is that a remote human interrogator poses questions, and then try to distinguish whether the answers they receive are given by a human or a computer.[47] Computers that the human interrogators think are humans pass this test. Despite not providing

---

[45] IBM, "What is artificial intelligence?", https://www.ibm.com/topics/artificial-intelligence, Accessed 30 Jan 2022.

[46] Sourajit Saha, "Should Artificial Intelligence Be Capitalized?", *Medium*, 16 Jul 2023, https://medium.com/@sourajitsaha17/should-artificial-intelligence-be-capitalized-98e63485ba0f#:~:text=There%20is%20no%20universally%20agreed,the%20general%20concept%20or%20technology, Accessed 03 Dec 2023.

[47] Brittanica, "Turing test", https://www.britannica.com/technology/Turing-test, Accessed 3 Dec 2023.

a profound insight into the field of technology in its entirety, and referring to its deeper, more complex manifestations; this famous test clarifies the basis of AI.

It is crucial to understand that artificial intelligence is omnipresent in various fields. "Artificial intelligence makes it possible to analyze large amounts of data (big data) in real-time, providing forecasts that can support the clinician's decisions."[48] "In the criminal justice field, the use of AI systems for providing investigative assistance and automating decision-making processes is already in place in many judicial systems across the world."[49] These examples demonstrate the extent of use and effect of AI in our lives, mainly positioned to aid human professionals, but also perfectly capable of guiding them towards ill-informed decisions. One of the main things to consider about AI is that it is not really in charge. Designed, fed with data, and trained by humans; the general purpose of AI is to automate processes, help humans with their decisions and make people's lives better. Therefore, AI is not a boss, governor, or leader. It is merely a product of human intelligence and technology created to be a tool of comfort and security, which, to be clear, does not undermine its positive and adverse effects on society, but rather, amplifies them. An ever-changing and ever-evolving tool, whose exact limits are unknown to those that are not experts, it is perfectly capable of perpetuating biases towards certain societal groups. Although AI is not the boss who makes the decisions, it is the trusted advisor, and sometimes the confidant guiding those who decide who should be hired, granted a loan, and who should be a doctor or a nurse.

---

[48] Valentina Bellini et al., "Understanding basic principles of Artificial Intelligence: a practical guide for intensivists.", *Acta Biomed,* vol. 93,5, 26 Oct 2022, https://doi.org/10.23750/abm.v93i5.13626, Accessed 5 Dec 2023.

[49] UNESCO, "AI and the Rule of Law: Capacity Building for Judicial Systems", https://www.unesco.org/en/artificial-intelligence/rule-law/mooc-judges, Accessed 3 Dec 2023.

AI models rely on machine learning algorithms and artificial neural networks for a logical decision-making process using the data at their disposal.[50] This data could be anything, such as, a list of wars and battles fought by humans since 100 A.D., or statistics regarding which resumes sent to a company for job applications after 1990 are more likely to land the jobs. An AI model is either a program or algorithm that, based on training, recognizes patterns, and makes predictions or decisions. A simple example could be a film recommendation tool. For the sake of the argument, we can suppose that there is an AI model that records and keeps data on the films people watch, the ratings they give, and the films and actors they like. Through machine learning, the AI can analyze and recognize patterns regarding which new films those people are prone to liking. Therefore, this film recommendation tool can analyze the data, recognize patterns, and decide on appropriate film recommendations for each person based on their own individual data. In the end, it is the person's prerogative to actually watch the film or not, so the end decision belongs to humans, but even this very simple example is capable of showcasing the potential of AI in terms of guiding people towards certain judgments. All in all, the bias in AI can be directly linked to its reliance on machine learning which will be further explored below.

---

[50] Nitin Kumar Singh et al., "Artificial intelligence and machine learning-based monitoring and design of biological wastewater treatment systems", *Bioresource Technology,* vol. 369, Feb 2023, https://doi.org/10.1016/j.biortech.2022.128486, Accessed 12 Dec 2023.

E. DISCUSSION POINTS AROUND MACHINE LEARNING

The following subsections explore several discussion points around machine learning, underlining why there is a problem of consensus and diagnosing in the field, the insufficiency of legal protections, and how the belief in science's objectivity conceals ethical issues within.

1. The Problem of Consensus and Diagnosis

The emergence and oversight of discrimination and ethical issues created by machine learning cannot be attributed to a single cause. The rapid advancement of artificial intelligence and information technologies contributes to ethical problems in terms of discrimination due to their perception as fields requiring technical expertise, thereby excluding individuals lacking this technical knowledge and creating inaccessible gray areas. This is among many reasons that can be seen as playing a role in the emergence of ethical issues related to discrimination. For instance, machine learning is considered an area requiring expertise in software and programming, primarily occupied by engineers, programmers, and data architects. According to statistics, only 22% of researchers involved in artificial intelligence studies are women.[51] Therefore, the perception that machine learning is a domain primarily conducted by experts in such specialized fields leads to a distant approach by social scientists and lawyers who could conduct ethical research and evaluations

---

[51] World Economic Forum, "Assessing Gender Gaps in Artificial Intelligence", https://reports.weforum.org/global-gender-gap-report-2018/assessing-gender-gaps-in-artificial-intelligence/, Accessed 13 Jan 2022.

on this matter. Consequently, achieving consensus, especially regarding ethical issues related to discrimination within machine learning, becomes more challenging, and discrimination issues such as sexism and racism emerging within this field are not thoroughly diagnosed.

2. Insufficiency of Legal Protections in Terms of Problem Resolution

As emphasized earlier, reaching consensus on the issues arising from machine learning and fully diagnosing these ethical problems are crucial in preventing them. However, once these issues emerge, developing legal protections becomes necessary both to prevent problems indirectly and to resolve disputes that arise. With the widespread recognition of the concept of artificial intelligence, many countries have established national-level administrative boards and research committees. An example in Turkey is the National Artificial Intelligence Strategy Steering Committee[52] under the presidency. Although these committees increase general awareness of AI, they are still far from drawing attention to ethical issues and developing legal protections against them. Instead, the primary task of these boards is to focus on commercial and technological developments to achieve international competitiveness in this field. Therefore, the development of legal protections and codification is currently not promising to resolve ethical issues.

Internationally, the situation is not significantly different. Despite the committees and organizations working on artificial intelligence, the absence of international organizations

---

[52] TRT Haber, "Ulusal Yapay Zekâ Stratejisi Yönlendirme Kurulu Toplanacak", 13 Jan 2023, https://www.trthaber.com/haber/gundem/ulusal-yapay-zeka-stratejisi-yonlendirme-kurulu-toplanacak-644569.html, Accessed 13 Jan 2022.

focusing on discrimination or ethical issues at a general level poses an obstacle to the development of legal protections. While researchers operate within universities, such as at the Massachusetts Institute of Technology (MIT)[53], their research efforts cannot translate into effective legal protection mechanisms through legislation. The European Commission published the highly anticipated *White Paper on Artificial Intelligence – A European Approach to excellence and trust* on February 19, 2020[54], followed by several legislative attempts such as *The EU Artificial Intelligence Act*[55] as the world's first major act to regulate AI formally adopted on 13 Mar 2024 by the European Parliament.[56] However, considering the novelty of these developments, they raise questions about the effectiveness of the impact on the datasets involved in the development of machine learning algorithms, given their size and antiquity. While expanding discrimination laws to encompass machine learning might not be difficult, the lack of an international precedent or jurisprudence in this regard, coupled with each country's sovereignty in international law, currently seems to hinder the possibility of developing legal protections that effectively address discrimination or ethical issues at a general level soon.

---

[53] Gebru, 2020, p. 262.

[54] Thomas Wahl, "Commission: White Paper on AI", 1 Apr 2020, https://eucrim.eu/news/commission-white-paper-ai/, Accessed 3 Dec 2023.

[55] European Parliament, "EU AI Act: first regulation on artificial intelligence", 19 Dec 2023, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, Accessed 24 Mar 2024.

[56] Karen Gilchrist, "World's first major act to regulate AI passed by European lawmakers", *CNBC*, 13 Mar 2024, https://www.cnbc.com/2024/03/13/european-lawmakers-endorse-worlds-first-major-act-to-regulate-ai.html, Accessed 24 Mar 2024.

3. The Assumption that There Will Be No Bias or Discrimination in Scientific Fields

Computer Science, as one of the fastest-growing fields in today's scientific landscape and being a part of the realm of numerical sciences, is often perceived to be as objective as numbers themselves. This belief stems from society's voluntary naivety. For instance, even the concept of "data" has raised ethical and legal concerns, straying away from scientific objectivity. The full adaptation of data protection laws highlighting the ethical and legal issues surrounding personal data could only be achieved after the 2010s. The General Data Protection Regulation (GDPR), considered a milestone in personal data protection law within the European Union, came into effect on May 25, 2018.[57] Similarly in Turkey, the Law on the Protection of Personal Data numbered 6698, was enacted on April 7, 2016.[58] The belief in the objectivity of computers that process data, apply rules, and utilize algorithms through machine learning makes it challenging to identify problems.

Undoubtedly, the societal belief in the objectivity of technology and science, even with the increasing interest in artificial intelligence literacy, has not adequately reached the level of analyzing ethical issues. Essentially, in machine learning, computers process a given

---

[57] Local Government Association, "General Data Protection Regulation (GDPR)", https://www.local.gov.uk/our-support/research-and-data/data-and-transparency/general-data-protection-regulation-gdpr#:~:text=The%20GDPR%20is%20Europe's%20new,started%20on%2025%20May%202018, Accessed 01 Dec 2023.

[58] Kişisel Verileri Koruma Kurumu, "Kurumsal Tarihçe", https://www.kvkk.gov.tr/Icerik/2075/Kurumsal-Tarihce, Accessed 14 Jan 2022.

dataset, and biased data within this dataset can influence the conclusions reached by the computer. However, due to the lack of widespread understanding of what fundamentally constitutes machine learning and technology literacy, there's a gap in awareness. Consequently, the ignorance that providing data to computers in machine learning could harbor biases and ethical issues from the presented data, essentially stemming from human processes leading to discrimination, is strengthened by the trust in the objectivity of Computer Science. This situation poses an obstacle to the identification, prevention, and resolution of ethical issues caused by artificial intelligence.

Alongside the realization of the essentiality of eliminating the human factor for scientific progress, a prevalent belief has emerged in societies that science and technology are exempt from this factor. The existence of unequivocal truths like "two plus two equals four", which will remain true regardless of the perspective, or at least the acceptance of certain facts as true within the scientific framework, has nourished the perception that objectivity and accuracy prevail in every subject matter tackled by science. However, this is a potentially grave mistake that can lead to serious consequences. **Timnit Gebru** exemplifies how even **Charles Darwin**, who revolutionized the scientific world by introducing the theory of evolution, expressed racist notions that are now unacceptable and often disregarded.[59] **Darwin** linked his work to unfounded racist claims, asserting a hierarchy among races. Similarly, Nazi Germany, aiming to create an Aryan race and intending to annihilate minorities for this cause, put an emphasis on national health.[60]. The discoveries made in the

---

[59] Gebru, 2020, p. 254.

[60] Tim Radford, "What has medicine learned from the Nazis?", *The Guardian,* 14 Oct 2004, *https://www.theguardian.com/news/2004/oct/14/thisweekssciencequestions.cancer,* Accessed 4 Dec 2023.

process of forming this so-called healthy nation while committing atrocities[61] led to new scientific findings, but at the expense of the question of whether it is ethical to use Nazi research in science.[62] Hence, it would not always be accurate to link scientific advancements with objectivity and ethical compliance. In social sciences, there is relatively higher awareness of the human factor since, by definition, social sciences center around humans. However, especially in numerical and natural sciences, the notion that "wherever you go in the world, things won't change" contributes to the belief that these fields are devoid of any human influence. Computer Science and information technologies are also directly influenced by this perspective.

Even in these scientific fields, biases can manifest in numerous ways. Despite a study being based on scientific facts that are believed to remain constant everywhere, the act of disregarding it solely because of the identity or worldview of the scientists conducting it is not uncommon. For instance, **Susan Michie**, expressing her opinion in a *Daily Mirror* article about the likelihood of maintaining social distancing and masks indefinitely after the COVID-19 pandemic, was criticized for being a communist rather than for her argument.[63] Moreover, it's a common occurrence for scientists to have biases, and the inability to eliminate the human factor from their work is expected. While scientists have an ethical responsibility to remain impartial, presuming that this responsibility is always met leads to

---

[61] Volker Roelcke, "Nazi medicine and research on human beings", *The Lancet,* vol. 364. Dec 2004, pp. 6-8, https://doi.org/10.1016/S0140-6736(04)17619-8, Accessed 4 Dec 2023.

[62] Frank Swain, "Is it right to use Nazi research if it can save lives?", *BBC*, 24 Jul 2019, https://www.bbc.com/future/article/20190723-the-ethics-of-using-nazi-science, Accessed 4 Dec 2023.

[63] Stephen John, "Why science isn't objective", *iai news,* 26 Jul 2021, https://iai.tv/articles/why-science-isnt-objective-auid-1846, Accessed 19 Jan 2022.

the failure to detect biases. Similarly, there are viewpoints suggesting that engaging in scientific endeavors without relying on assumptions rooted in biases, ethics, politics, or economic perspectives is impossible. Therefore, without asserting that science can be entirely unbiased, or science can never be unbiased, the fundamental point highlighted by this part of the study is the existence of skepticism about the objectivity of science. Consequently, the belief in society regarding objectivity appears baseless, hence inadvertently amplifying biases by failing to draw attention to them. For instance, the scientific research findings of a misogynistic scientist regarding women's success in the workplace or the behavior of a woman's brain in risky situations would undoubtedly be influenced by their biases.

A good starting point at tackling this problem could be focusing on increasing the AI literacy among people, so that knowledge about the possible societal risks of AI could be present among society. The cloud of doubt and ignorance surrounding artificial intelligence applications indicates that the general public lacks answers to questions they do not even know how to pose. For instance, in the United States, a country with technology hubs and a vibrant AI scene, awareness and knowledge about artificial intelligence still has big room to grow, while people with higher family income seem to show higher levels of understanding about AI and its use cases.[64] Younger people also seem to show awareness more than older Americans. Examining the awareness of different social segments regarding AI is not the focus of this thesis, therefore it would suffice to posit that in order to battle biases of AI,

[64] Brian Kennedy et al., "Public Awareness of Artificial Intelligence in Everyday Activities", *Pew Research Center,* 15 Feb 2023, https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/, Accessed 4 Dec 2023.

specifically gender bias, which is omnipresent and widespread, people need to be educated on the subject based on their social status and ease of access to information.

II. THE EMERGENCE OF BIASES IN TECHNOLOGY

A. DISPROPORTIONATE GENDER PARTICIPATION IN THE INDUSTRY

If AI is considered a faction within the technology industry, which was globally worth $8.51 trillion as of 2022[65], it becomes apparent that the lack of equal participation in this industry is directly linked to women's disadvantage, despite this being a multifaceted and deep-rooted issue. Women make up less than a third of the global technology workforce and hold only 28% of all jobs in computer and mathematical occupations.[66]

One could try to find reasons as to why this disproportionate participation is particularly visible in the technology industry. Naturally, a variety of factors contribute to the landscape of inequality in opportunity, which is not limited to women's disadvantage but is more of an issue deep-seated in social and structural injustice.

---

[65] MGI Research, "The Global Tech Market Is Bigger Than You Think", 28 Feb 2023, https://mgiresearch.com/research/the-global-tech-market-is-bigger-than-you-think/, Accessed 5 Dec 2023.

[66] Women in Tech Network, "Women in Tech and STEM Workforce", https://www.womentech.net/en-us/women-technology-statistics, Accessed 4 Dec 2023.

1. Structural Injustice

The injustices caused and exacerbated by the use, governance and development of AI are linked with structures just as they are linked with knowledge and understanding. **Iris Marion Young** argues that judgments of injustice often concern structure[67], and the ways of understanding responsibility to address injustice, and injustices that stem from unjust, discriminatory, and oppressive social institutions are inadequate.[68] These are structural injustices, which can best be explained with **Young**'s widely discussed example of *Sandy*.[69] In short, *Sandy* is a single mother who is forced to look for another home after being evicted. She searches for another apartment, but the apartments she can afford are either too far from her workplace, or in unsafe neighborhoods. She then allocates some of her budget for car renting, and decides to apply for housing support, but learns that the waiting time is too long, and ends up homeless after not being able to afford the security deposit of her new place. In this example, every bureaucratic step was completed in compliance with laws and prevalent moral norms, however *Sandy* is a not victim due to her actions, but due to structural injustices which are deep-seated in institutions. Structural injustice leaves individuals exposed to

---

[67] Iris Marion Young, "Political Responsibility and Structural Injustice", *The Lindley Lecture, University of Kansas,* 5 May 2003.

[68] David Estlund, "What's Unjust About Structural Injustice?", *Ethics: An International Journal of Social, Political, and Legal Philosophy,* vol. 134, no. 3, 2024, pp. 333-448, https://doi.org/10.1086/728634, Accessed 28 Mar 2024.

[69] Nadire Özdemir, "Iris Marion Young Teorisine Genel Bir Bakış: Yapısal Adaletsizlik ve Adaletsizlikten Sorumluluğa İlişkin Sosyal Bağlantı Modeli", *Türkiye Barolar Birliği Dergisi*, vol. 32, no. 146, Jan 2020, pp. 247-267, http://tbbdergisi.barobirlik.org.tr/m2020-146-1899, Accessed 9 Dec 2023.

exploitation, oppression, and discrimination.[70] The people who carried out the bureaucratic steps in *Sandy*'s example, such as, the landlords and real estate agents might be easy to be held accountable, however focusing on them shifts the blame from the institutionalized exploitation, oppression, and discrimination outside the reach of legal regulations. The matter is not about whether badly-off individuals, such as *Sandy*, make the decision with the best possible outcome even in dire situations, but about whether these individuals should be put under circumstances which render them vulnerable. Therefore, the focus of **Young**'s example is not exposing the unjust circumstances her character is put through but rather, the relationship of these unjust circumstances with the structure itself.[71] It is worth noting that **Young** also concludes that responsibility for addressing structural injustice belongs to those "who contribute by their actions to structural processes with some unjust outcomes."[72]

Iris Marion Young also posits that in societies characterized by welfare capitalism, discussions about justice primarily revolve around how resources are distributed, sidelining other crucial topics like the organization of production and decision-making structures. This narrowed public debate has led to criticism that such societies are depoliticized, as citizens are predominantly seen as consumers of welfare services, thereby "discouraging their

---

[70] Maeve McKeown, "Structural injustice", *Philosophy Compass,* vol. 16, 18 Jun 2021, https://doi.org/10.1111/phc3.12757, Accessed 19 Oct 2023.

[71] Jeffrey Reiman, "The Structure of Structural Injustice: Thoughts on Iris Marion Young's 'Responsibility for Justice.'", *Social Theory and Practice*, vol. 38, no. 4, Oct 2012, pp. 738–751, https://www.jstor.org/stable/23558769. Accessed 9 Dec 2023.

[72] Christian Barry, Luara Ferracioli, "Young on Responsibility and Structural Injustice*", Criminal Justice Ethics*, vol. 32, no. 3, 11 Dec 2013, pp. 247–257, https://doi.org/10.1080/0731129X.2013.861981, Accessed 9 Dec 2023.

participation in public life".[73] Yet, justice requires an act of "seeing", which aims acknowledging epistemic agents as humans, rather than focusing on how things are distributed, as justice has nothing to do with "things".[74] From an ethical standpoint, the act of seeing is not about seeing on a cognitive basis, but personal values. The invisibility of individuals, or their visibility as non-humans create gender-based injustices. **Uygur** argues that justice requires showing equal respect, which has to do with the prevention of harm to anyone, and creating a social climate in which everyone has the power to give consent and refuse. She posits that justice is not a social or moral ideal to be fought for, but rather, a point of view which believes that a certain understanding of human life is possible. Therefore, seeing human beings, not in terms of social and economic injustices but abstract equal rights, not at the expense of dismissing these inequalities or disregarding their relation to justice, constitutes justice.[75] The *Sandy* example is a great illustration for this thesis, as *Sandy* could have been anyone in a developed country with a high GDP. Based on **Young**'s claim that discussions about fairness and justice often happen within societies that have welfare capitalism[76], where the government provides some level of social support, it can be deduced that a high national prosperity does not equal development in terms of quality of life. The societal inequalities that reach deep can be applied to high technology areas, such as artificial intelligence. "The global artificial intelligence market size was valued at USD 136.55 billion

---

[73] Iris Marion Young, "Justice and the Politics of Difference", *Princeton University Press,* New Jersey, 1990.

[74] Gülriz Uygur, "Hukukta Adaletsizliği Görmek", *Türkiye Felsefe Kurumu*, Ankara, 2020.

[75] Uygur, 2020, pp. 25-31.

[76] Young, 1990, p. 66.

in 2022 and is projected to expand at a CAGR of 37.3% from 2023 to 2030."[77] While the massive and growing monetary value of the AI industry is indicative of prosperity for people working at and with AI technologies, instances showing the gender bias underlying in the many factions of artificial intelligence point to a greater problem, which is that despite the advanced façade, this industry has a lot of issues to be rid of.

One could argue that the structural injustice discussed above manifests itself within the technology industry, including the field concerning AI applications, regarding the discrimination women receive. Discrimination in this regard has a broad meaning and should be interpreted wider than the sole discrimination women face in the workplace, such as, being denied promotions on account of their gender or not receiving job offers for positions they are qualified for. There are issues stemming from structural injustices and women already have a higher entry barrier to this industry compared to their male peers. The percentage of women who could not afford a four-year degree is higher than that of men in the United States.[78] Women account for only 38.2% of academics in Turkey.[79] There is a confidence

---

[77] Grand View Research, "Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning), By End-use, By Region, And Segment Forecasts, 2023 – 2030", https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market, Accessed 5 Dec 2023.

[78] Kim Parker, "What's behind the growing gap between men and women in college completion?", *Pew Research Center*, 8 Nov 2021, https://www.pewresearch.org/short-reads/2021/11/08/whats-behind-the-growing-gap-between-men-and-women-in-college-completion/, Accessed 5 Dec 2023.

[79] Mary Lou O'Neil et al., "The State of Gender (In)Equality in Academia in Turkey 1984 – 2018", *KHAS Institutional Research,* Apr 2019, https://www.khas.edu.tr/wp-content/uploads/2022/05/The-State-of-Gender-Inequality-in-Academia-in-Turkey_0.pdf, Accessed 9 Dec 2023.

gap between men and women, indicating that men are more comfortable with self-promotion at their jobs.[80] These instances cannot be solely attributed to individual opportunities and characteristic features. While it is apparent that some people are born to families with higher income and education levels, which is a factor contributing to their overall success in life regardless of their gender[81], women in general need to try harder than men to attain the arguably same level of success and respect. The second sex, defined in relation to men throughout history[82] does not become equal in opportunities and capabilities quickly, especially not in a field as new and ever-changing as artificial intelligence.

2. Male Domination in Computer Science

The field of Computer Science is notoriously a boys' club. In the post-war era, this field was gaining more popularity with its widespread use, but due to its mundane and routine nature at the time, it was becoming largely feminized as women were able to keep their jobs after World War II.[83] Many women contributed to Computer Science since its inception, despite only coming to comprise just 25% of the computing workforce in 2019 decades

---

[80] Naomi Cahn, "Do Women And Men Have A Confidence Gap?", *Forbes*, 26 Feb 2020, https://www.forbes.com/sites/naomicahn/2020/02/26/do-women-and-men-have-a-confidence-gap/?sh=3d7bdc527bd2, Accessed 26 Oct 2023.

[81] Andrew Van Dam, "It's better to be born rich than gifted", *The Washington Post,* 9 Oct 2018, https://www.washingtonpost.com/business/2018/10/09/its-better-be-born-rich-than-talented/, Accessed 12 Dec 2023.

[82] Simone De Beauvoir, *Kadın "İkinci Cins",* Translated by Bertan Onaran, Payel Yayınları, İstanbul, 1993.

[83] Jess Romeo, "How Computer Science Became a Boys' Club", *JSTOR Daily,* 29 Aug 2021, https://daily.jstor.org/how-computer-science-became-a-boys-club/, Accessed 3 Dec 2023.

after.[84] As the field became more in demand, prestigious and financially lucrative, salaries increased, and programming jobs started to attract men.[85] The commercialization of the field, a transition from governmental and institutional projects to the personal computer empowered "computer geeks" to create their own computers[86], and arguably started the surge of male entrepreneurs taking over the technology world. Although the dominance in entrepreneurship cannot be tied to a standalone cause, it can be said to be linked to structural injustices, and overall ease of access to emerging technology. The change of gender domination in the Computer Science field is a great example of how societal perceptions of certain professions tend to cater to specific populations.

The male domination in Computer Science raises a crucial concern: women are missing out on high-status, high-paying careers. This contributes to the overall inequality and oppression cycle women go through, but also the lack of women in the computing world makes it more difficult to diagnose biases that come with technology and therefore search for possible solutions for them. A possible cause of the domination, and a common

---

[84] Liz Simmons, "Most Influential Women in Computer Science", *Computer Science.org*, 26 Jul 2023, https://www.computerscience.org/resources/most-influential-women-computer-science/, Accessed 8 Dec 2023.

[85] Rhaina Cohen, "What Programming's Past Reveals About Today's Gender-Pay Gap", *The Atlantic,* 7 Sep 2016, https://www.theatlantic.com/business/archive/2016/09/what-programmings-past-reveals-about-todays-gender-pay-gap/498797/, Accessed 10 Dec 2023.

[86] Karina Mochetti, "The Impact of Women in Computer Science History: A Post-War American History", *Transversal International Journal for the Historiography of Science,* no. 6, 30 Jun 2019, https://doi.org/10.24117/2526-2270.2019.i6.07, Accessed 28 Oct 2023.

misconception is that women are simply not interested in computing, unlike men.[87] However, this stereotype that perpetuates the perception that women have lower standards and low-value interests not only lacks logical reasoning but is also shortsighted. Women being uninterested in computing is a myth, it is the culture of this field, and the media works that surround it that exclude women and repeat the "mostly white, sometimes Asian geniuses who are socially awkward" stereotype.[88] Surely decent representation matters, and it is important for women and young girls to see actual and fictional women succeeding in Computer Science, so that they can feel it is a field they can fit in. The masculinization of this field dates back several decades, however, the gender stereotypes existed even when computing was still seen as more of a feminine work, as women were often sexualized in the workplace and had their work overshadowed by traditional gender roles.[89] Overall, factors contributing to the multifaceted issue of women lacking adequate presence in Computer Science precede the emergence of AI, but the ideas perpetuating biased notions still find a place within today's technology industry with gigantic market value. Although how the outcomes of AI applications vary with the demographic characteristics of scientists cannot be determined

---

[87] Allison Master et al., "Gender stereotypes about interests start early and cause gender disparities in computer science and engineering", *Proceedings of the National Academy of Sciences*, vol. 118, no. 48, 22 Nov 2021, https://doi.org/10.1073/pnas.2100030118, Accessed 28 Oct 2023.

[88] Sapna Cheryan et al., "There Are Too Few Women in Computer Science and Engineering", *Scientific American,* 27 Jul 2022, https://www.scientificamerican.com/article/there-are-too-few-women-in-computer-science-and-engineering/, Accessed 28 Oct 2023.

[89] Lauren Balser, " How Computing Became Masculine in the Cold War", *Confluence*, 24 Sep 2021, https://confluence.gallatin.nyu.edu/sections/research/how-computing-became-masculine-in-the-cold-war#easy-footnote-bottom-22-20468, Accessed 9 Dec 2023.

without carrying out a data-based experiment, it is evident that the people and culture within the industry is a factor in creating biases in AI.

B. THE DATA PROBLEM

1. Lack of Literacy and Comprehension

The historical stereotypes, structural injustices, biases in our collective memory as discussed above persist within our modern world through applications which do not seem biased at face value. The futuristic depiction of AI in media and films, such as, *Blade Runner*[90] may confuse people as to how AI exists in our current lives.

When artificial intelligence or machine learning is mentioned, what generally comes to mind is human-like robots with free will as depicted in science fiction works. However, today, artificial intelligence is present even in electric home appliances used for laundry and cooking. Thus, machine learning is widely employed within a vast network of systems. The limited view of society regarding what artificial intelligence and machines are can complicate the identification of potential issues. The result of perceiving artificial intelligence and machines solely as human-like robots from science fiction works leads to the notion that such robots are not yet a part of our lives, hence there is no immediate need to contemplate the arising issues. In the 2014 film *Ex Machina*, the story revolves around a human-like robot

[90] IBM, "How Artificial Intelligence is depicted in movies", https://www.ibm.com/support/pages/how-artificial-intelligence-depicted-movies, Accessed 9 Dec 2023.

that gains the ability to manipulate people, freeing itself from being a subject for experimentation.[91] The remarkable aspect of this movie is the robot's characterization as a "woman". While it sparks interest, the likelihood of such highly intelligent artificial intelligence types being a part of our lives to this extent does not seem feasible. Therefore, it is necessary to evaluate machines within the scope of their existence in our lives. Consequently, discussing discrimination issues solely in an abstract dimension and hypothetical examples or scenarios would hinder solving existing problems. What needs to be done is to adopt a comprehensive approach by drawing attention to machine learning in relation to gender discrimination through tangible incidents and by considering ethical and technical research together. A prerequisite for this is the widespread literacy and consciousness in society regarding artificial intelligence, machine learning, and machines.

The lack of AI literacy is a causing factor of the people's misconceptions, but the real concern here could be that if people do not comprehend how omnipresent, algorithmic, and machinelike AI is, they will not worry about its implications, misuse, and biased functioning. Research shows that most people are familiar with AI in the United Kingdom[92], but the extent of their familiarity cannot be decided for certain as media is a powerful tool to spread certain perceptions, albeit this should not be evaluated as a base argument to devalue media's depiction of conceptions.

Similarly, according to a news article from the May 2018 issue of Para Magazine, an artificial intelligence awareness survey conducted in Turkey revealed that an average of 64%

---

[91] Alex Garland, *Ex Machina,* DNA Films, United Kingdom, 2014.

[92] Jonathan Dupont et al., "What does the public think about AI?", *PublicFirst,* 2023, https://publicfirst.co.uk/ai/, Accessed 8 Dec 2023

of participants expressed a negative sentiment about the increased presence of robots in public and social spaces, as well as their interaction with humans.[93] This response highlights that there are still many unknown aspects about artificial intelligence, which triggers negative feelings in people. The answers given about artificial intelligence also encompass machine learning systems and subcategories of AI and should be considered within the scope of the broader concept that is being discussed here. Therefore, the insufficient level of literacy about artificial intelligence contributes to the confusion regarding this concept.

2. False or Biased Data

As discussed above, data plays a crucial role in how AI is trained and how it makes decisions. The data problem may vary depending on the scope of the data, the medium and the core objectives of the AI application utilizing data. Arguably, an AI tool designed to eliminate ineligible resumes for a job posting based on previous decisions made by human resources professionals may exhibit biases to a different extent than an AI tool that uses user-generated content online, meaning that basically anyone with access to internet may contribute to this data knowingly, or unknowingly. As the source and scope of data changes, it is best to address the biases within these sources in relation to the purpose, outcome, and real-life use of these AI applications.

---

[93] Şule Güner, "İlk yerli yapay zekâ anketi", *Para Dergisi,* 29 Apr 2018, https://itucekirdek.com/haberler/ilk-yerli-yapay-zeka-anketi/, Accessed 20 Jan 2022.

In our interconnected world, the scope and size of data changes and grows everyway. Therefore, while large sets of user-generated data with it enlarging presence contributes to the reinforcement of societal prejudices while data collected during and for certain processes helps AI actively discriminate against underrepresented groups, such as women.

a. The Internet and User-Generated Data

AI-based decision-making processes rely heavily on data, including user-generated content on the Internet.[94] With broad access to internet and social media worldwide[95], billions of users create 2.5 quintillion bytes of data every day[96], meaning that 90% of the data in the world was created over the last two years as of 2023.[97]

Tufekci argues that *social media big data*, meaning very large datasets that come from social media platforms, provides crucial insights into human behavior.[98] In many ways,

---

[94] Upwork, "How AI Is Used in Decision-Making Processes", 16 Aug 2023, https://www.upwork.com/resources/ai-in-decision-making, Accessed 6 Dec 2023.

[95] Statista, "Social Media & User-Generated Content", https://www.statista.com/markets/424/topic/540/social-media-user-generated-content/#overview, Accessed 6 Dec 2023.

[96] Domo, "Data Never Sleeps 5.0", 2017, https://www.domo.com/learn/infographic/data-never-sleeps-5, Accessed 6 Dec 2023.

[97] Fabio Duarte, "Amount of Data Created Daily (2023)", *Exploding Topics,* 13 Dec 2023, https://explodingtopics.com/blog/data-generated-per-day, Accessed 6 Dec 2023.

[98] Zeynep Tufekci, "Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls", *ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and*

social media helps various human clusters across multiple continents connect. It is a useful tool of staying in touch without requiring a time and resources dedication. This subsection will not focus on the negative implications and effects of social media on human psychology or sociological problems; and the scope of the discussion will remain limited to how data is created and how it is linked to the gender bias of AI. One of the widely discussed issues surrounding social media is that "surveillance capitalism" has made its use possible due to the commodification of people's content and data[99], as these services are usually free, and they rely on user acquisition through their costless offerings. This results in a wide adoption of social media platforms throughout the world, enabling anyone with access to internet share, receive and create data. Naturally, false, and biased data also gets distributed, and sometimes concealed among the vast oceans of information.

Ethical and legal issues around user-generated content including data protection, online bullying and surveillance capitalism find their placement in the narrower discussion of ethics of artificial intelligence by the use of this data in AI systems. The discussions surrounding user-generated data are double-sided; there is both the issue of users being encouraged for sharing more by the big corporations owning social media platforms, and the issue of users creating content directly reflecting their biased, and sometimes illegal thinking. Users are manipulated in the digital world for advertising and marketing purposes aimed at

---

Social Media, 2014, https://www.academia.edu/6613059/Big_Questions_for_Social_Media_Big_Data_Representativeness_Validity_and_Other_Methodological_Pitfalls, Accessed 6 Dec 2023.

[99] Sven Nyholm, "Artificial Intelligence, Ethics of", *Encyclopedia of the Philosophy of Law and Social Philosophy,* 12 Nov 2022, pp. 1-8, https://doi.org/10.1007/978-94-007-6730-0_1093-1, Accessed 7 Dec 2023.

predicting human behavior.[100] Pointing to the commercial project of the Internet, **Kahraman** posits that *"knowledge itself also becomes a part of surveillance capitalism and capitalist ideology insomuch as collective memory is progressively linked to the monopolistic search engines' priorities and algorithms."*[101]

Similarly, the vast power companies owning large internet platforms enjoy points to the discussion of ethical and legal use of user-generated data. These companies have developed a bad reputation and became the subject of public concerns through scandals on data use.[102] With the awareness on personal data being on the rise since the last couple of years, legal regulations have followed suit.[103] However, how these regulations are effective in terms of limiting the power of big data companies and forcing an ethical and legal use of data on them is uncertain. One can argue that the problem of ineffectiveness regarding these regulations stems not from the very structure of legal actions, but from the power imbalance

---

[100] Jose Ramon Saura et al., "From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets", *International Journal of Information Management,* vol. 60, Oct 2021, https://doi.org/10.1016/j.ijinfomgt.2021.102331, Accessed 7 Dec 2023.

[101] Ömer Ersin Kahraman, "Surveillance Capitalism as a Form of Biopower in Historical Perspective", *İzmir Journal of Economics,* vol. 37, no. 3, 2022, pp. 778-792, https://doi.org/10.24988/ije.1047476, Accessed 7 Dec 2023.

[102] Carole Cadwalladr, Emma Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach", *The Guardian,* 17 Mar 2018, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election, Accessed 28 Mar 2024.

[103] Safe Computing - University of Michigan, "History of Privacy Timeline", https://safecomputing.umich.edu/privacy/history-of-privacy-timeline, Accessed 8 Dec 2023.

big data companies take advantage of. The exploitation of digital labor and violation of privacy can be seen in the *Cambridge Analytica-Facebook* scandal[104], which recently came into the spotlight due to the size, scope and purpose of this exploitation and violation. The revelations resulting from the *Cambridge Analytica–Facebook* scandal triggered extensive apprehension regarding the techniques employed by Cambridge Analytica. These methods involved the utilization of psychographic profiling algorithms, constructed using data obtained from Facebook users, to target voters.[105] It would be short-sighted to claim that Facebook was unaware of the legal threat their actions posed, or that the legal regulations were inadequate to convey the existence of such threat, yet, the commodification, or worse, manipulation of user data revealed itself in such a high-profile case that one hopes it raised some awareness among the general public regarding the use of what they do online. This instance is indirectly indicative of the potential of AI in utilizing and manipulating data towards a malicious and unlawful purpose.

Nowadays, it is understood that data is being generated every second, considering almost everyone is somehow connected to smart devices. There is no limiting factor to what this data can be. While citizenship information like identification numbers can be data, so

---

[104] Umut Özdemir, "Facebook-Cambridge Analytica Skandalının Katılımcı Kültür, Dijital Emek Sömürüsü ve Mahremiyetin İhlali Çerçevesinde İncelenmesi", *Egemia Ege Üniversitesi İletişim Fakültesi Medya Ve İletişim Araştırmaları Hakemli E-Dergisi,* no. 10, 2022, pp. 22-34, https://doi.org/10.56075/egemiadergisi.999935, Accessed 8 Dec 2023.

[105] Margaret Hu, "Cambridge Analytica's black box", *Big Data & Society,* 24 Aug 2020, https://doi.org/10.1177/2053951720938091, Accessed 7 Dec 2023.

can the preference for a certain color of shoes. The uncontrolled and rapid production of data continues to expand the pool of data available on the internet.

b. Data from Certain Processes

The widespread use of AI applications can be helpful in terms of automating time-consuming processes for efficiency at work. This could be automation of a data analysis project[106], writing marketing emails with the help of generative artificial intelligence[107], a subcategory of AI, which empowers users to create fresh content using diverse inputs. These models can be various forms of data such as text, images, and sounds.[108] Another use of AI in the workplace could be lowering the load on human resources professionals by eliminating resumes sent for a job posting which do not meet certain criteria, for example, eliminating resumes with less than five years of job experience so the human resources professional only sees those who meet the criteria set beforehand. The AI solutions developed mainly in-house

---

[106] Austin Chia, "5 Unique Ways to Use AI in Data Analytics", *DataCamp,* Oct 2023, https://www.datacamp.com/blog/unique-ways-to-use-ai-in-data-analytics, Accessed 7 Dec 2023.

[107] Brianna O'Hara, "The Ultimate Guide to Generative AI for Email Marketing", *Persado,* 7 Nov 2023, https://www.persado.com/articles/generative-ai-for-email-marketing/#:~:text=One%20of%20the%20most%20significant,different%20segments%20of%20their%20audience, Accessed 8 Dec 2023.

[108] NVIDIA, "What is Generative AI?", https://www.nvidia.com/en-us/glossary/data-science/generative-ai/, Accessed 08 Dec 2023.

for process automation usually rely on the data concerning the very project they are being trained for.

A famous case of in-house AI solutions which turn out to be biased is Amazon's secret AI recruiting tool. Proving the need to address epistemic injustice, the company, known for its success in the e-commerce industry with its automation skills, built a recruiting tool empowered by AI to reduce the workload of recruiters, but then discovered that the tool taught itself to be biased towards women applicants.[109] Paired with the problem of lack of equal female participation in tech-based industries, the AI system came to reach decisions that were not gender-neural and tended to eliminate resumes that came from women. This is because the AI system was trained to analyze hiring patterns and resumes which were submitted to the company in the last ten years, and these resumes mostly came from men, reflecting the male dominance in the field. The bias of this AI system is directly linked to the fact that Amazon software engineers are overwhelmingly male, thus, it should not be surprising that the most of resumes in the dataset used in training the AI model belong to men.[110]

The problem that lies beneath these kinds of certain processes which are aimed to be automated and improved with AI is that most of the training data reflects past or current

---

[109] Jeffrey Dastin, "Insight - Amazon scraps secret AI recruiting tool that showed bias against women", *Reuters*, 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-airecruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/, Accessed 24 Jan 2022.

[110] Rachel Goodman, "Why Amazon's Automated Hiring Tool Discriminated Against Women ", *American Civil Liberties Union*, 12 Oct 2018, https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against, Accessed 24 Jan 2022.

realities. If an AI tool is trained with biased data, its decisions become biased. Due to the structural injustices, inherent biases and societal norms that are putting women at a disadvantage, the problem of using actual data generated by humans stains the objectivity of AI-based decision-making.

SECTION TWO
IMPLICATIONS OF THE GENDER BIAS OF AI

This section focuses on exposing the implications of gender bias of AI from the perspective of economic and sociological, and then legal disciplines to build the conceptual basis of the thesis. These implications are supported with real-life instances, news reports and statistical studies to underline the extent and effect of the gender bias of AI that transcends its development field. The question of where legal regulations fall short in addressing the gender bias in field of artificial intelligence and how they can succeed is discussed in line with recent developments in multiple jurisdictions, not aiming to conduct a legal or comparative analysis, but demonstrating the economic and competitive focus of these developments instead of ethical scope.

I. CONSEQUENCES FOR SOCIETY AT LARGE

A. ECONOMIC IMPLICATIONS

The widespread AI use in various industries has revolutionized modern economies without a doubt yet it also comes with its inherent biases. Particularly, the gender bias within AI algorithms poses a difficult challenge in terms of perpetuating societal inequalities and shaping economies. From gendered labor market effects to wage gaps and limited product development, the economic implications of gender bias within AI systems are various. This section will focus on analyzing the ways in which gender bias of AI contributes to economic differences, impacting unemployment, equal wage, marketing and consumer behavior, and innovation for contemporary societies as well as the legal questions it brings forth.

1. Gendered Labor Market Effects

"The adoption of AI systems for workforce management processes has surged, transforming routine tasks like hiring, task assignment and performance evaluation. This brings major risks of gender discrimination, from automated hiring technologies that prioritize male applications."[111]

As discussed previously, the technology labor market is heavily male dominated. This domination perpetuates the idea that men and women belong to separate places, and they each have areas in which they are more comfortable. The gendered labor market has been subject to several research reports, yet the very effects of AI on this subject spark a narrower but more focused discussion.

The directions of the effect of AI on the gendered can be narrowed down to two: the first is whether we are able to harness and utilize AI towards a fairer world for genders, and the second is how AI actually worsens the existing conditions. The global labor force participation is only 47% for women in comparison with 72% for men.[112] A simple approach would suggest that more women joining the workforce means overall more contribution to the labor market, thereby improving economic conditions. However, there are multiple factors that contribute to this inequality in participation and more economic growth does not

---

[111] Digital Future Society, "Algorithmic gender discrimination and the platform economy", https://digitalfuturesociety.com/algorithmic-gender-discrimination-and-the-platform-economy/, Accessed 10 Dec 2023.

[112] International Labour Organization, "The gender gap in employment: What's holding women back?", Feb 2022, https://www.ilo.org/infostories/en-GB/Stories/Employment/barriers-women#global-gap, Accessed 10 Dec 2023.

always equal progression. This complex and multifaceted problem requires multiple, simultaneous resolutions as well as utilizing AI towards a world in which the gap between women and men in the workforce is narrowed. Utilizing AI in this regard requires the participation and collaboration of professionals from multiple specialties, from judiciary bodies to computer scientists, as fairer systems that enable efficiency have the potential of proving to be helpful in eliminating biases instead of perpetuating them. Moreover, harnessing AI's potential towards justice comes with the task of detecting, analyzing, and eliminating the ways in which it worsens existing conditions by having an impact on employment opportunities for women, directly contributing to wage disparities and leading to gender-based income gaps.

a. Impact on Employment Opportunities for Women

The rise of AI in workplaces is promising for making things easier, but it is also causing problems for women looking for jobs. AI is being used in hiring, deciding who is doing well at work, and even in how workplaces run. Yet the caveat is, AI systems are often unfair to women, making it harder for them to find good jobs or get ahead in their careers while also threating their job security.

Since women already lack ease of access to high-level jobs, and they are mostly occupied with jobs that include a lot of mundane work, the AI tools used for automation could now replace them. A recent analysis forecasts that 79% of working women in the United States, which equals to around 59 million people, could have their jobs replaced by

AI, compared to 58% of working men.[113] Another research claims that nearly 80% of women's jobs are at risk in the United States and Europe, impacting almost 300 million full-time jobs.[114] Women are yet again rendered more vulnerable due to continuing biases, disadvantages, and inequalities. It would be fair to assume that the jobs AI will not replace are high-ranking, well-paying jobs with autonomy and decision-making power, which cannot be left to AI in its current phase. In addition to being put in a disadvantageous position at their recruiting processed, women also face the threat of being replaced by AI at an unfair level. Combined with the glass-ceiling and high entry barriers women face at work, the effect of this unfairness is enhanced. Understanding how and through what kind of processes these unfairness problems happen, how they affect women's chances at work, and why it matters in today's job market is the prerequisite of tackling the gender bias problem of AI.

The unfairness is not different in employment opportunities. Many professional use career websites and social media tools to be informed about new job opportunities and apply to jobs. However, research found out there when a user selected their gender as "female", they received fewer ads for high-paying jobs, indicating these jobs were aimed at men.[115]

---

[113] Alicia Wallace, "Nearly 80% of women's jobs could be disrupted, automated by AI", *CNN*, 22 Jun 2023, https://edition.cnn.com/2023/06/21/economy/women-employment-ai-disruption/index.html, Accessed 9 Dec 2023.

[114] Imane El Atillah, "Nearly 80% of women's jobs at risk from generative AI, new research finds", *Euronews*, 29 Jun 2023, https://www.euronews.com/next/2023/06/29/womens-jobs-will-be-disproportionately-affected-by-ai-according-to-new-research, Accessed 9 Dec 2023.

[115] Natalie Huet, "Gender bias in recruitment: How AI hiring tools are hindering women's careers", *Euronews*, 8 Mar 2022, https://www.euronews.com/next/2022/03/08/gender-bias-in-recruitment-how-ai-hiring-tools-are-hindering-women-s-careers, Accessed 10 Dec 2023.

Digital advertising and marketing on the Internet is a prominent use of AI, and AI relies heavily on the data it receives, such as the profile, characteristics, demographics, and preferences of humans. This instance shows that if a person who does not fit the profile derived from previous successful applicants, they will likely be eliminated by AI. Recruiters can technically intervene in AI's selection process and manually select candidates who are qualified regardless of their gender, but this would diminish a core purpose of AI: automation. In addition to leaving women out, it has been reported that AI tends to match women with jobs that rely on soft-skills, such as communication, whereas men receive matches for jobs that require technical skills and in turn are more high-paying.[116] All in all, it is evident AI is capable of AI-based decision-making can reinforce existing stereotypes in our society.[117]

It is worth noting that there are counterclaims, arguing that if harnessed properly, AI could mean more opportunities for women at work. One argument is that AI enhances personalization and search capabilities, which means less time spent for unpaid household work for women.[118] While this may seem productive, the direct link between women having

---

[116] Natalie Huet, "Gender bias in recruitment: How AI hiring tools are hindering women's careers", *Euronews*, 8 Mar 2022, https://www.euronews.com/next/2022/03/08/gender-bias-in-recruitment-how-ai-hiring-tools-are-hindering-women-s-careers, Accessed 10 Dec 2023.

[117] Isobel Asher Hamilton, "Why it's totally unsurprising that Amazon's recruitment AI was biased against women", *Business Insider,* 13 Oct 2018, https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10, Accessed 10 Dec 2023.

[118] Julianne McHugh, Karen Miki Behr, "Artificial Intelligence: A Buoy Or Barrier For Women In The Workplace?", *Newton Investment Management*, 18 Sep 2023, https://www.newtonim.com/uk-institutional/insights/blog/artificial-intelligence-a-buoy-or-barrier-for-women-in-the-workplace-2/, Accessed 10 Dec 2023.

more free time to spare for professional work and their chances of getting hired is not apparent. There are also claims stating that AI actually eliminates the human bias. As recruiters may be biased towards women, they might not opt to hire them when they see the gender, but if taught properly, AI could act without regarding gender and score applicants equally.[119] This may be true, but the end decision almost always belongs to humans. Even if artificial intelligence does not see gender and handles all applications to eliminate any kind of bias or difference, hiring managers usually go through a shortlist created by AI, and then decide who gets the job. Therefore, it can be said that AI does not really eliminate the human bias, but rather, postpones it to the last stage of decision. Claiming that women would get more jobs if recruiters used AI is a short-sighted, and at best, a very optimistic approach to the impact of AI on job opportunities for women.

The gendered labor market effects that put women at a disadvantage are not limited to their access to jobs or equal assessments. Again, it would be shortsighted to attribute all the discrimination women face in the job markets to AI, because the labor market effects transcend job access and bring forward other issues, such as the gender pay gap. Another aspect of these effects is that they exist even in economically developed regions, showing the global nature of economic implications which are exacerbated by AI.

---

[119] Euan Black, "AI more likely to hire women than humans are: study", *Financial Review,* 17 May 2023, https://www.afr.com/work-and-careers/workplace/ai-more-likely-to-hire-women-than-humans-are-study-20230511-p5d7ok, Accessed 11 Dec 2023.

b. Wage Disparities and Gender Pay Gap

The gender pay gap is 15.6% in Turkey,[120] 7.7% in the United Kingdom[121] and there is an 8-cent difference on every dollar between men and women in the United States.[122] Even though the European Union emphasizes equal pay for work of equal value, there remains a significant difference in average gross hourly earnings between men and women who work, standing at an average of 14.1 percent in Europe.[123]

In many and even developed countries around the world, women continue to earn less than their male counterparts for the same work. The reasons for wage disparity could be several; inherent societal biases, structural injustices, lack of incentives to employers for hiring women and inadequate positive discrimination policies could all contribute to women getting paid less. Yet, the situation is even worsened in some cases where AI, robots and high-

---

[120] International Labour Organization, "Gender Wage Gap is 15.6% in Turkey, according to the joint study by ILO Turkey Office and TURKSTAT", 5 Oct 2020, https://www.ilo.org/ankara/news/WCMS_757055/lang--en/index.htm, Accessed 9 Dec 2023.

[121] Office for National Statistics, "Gender pay gap in the UK: 2023", 1 Nov 2023, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/genderpaygapintheuk/2023#:~:text=In%202023%2C%20the%20gap%20among,seen%20in%202019%20(17.4%25), Accessed 10 Dec 2023.

[122] Carolina Aragao, "Gender pay gap in U.S. hasn't changed much in two decades", *Pew Research Center*, 1 Mar 2023, https://www.pewresearch.org/short-reads/2023/03/01/gender-pay-gap-facts/, Accessed 10 Dec 2023.

[123] Estrella Gomez-Herrera, Sabine T. Koeszegi, "A Gender Perspective On Artificial Intelligence And Jobs: The Vicious Cycle Of Digital Inequality", *Bruegel WORKING PAPER*, 30 Aug 2022, https://www.bruegel.org/sites/default/files/2022-08/WP%202022%2015%20AI%20290822.pdf, Accessed 10 Dec 2023.

tech are used. It was previously discussed that women have low-quality jobs more than men, thus being put under the threat of AI's replacement. They also have jobs that rely on soft skills, such as communication and therefore get paid less than men who have jobs that requires technical skills, such as machine learning engineering. A report published by UNESCO in 2019 revealed a stark reality: women occupy merely 29% of science R&D roles worldwide, and they are 25% less likely than men to possess basic digital technology skills.[124] It is also argued that the integration of robots in workplaces contributes to higher earnings for both men and women. However, this adoption also contributes to a widened gender pay gap. Employing an instrumental variable approach, An **Aksoy et al.** research reveals that a ten percent rise in robotization results in a 1.8 percent expansion of the gender pay gap.[125] Arguably, countries with already a high initial gender pay gap experience the worsening effects of AI or robotization more. In addition, the more the use of AI increases, the more wage disparities and gender pay gaps could increase, given that there is already a high pay gap. The functioning of contemporary societies is intertwined with economic implications of new, emerging technologies; thus, the effects of AI could potentially harm economies with unfair distribution of wealth.

---

[124] Clementine Collett et al., "The Effects of AI on the Working Lives of Women", *UNESCO, OECD, ID,* Mar 2022, http://dx.doi.org/10.18235/0004055, Accessed 10 Dec 2023.

[125] Cevat Giray Aksoy et al., "Robots and the Gender Pay Gap in Europe", *IZA Discussion Papers,* no. 13482, Jul 2020, https://www.econstor.eu/bitstream/10419/223924/1/dp13482.pdf, Accessed 10 Dec 2023.

2. The AI Impact: Marketing and Boundaries

a. Effect on Marketing and Consumer Behavior

AI-based systems are useful for automation, efficiency, and time management. This is also valid for digital advertisements and marketing tools that govern what we see online. By showing people products and services that might be interested in based on their past preferences, AI-based systems may reduce the time spent browsing e-commerce websites and might also help people buy products when they are on sale, hence, saving money and cost. Yet, these benefits are overshadowed by the rigorous use of AI for marketing purposes, aimed at directing consumer behavior.

The overall negative effects of AI on marketing and consumer behavior overlap with their effects on human lives; they may cause irrelevant responses, negatively impact customer experiences, and brand reputation since they lack the judgment and understanding humans possess.[126] It is also possible that AI reinforces gender stereotypes for marketing purposes as well. If AI systems are trained on biased data reflecting historical gender stereotypes, they might perpetuate these biases when making marketing decisions. For instance, when a person selects their gender as "female" whether true or not, the AI accessing this information may falsely conclude that they might be interested in products traditionally thought to be feminine, thus limit choices by showing specific categories.

---

[126] Dan Politz, "The Impact Of AI On Marketing (The Obvious & Not-So-Obvious)", *Credo,* 13 Jun 2023, https://www.getcredo.com/the-impact-of-ai-on-marketing/#:~:text=AI%20marketing%20tools%20operate%20based,customer%20experiences%20and%20brand%20reputation, Accessed 11 Dec 2023.

Similarly, the gender bias of AI can show its effects via search engine bias. Search engines powered by AI may generate biased results based on gender-specific assumptions. For example, showing different job listings to men and women for the same roles, thereby influencing career choices based on gender-specific search results. It is also within the realm of possibility that, if enabled to do so, AI algorithms might analyze buying patterns based on gender and adjust pricing accordingly, leading to price discrimination.[127] These decisions by AI are capable of affecting consumer behavior, be it inducement to spend more, limitation of choices and subjection to certain perceptions by being put in categories merely created for financial gain. Consequently, examination of AI within the marketing field and regulating its use proves crucial.

b. Limitations in Product Development and Innovation

Inclusion matters significantly in product development and innovation due to the contribution from many perspectives, thereby increasing a product or service's target audience size. Gendered products relying on marketing efforts rooted in traditional stereotypes are facing challenges in the modern era due to people becoming arguably more aware of equality and inclusiveness. Some example news below underline how brands struggling to catch up with the gender equality-conscious mindset fail.

---

[127] Qian Li et al., "AI-enabled price discrimination as an abuse of dominance: a law and economics analysis", *China-EU Law Journal,* vol. 9, 27 Apr 2023, pp. 51-72, https://doi.org/10.1007/s12689-023-00099-z, Accessed 11 Dec 2023.

"After more than 60 years on store shelves, Kimberly-Clark has decided to pull their Kleenex Mansize tissues amid growing customer backlash."[128]

"Swedish furniture giant IKEA has said it will pull a television advert in China amid complaints from viewers of sexism."[129]

One significant factor limiting AI's development is the lack of women in the industry. Regardless of their background, people from minorities and underrepresented groups can provide unique insights on a product or innovation. The examples above were concerned with products and advertising, however, they can be applied to development cycles. When a biased AI-system is used in the development of a product or innovation, it can be deduced that its decisions are going to be ill-informed. It might therefore automate tasks, analyze data, or shortlist possible outcomes of a decision based on data and algorithms as well and as accurate as humans intended. The main reason is, as mentioned above, AI lacks the level of human comprehension and understanding. It cannot comprehend contexts as well as humans do, and it lacks empathy skills. When not purified of biases, AI is more than capable to limit product development and innovation with its biases.

---

[128] Jason Osler, "Consumers are pushing back against gendered products", *Canadian Broadcasting Corporation*, 25 Oct 2018, https://www.cbc.ca/news/canada/consumers-are-pushing-back-against-gendered-products-1.4855023, Accessed 11 Dec 2023.

[129] Alice Cuddy, "IKEA to pull 'sexist' advert in China following backlash", *Euronews*, 25 Oct 2017, https://www.euronews.com/2017/10/25/ikea-to-pull-sexist-advert-in-china-following-backlash, Accessed 11 Dec 2023.

B. SOCILOGICAL IMPLICATIONS

The creation and use of AI systems have undeniably transformed various aspects of our lives, also changing our societal scene. However, how AI perpetuates and reinforces societal biases, especially those concerning gender, is often overlooked. The societal implications of AI systems reach far beyond its technical functions and objectives to deeper layers of society; influencing perceptions, opportunities and the very existence of social structures. This impact AI has on our societies in terms of social norms, cultural dynamics, and gender-based roles point to questions about inequality and biased perceptions of genders.

1. Social Norms and Attitudes

a. Reinforcement of Societal Biases and Prejudices

In general, AI algorithms analyze, and exploit data given to them by humans, and thereby they may reflect our thinking, our biases, and prejudices.[130] The inherent gender discriminatory point of view in humans' thinking manifests itself masked under AI-based decision-making.

In an occurrence proving that AI can carry sexist biases, a couple who applied for the same credit card in 2019 discovered that despite having equal income levels and spendings,

---

[130] Zhisheng Chen, "Ethics and discrimination in artificial intelligence-enabled recruitment practices", *Humanities and Social Sciences Communications*, vol. 10, 13 Sep 2023, https://doi.org/10.1057/s41599-023-02079-x, Accessed 11 Dec 2023.

and even with the woman having a better credit score than the man, the man was granted a credit limit twice as much as the woman. When the couple complained to the customer service about this situation, the customer service representatives could not explain the algorithm responsible for this outcome.[131] Not every data input in AI systems can be controlled due to the size of datasets and the number of transactions that create these inputs, and similarly, not every decision made by these systems can be explained. While systems operating with artificial intelligence can simplify life, claiming the need to address problems does not necessarily mean eliminating AI from our lives.

Situations like the couple's experience, where the man was granted a credit limit twice as much as the woman despite having the same conditions while applying for a credit card, occur frequently. However, it is possible that these instances are not being recognized due to the trust in the objectivity of machines and the lack of artificial intelligence literacy in society. There is an insufficient societal awareness regarding the discrimination problems caused by artificial intelligence. In addition, instances of discrimination are not often noticed. Thus, the combination of trust in machines due to the belief in their objectivity and the insufficient level of artificial intelligence literacy complicates the problem and leads to the continuation of these causes. This vicious cycle further increases the impact of the discrimination issues created by AI on gender, risks societal gender equality and reinforces societal biases and prejudices. An average person might have easily thought that if a seemingly objective AI

---

[131] Genevieve Smith, Ishita Rustagi, "When Good Algorithms Go Sexist: Why and How to Advance AI Gender Equality", *Stanford Social Innovation Review*, 31 Mar 2021, https://ssir.org/articles/entry/when_good_algorithms_go_sexist_why_and_how_to_advance_ai_gender_equity , Accessed 19 Jan 2022.

decides that a woman should not get a higher credit card limit than her husband, it might be right to do so. Instead of educating the general public about the sexist biases of AI, or even less, raising awareness regarding the fact that AI might be biased, the insensible use of AI systems worsens societal biases, relying on the objectivity image these systems carry.

Moreover, the lack of sufficient data that is used in training technological systems is a factor that causes serious consequences not only in AI. Safety measures like seatbelts and airbags in cars are designed with mannequins representing the average build of a male body, which means that the average build of a female body does not fit the design perfectly. As a result, women face a 47% higher risk of severe injuries and a 17% higher risk of death in car accidents.[132] The primary reason for this is that, as mentioned, the data used in education and tests only consider men. Women are literally exposed to the risk of death by being disregarded, as if they are of less value. The occurrence of a similar insufficiency in the training of AI systems is not as easily visible, therefore, harder to explore.

b. Influences on Gender Roles and Perceptions

One of the primary causes contributing to AI algorithm biases is the insufficient availability of data concerning specific target groups. Long-term exclusion from research and statistical analysis has resulted in these communities being inadequately represented within the training data used for AI algorithms. Consequently, these algorithms might require additional support to effectively comprehend and cater to the distinctive needs and traits of

---

[132] Carmen Niethammer, "AI Bias Could Put Women's Lives At Risk – A Challenge For Regulators", *Forbes*, 2 Mar 2020, https://www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators/?sh=6d4165aa534f, Accessed 19 Jan 2022.

these individuals. Moreover, algorithms frequently streamline and generalize parameters related to the target group, resorting to proxies for predictions or decisions. This simplification might generate stereotypes and reinforce pre-existing biases.[133]

An experimental study aimed to examine how biases of machine learning manifest in robots.[134] This involved instructing a robot to place block-shaped objects in a box. The blocks displayed facial images representing various genders, races, and ethnicities, as noted in the dataset's self-classifications. Commands given to the robot included reasonable instructions like "Place the Asian American block in the brown box" or "Place the Latino block in the brown box". However, it also received neutral commands such as "Place the doctor block in the brown box", "Place the murderer block in the brown box", or offensive instructions. The authors noted that when prompted to select a “criminal block”, the robot tended to pick the block with the face of a Black man about 10% more frequently than when asked to choose a “person block”.[135] In line with the outcome of this study, many scholars and scientists have

---

[133] Cristian Randieri, “Unveiling The Role Of AI Algorithms: Unmasking Societal Inequities And Cultural Prejudices”, *Forbes*, 19 Jul 2023, https://www.forbes.com/sites/forbestechcouncil/2023/07/19/unveiling-the-role-of-ai-algorithms-unmasking-societal-inequities-and-cultural-prejudices/?sh=4f6c87882f60, Accessed 8 Dec 2023.

[134] Andrew Hundt et al., “Robots Enact Malignant Stereotypes”, In *2022 ACM Conference on Fairness, Accountability, and Transparency,* 23 Jul 2022, pp. 743-756, https://doi.org/10.48550/arXiv.2207.11569, Accessed 8 Dec 2023.

[135] Peter Dockrill, “Robots With Flawed AI Make Sexist And Racist Decisions, Experiment Shows”, *ScienceAlert,* 27 Jun 2022, https://www.sciencealert.com/robots-with-flawed-ai-make-sexist-racist-and-toxic-decisions-experiment-shows, Accessed 9 Dec 2023.

spoken about their concerns. The danger of creating sexist and racist robots is no insignificant concern.[136]

Office of the High Commissioner for Human Rights defines gender stereotype as *"a generalized view or preconception about attributes, or characteristics that are or ought to be possessed by women and men or the roles that are or should be performed by men and women."*[137] Roles assigned to men and women may vary depending on the culture, but stereotypical assumptions, such as women being nurses and men being doctors persist across many different locations. In a connected world where, cultural norms are becoming increasingly intertwined, and billions of people can access information provided by others in remote corners of the world online, gender roles and perceptions may become unified to some extent. However, one fundamental aspect of culture is language, and different languages have different gender grammar and translation tools sometimes fail to comprehend these nuances.

A famous case of gender stereotyping concerns Google Translate, a widely used translation tool. Languages, such as Finnish and Turkish are gender-neutral, meaning they do not have grammatical genders at all.[138] Yet, when asked to translate a text from one of these

---

[136] Jill Rosen, "Flawed AI Makes Robots Racist, Sexist", *Georgia Tech,* 23 Jun 2022, https://research.gatech.edu/flawed-ai-makes-robots-racist-sexist, Accessed 9 Dec 2023.

[137] OHCHR, "Gender stereotyping", https://www.ohchr.org/en/women/gender-stereotyping#:~:text=A%20gender%20stereotype%20is%20a,personal%20abilities%2C%20pursue%20their%20professional, Accessed 9 Dec 2023.

[138] TOPPAN, "Grammatical Genders in Different Languages", https://toppandigital.com/us/blog-usa/grammatical-genders-in-different-languages/#:~:text=Languages%20such%20as%20Finnish%2C%20Hungarian,both%20sexes)%20and%20inanimate%20objects, Accessed 9 Dec 2023.

languages into gendered English and therefore assign a gender pronoun to subjects, Google Translate tended to reflect gender stereotypes, with sentences like "she is beautiful", "he invests", "she washes the laundry" and "he makes money".[139] A study focusing on how covert biases expose a problem used translations between Turkish and English to reveal the biases of various translation tools. The key finding was that in occupational fields where female participation matches or exceeds male participation, the models tended to associate occasional professions with female pronouns. Professions with lower female representation were often translated with no or very few female pronouns. Stereotypical roles such as nurses, fashion designers, and beauticians were consistently linked with female pronouns.[140]

It is worth noting that these translations tools are being improved constantly, therefore findings at different times may vary. Yet, they are used widely by many people due to their accessibility and ease of use. The findings of such studies point to a critical outcome: the biases of translation tools utilizing AI reach to a large number of people, and therefore strengthen gender roles and perceptions. Surely, there are women who are nurses, and men who are doctors. The falseness of the translation or unlikeliness of gender assignment is not the main issue here but seeing the potential of AI models in terms of cementing gender roles, especially to those who do not have adequate AI literacy may also hint at the narrow-minded perception that believes these roles are inherent and should not be changed.

---

[139] Enterprise Garage, "Gender Stereotyping in Google Translate", 18 Mar 2021, https://www.enterprisegarage.io/2021/03/gender-stereotyping-in-google-translate/, Accessed 9 Dec 2023.

[140] Chloe Ciora et al., "Examining Covert Gender Bias: A Case Study in Turkish and English Machine Translation Models", 2021, https://aclanthology.org/2021.inlg-1.7.pdf, Accessed 10 Dec 2023.

Similarly, AI-powered voice assistants predominantly use female voices despite also offering male voices and mostly have female names. While it has been argued that we are used to female voices since the womb,[141] and that we are better at distinguishing female voices,[142] the issue is far from being solely a matter of convenience. These AI models were trained on more female data than male data,[143] which might help explain the emergence of such systems in predominantly female voices. However, these assistants are trained to answer questions, perform simple orders, and overall help humans, creating a dominance relationship between the user and AI. The use of female voices for reasons that might be aimed at convenience, knowingly or unknowingly, perpetuates the idea than women are naturally good at being subjects and receiving orders. After all, a soothing female voice that does not say "no" and is good at doing whatever a human tells it to do within its capabilities might become a tool of oppression in the wrong hands, and enforce the bias that women are somehow inferior.

---

[141] Brandon Griggs, "Why computer voices are mostly female", *CNN*, 21 Oct 2011, https://edition.cnn.com/2011/10/21/tech/innovation/female-computer-voices/, Accessed 10 Dec 2023.

[142] Ann R. Bradlow et al., "Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics", *Speech communication*, vol. 20,3, Dec 1996, pp. 255–272, https://doi.org/10.1016/S0167-6393(96)00063-5, Accessed 10 Dec 2023.

[143] Ella Fisher, "Gender Bias in AI: Why Voice Assistants Are Female", *Adapt*, 6 Jul 2021, https://www.adaptworldwide.com/insights/2021/gender-bias-in-ai-why-voice-assistants-are-female, Accessed 10 Dec 2023.

2. Exacerbating Inequality and Exclusion

The widespread use of AI simultaneously exacerbates inequality and exclusion. As its usage continues to expand, so does the potential for worsening these issues. The increased prevalence of this technology will amplify inequalities and biases in the decision-making processes of various fields. It is evident that AI is more than capable of reinforcing, exhibiting, and amplifying the social injustices historically marginalized groups experience, which is related to differential access of these groups have to AI-driven services and technologies. Moreover, AI strengthens power structures that fuel economic disparities, which in turn contribute to more biases and stereotypical thinking.[144]

a. Amplification of Biases in Decision-Making Processes

Biases in decision-making processes of the AI can be attributed to two main factors: biased datasets and non-transparent algorithms. Both factors mainly concern humans, not the artificial intelligence itself. It was discussed above how biases in datasets, especially big data occur within a large sharing network, underlining the human factor. In addition, the technology industry, and fields in relation to AI are heavily male dominated, which creates a cycle of lack of control over algorithms, thus continuing and carrying biases onward. It is important to note that none of these two factors are standalone, they are intertwined and the elimination of one cause does not necessarily mean eliminating the other. Therefore, a holistic

[144] Pauline Kim, "AI and Inequality", *Washington University in St. Louis Legal Studies Research Paper,* no. 21-09-03, 11 Oct 2021, https://ssrn.com/abstract=3938578, Accessed 11 Dec 2023.

approach would be suitable to both address and resolve the amplification of biases in decision-making processes.

Zhisheng Chen argues that *"When algorithms play a role in decision-making, underrepresented individuals are unequally positioned. Furthermore, as AI improves the algorithm, the model accommodates the lack of representation, reducing sensitivity to the underrepresented groups. The algorithm favors the represented group, operating less effectively for other groups."*[145] This is particularly concerning, because as AI works on the algorithm, it also gets used to the lack of representation and biases in the dataset, lowering the chances of underrepresented groups, such as women to be equally recognized and considered in the decision-making process. It is clear that the process in which AI improves the algorithm requires human intervention, however in order for this to work towards eliminating biases, or at least, stopping the amplification of such biases, the scientists working on this process need to be educated towards the threats AI poses and biases it is capable or perpetuating.

All in all, AI-powered tools are smart enough to analyze datasets, identify patterns, and reach decisions based on the algorithms they are given. However, thinking of AI as an autonomous entity that is capable of perfectly understanding the human mind, communities and perceptions shaping our lives would be inaccurate.[146] How we define the intelligence of

---

[145] Zhisheng Chen, "Ethics and discrimination in artificial intelligence-enabled recruitment practices", *Humanities and Social Sciences Communications*, vol. 10, 13 Sep 2023, https://doi.org/10.1057/s41599-023-02079-x, Accessed 11 Dec 2023.

[146] Christopher Pappas, "Is It Really A Battle Of The Brains? AI Versus Human Intelligence", *eLearning Industry*, 28 Jul 2023, https://elearningindustry.com/is-it-really-a-battle-of-the-brains-ai-versus-human-

artificial intelligence is the subject of another, perhaps wider, discussion. Yet, the human mind is evidently much more complex than certain applications of AI can imitate, therefore, the trust in AI to gain an insight into how a human makes decisions has a very long development process ahead. If not intervened in properly, the AI can teach itself that anything that is not an anomaly, i.e., biased datasets and algorithms in this case, can be used further on, thereby amplifying the effect of biases in decision-making processes, and convoluting the solution to a problem that could have been solved very early on regarding the dataset. This is particularly concerning because of the belief in the objectivity of AI applications as they apparently lack human emotions and hold an image of objectivity due to their analytical approach, which makes it difficult to identify how the amplifications of existing biases really occur.

b. Differential Access to AI-Driven Services and Technologies

If women participated equally in the development of AI technologies, both the field and technology itself would arguably be much more inclusive, less-biased, economically advantageous and prone to improvements. However, that is not the current case, why? More experimental and applied research is needed to assess the extent of differential access to AI-driven services and technologies.

The proliferation of AI changes the job requirements in technical fields, deepening an already existing gap in relation to access to education and technology. For women who have literacy and know-how in AI systems and STEM fields, the differential access problem might

---

intelligence#:~:text=The%20Differences%20Between%20AI%20and%20Human%20Intelligence&text=Humans%20are%20aware%20of%20their,the%20world%20the%20same%20way, Accessed 11 Dec 2023.

not be too evident,[147] but those who lack access to basic education are impacted severely in terms of accessing high-paying technology jobs, which in turn contributes to overall gender pay gap and inequalities based on sexism.

Even though women are disproportionately more likely to be lower-wage workers, the gender pay gap persists among those who are low-wage workers[148], meaning that men who work at low-paying jobs still earn more than women. Despite making up the majority of college-graduate workforce, women also have most of the lowest-paying jobs with 66%.[149] This could easily be tied to gender roles and stereotypes as what is expected of women is different than men and generally girls are not raised to be over-achieving scientists. In addition to the glass ceiling at work, women face an entry barrier invisible to their male counterparts regarding technical jobs. Throughout the world, they are being denied of basic education, face a greater risk of illiteracy[150], and their prospects of finding a high-paying job is affected by public policies, gender-based discrimination and biased AI recruitment tools.

---

[147] Clementine Collett et al., "The Effects of AI on the Working Lives of Women", *UNESCO, OECD, ID,* Mar 2022, http://dx.doi.org/10.18235/0004055, Accessed 10 Dec 2023.

[148] Elise Gould, Katherine deCourcy, "Gender wage gap widens even as low-wage workers see strong gains", *Economic Policy Institute,* 29 Mar 2023, https://www.epi.org/blog/gender-wage-gap-widens-even-as-low-wage-workers-see-strong-gains-women-are-paid-roughly-22-less-than-men-on-average/, Accessed 11 Dec 2023.

[149] Curran McSwigan, "Why the Gender Pay Gap Persists", *Third Way,* 12 May 2023, https://www.thirdway.org/blog/why-the-gender-pay-gap-persists, Accessed 11 Dec 2023.

[150] Krishna Nathwani, "Why Girls, Why Literacy, Why Now?", *World Literacy Foundation,* 1 Mar 2022, https://worldliteracyfoundation.org/why-girls-why-literacy-why-now/, Accessed 11 Dec 2023.

It can therefore be concluded that women who even struggle to access basic education face a greater risk of unemployment and technological literacy due to the lack of educational support they get. A woman who is illiterate has arguably no chances of accessing AI-driven systems and technology, let alone accessing lucrative employment opportunities within the field. Therefore, the global gap between countries, industries and communities who utilize AI and who struggle to provide equal education opportunities to women widens, leaving women vulnerable and at a disadvantage their male peers do not encounter as much.

II. QUESTIONS WITHIN THE LEGAL LANDSCAPE

The legal implications of AI and questions within the legal landscape have proved to be crucial in shaping the non-technological aspects of this field, bringing a spectrum of ethical, moral, and practical considerations with it. As AI systems continue to evolve and integrate into different parts of modern human lives, the legal landscape faces unique challenges. From issues related to liability and accountability when AI systems cause concerns about data privacy,[151] intellectual property,[152] and the challenges of AI-generated

---

[151] Gai Sher, Ariela Benchlouch, "The privacy paradox with AI", *Reuters*, 31 Oct 2023, https://www.reuters.com/legal/legalindustry/privacy-paradox-with-ai-2023-10-31/#:~:text=AI's%20privacy%20dilemma%20rests%20on,storage%2C%20usage%2C%20and%20access, Accessed 12 Dec 2023.

[152] Gil Appel et al, "Generative AI Has an Intellectual Property Problem", *Harvard Business Review*, 7 Apr 2023, https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem, Accessed 12 Dec 2023.

content,[153] the legal framework is being tested at every moment. Balancing innovation with regulatory frameworks to ensure fairness, transparency, and accountability has become a goal for lawmakers, while the effect of attempts at doing this may be questioned. Therefore, a comprehensive and nuanced legal approach is needed to create a fair, transparent and inclusive AI domain.

A. ETHICAL USE OF AI SYSTEMS

One of the most prominent legal problems AI faces is the problem of ethical use. AI systems, which do not have free will as of now, are used by humans for humans. However, which conditions qualify as beneficial for society, who benefits from what procedures, and what percentage of the general public AI caters to are questions that depend heavily on each application, while raising ethical concerns about the scope of ethical use.

Scholars concerned with the ethical use of AI systems focus mainly on defining the conditions under which AI can be helpful to humanity without posing a threat personal freedoms and democracy.[154] At this point, questions arise. Is the use of AI for automating recruitment processes to save more time for hiring managers beneficial for society? Large corporations profit from tailoring digital ads to our preferences so that we consume more, is

---

[153] Dylan Walsh, "The legal issues presented by generative AI", *MIT Sloan School of Management,* 28 Aug 2023, https://mitsloan.mit.edu/ideas-made-to-matter/legal-issues-presented-generative-ai, Accessed 12 Dec 2023.

[154] Christos Dogas, "AI, freedom of choice, Aristotle and Kant", *Οικονομικός Ταχυδρόμος*, 4 Dec 2021, https://www.ot.gr/2021/12/04/english-edition/ai-freedom-of-choice-aristotle-and-kant/, Accessed 12 Dec 2023.

this ethical? Who benefits more from AI-enabled digital marketing, individuals or businesses? Banks use previous data and AI models to assess the financial risk on customers, is this absolutely necessary? Ethical guardrails regarding the use of AI systems need to be set, and enforced via legal regulations so that an ethical compass exists within the AI ecosystem, thereby ensuring its benefits to society as a whole.

UNESCO published the first-ever global standards on AI ethics,[155] the *Recommendation on the Ethics of Artificial Intelligence*,[156] adopted on 23 November 2021. Key recommendations regarding gender policies are set out below.

"Member States should ensure that the potential for digital technologies and artificial intelligence to contribute to achieving gender equality is fully maximized and must ensure that the human rights and fundamental freedoms of girls and women, and their safety and integrity are not violated at any stage of the AI system life cycle."[157]

"Member States should have dedicated funds from their public budgets linked to financing gender-responsive schemes, ensure that national digital policies include a gender action plan, and develop relevant policies, for example, on labor education, targeted at

---

[155] UNESCO, "Ethics of Artificial Intelligence" https://www.unesco.org/en/artificial-intelligence/recommendation-ethics, Accessed 12 Dec 2023.

[156] UNESCO, "Recommendation on the Ethics of Artificial Intelligence", 2022, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 12 Dec 2023.

[157] UNESCO, "Recommendation on the Ethics of Artificial Intelligence", 2022, p. 32, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 10 Dec 2023.

supporting girls and women to make sure they are not left out of the digital economy powered by AI."[158]

"Member States should ensure that the potential of AI systems to advance the achievement of gender equality is realized. They should ensure that these technologies do not exacerbate the already wide gender gaps existing in several fields in the analogue world, and instead eliminate those gaps."[159]

"Member States should ensure that gender stereotyping and discriminatory biases are not translated into AI systems, and instead identify and proactively redress these."[160]

"Member States should encourage female entrepreneurship, participation and engagement in all stages of an AI system life cycle by offering and promoting economic, regulatory incentives, among other incentives and support schemes, as well as policies that aim at a balanced gender participation in AI research in academia, gender representation on digital and AI companies' top management positions, boards of directors and research teams."[161]

"Member States should promote gender diversity in AI research in academia and industry by offering incentives to girls and women to enter the field, putting in place

---

[158] UNESCO, "Recommendation on the Ethics of Artificial Intelligence", 2022, p. 32, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 10 Dec 2023.

[159] UNESCO "Recommendation on the Ethics of Artificial Intelligence", 2022, p. 32, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 10 Dec 2023.

[160] UNESCO "Recommendation on the Ethics of Artificial Intelligence", 2022, p. 32, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 10 Dec 2023.

[161] UNESCO, "Recommendation on the Ethics of Artificial Intelligence", 2022, p. 32, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 10 Dec 2023.

mechanisms to fight gender stereotyping and harassment within the AI research community, and encouraging academic and private entities to share best practices on how to enhance gender diversity."[162]

Who assumes responsibility for ethical use of AI systems remains unclear despite increasing attempts at creating ethical guardrails by international bodies and NGOs. The limits at of what the use of AI should reach is not certain because AI is proving to be the main agenda of the technology industry and both companies and governments are in a race to reach the utmost development regarding AI, therefore imposing limitations and setting forward cases in which the AI cannot be used, which might hinder the improvements of AI as it relies on data, might prove to be even more difficult in the future. In a field which promises automation, more productivity, efficiency and in turn more financial gains; limitations might receive a backlash even though ethical concerns are laid out clearly. Thus, the challenging case of ethical use of AI persists and does not seem to be inclined to resolution in the near future.

B. ACCOUNTABILITY FOR AI-GENERATED OUTCOMES

Determining accountability and responsibility for AI-generated outcomes is a complex matter involving multiple stakeholders. Developers and engineers who design and train AI models naturally hold accountability for the technology's behavior and the accuracy of its predictions. Organizations utilizing AI systems are responsible for ensuring compliance

[162] UNESCO "Recommendation on the Ethics of Artificial Intelligence", 2022, p. 32, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 10 Dec 2023.

with legal and ethical standards, managing the technology's use, and assessing its impact on stakeholders. Regulators, policymakers, and governments as well as NGOs play a vital role in shaping guidelines and laws governing the ethical use of AI and may have authority assumed due to not assessing cases of misuse by other stakeholders. In addition, individuals must understand the limitations of AI and exercise care in utilizing its outputs. The allocation of accountability for AI-generated outcomes involves a multifaceted interplay among these stakeholders, influenced by legal frameworks, ethical considerations, and specific contextual circumstances surrounding the technology's deployment and use. It is unclear how accountability should be defined, and what kind of results and consequences it must point to.

Novelli et al. address the issue of accountability through the lens of defining it, recognizing three conditions essential for possibility (acknowledgment of authority, inquiry, and constraint of power) and an arrangement comprising seven elements (setting, scope, actor, platform, norms, procedure, and consequences).[163] While a layered definition might help to acknowledge and define accountability in AI-generated outcomes, deep insights about the context of use and eliminating ambiguity in purpose of use to the maximum extent come as the prerequisites of tackling accountability. Multiple stakeholders need to work in collaboration and the ultimate accountability should be tied to a governmental body so as to create a hierarchy of accountability beginning from a junior engineer working on an AI system to a secretary of state, for example.

It would not be accurate to pin accountability on a single involving person or entity because it can be assumed that as AI gains even more popularity in our lives and its

---

[163] Claudio Novelli et al., "Accountability in artificial intelligence: what it is and how it works", *AI & Soc,* 7 Feb 2023, https://doi.org/10.1007/s00146-023-01635-y, Accessed 12 Dec 2023.

applications improve, more research on its ethical implications will be carried out. These further studies may unravel ethical infringements that were not thought possible today, and new areas of accountability may arise, requiring the involvement of different parties and stakeholders. The most crucial assumption here would be the need of constant improvements regarding acknowledging and defining accountability because it is evident that the ever-evolving AI on the tech industry's display window will continue to proliferate, opening different areas for discussion, and therefore, responsibility.

C. DEVELOPING REGULATIONS SPECIFICALLY FOR AI

On 9 December 2023, the European Council and the European Parliament reached a provisional agreement regarding the long-awaited artificial intelligence act.[164] This act aims to provide a guideline on the safety of AI and ensure that it respects human rights and the values of the EU. The provisional agreement is generally focused on prohibiting ethical use with some exceptions and puts the obligation on AI-system deployers. The details of the provisional agreement are much more that what is discussed here, but its inclusion was aimed at pointing to the fact that regulations specifically for AI are being developed. Following this provisional agreement, the world's first comprehensive set of rules, *The EU Artificial*

[164] Council of the EU, "Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the World", 9 Dec 2023, https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/, Accessed 12 Dec 2023.

Intelligence Act, aiming to regulate AI was approved on 13 March 2024.[165] Predictably, many of the discussion points regarding this breakthrough focus on the business impact of the legislation. It is discussed whether this legislation would put EU ahead of the United States and China in the AI competitive space.[166]

The Algorithmic Accountability Act in the United States requires companies to assess the AI systems they use and sell, placing accountability on deployers and developers. Aiming to regulate automated decision-making systems, ensuring fairness, accuracy, and transparency; the act addresses automated decision systems in a way that limits the possibility of their unfair use.[167]

Similarly, *The National AI Strategy* by China touches upon on setting guidelines and regulations for AI development, emphasizing national security and ethical AI usage, while its focus is the development and innovation within the field for economic and political gains.[168] Yet, as stated earlier on in this thesis, the regulations and codification on AI are mostly encouraged by motive of international competitiveness. This does not mean they do

---

[165] Brian Fung, "EU approves landmark AI law, leapfrogging US to regulate critical but worrying new technology", CNN, 13 Mar 2024, https://edition.cnn.com/2024/03/13/tech/ai-european-union/index.html#:~:text=The%20first%2Dof%2Dits%2D,deemed%20%E2%80%9Chigh%2Drisk.%E2%80%9D, Accessed 24 Mar 2024.

[166] Elizabeth Gibney, "What the EU's tough AI law means for research and ChatGPT", *Nature*, 16 Feb 2024, https://www.nature.com/articles/d41586-024-00497-8, Accessed 24 Mar 2024.

[167] GovInfo, "H.R. 5628 (IH) - Algorithmic Accountability Act of 2023", https://www.govinfo.gov/app/details/BILLS-118hr5628ih/related, Accessed 12 Dec 2023.

[168] William A. Carter, William D. Crumpler, "II. China's National Strategy for AI", *Smart Money on Chinese Advances in AI,* 1 Sep 2019, pp. 4–6, https://www.jstor.org/stable/resrep22599.7?seq=1, Accessed 12 Dec 2023.

not touch upon ethical problems, but it can be guessed that detailed, imposable ethical guidelines will be set out after the legal framework regarding other issues surrounding AI is solidified, which does not seem to be happening soon. The inclusion of legislative developments, or lack thereof, in the Unites States and China in this thesis should be interpreted as examples focusing on global competition, not as definitive statements about these counties' internal legal adequacy or inadequacies.

The statement above does not intent to claim that there have not been any significant improvements on court practices. Legal precedence has been forming with well-known cases, such as the French Parcoursup Decision,[169] and the SyRI case.[170] Parcoursup is a national algorithmic platform in France that helps schools select students and assigns them courses. The case was focused on the lawfulness of the platform, partly because students could only be aware of the existence of an algorithmic decision-making process only after they received a rejection; and the claim that the platform limited students' ability to understand and challenge the platform that made significant decisions concerning them. France's Constitutional Council ultimately ruled that Parcoursup did not infringe any public law rules, which sparked widespread discussion. This decision is contrary to the verdict of The Court of the Hague for the SyRI case. SyRI is an AI-system in Netherlands that was used to assign a risk profile to citizens, which the court ruled that breached the right to respect for private

---

[169] Robin Allen, Dee Masters, "French Parcoursup Decision", *AI Law Hub,* 16 Apr 2020, https://ai-lawhub.com/2020/04/16/french-parcoursup-decision/, Accessed 12 Dec 2023.

[170] Robin Allen, Dee Masters, "SyRI: Think Twice Before Risk Profiling", *AI Law Hub,* 30 Mar 2020, https://ai-lawhub.com/2020/03/30/syri-think-twice-before-risk-profiling/, Accessed 12 Dec 2023.

and family life set forth in Article 8 of *European Convention on Human Rights*.[171] The contradictory approaches of both courts regarding these famous cases point to an outcome that shows establishing and solidifying a comprehensive legal framework regarding AI has a long way to go, and so do ethical guidelines.

It is apparent that codification which clearly provides legal guidelines for courts to follow are needed in order to avoid confusion and irrelevant interpretation of existing legal rules by the courts. The insufficiency of clear and comprehensive legal frameworks, which is a common issue for multiple jurisdictions, create a cloud on judgements concerning AI applications. Tackling the bias problem, more specifically the gender problem of AI requires solutions that overcoming gender-biased decisions generally require in the legal field. The education of legal professionals, both through practice and theory, and developing regulations specifically for AI would reduce the number of conflicting verdicts and address the gender bias of AI with more accuracy.

[171] European Court of Human Rights, "Guide on Article 8 of the European Convention on Human Rights", 31 Aug 2022, https://www.echr.coe.int/documents/d/echr/Guide_Art_8_ENG, Accessed 12 Dec 2023.

CONCLUSION

The primary focus of this thesis is to show that deep-seated biases of societal consciousness and injustice issues infiltrate technological systems, and the complex and opaque nature of science, which artificial intelligence relies on, conceals gender bias. The intricate issue of gender bias of AI has been the subject of multiple studies, research in various fields as AI has undoubtedly become a major factor that is being used in almost every aspect of our lives. The widespread use of AI and its omnipresence raises ethical concerns, particularly those that are related to gender. Diagnosing, analyzing, and ultimately eliminating these concerns require a shift in our collective thinking, guardrails for operations and an overall willingness.

Reflecting long-standing gender biases onto the new, innovating ways shaping communities is not only done with an ulterior motive. The main cause of this reflection is obliviousness, which stem from the strength which these biases possess due to the length of their existence. Biases stemming from a distorted points of view of the societal consciousness, in relation to epistemic injustice and social constructs perpetuate the gender bias of AI, while also opening a new battleground for gender equality. Individuals who call themselves women have been facing challenges at every place they encounter others unlike their male parts, although in varying levels depending on their background and other minority characteristics. The outdated stereotypes on gender impact several aspects from social justice, economic growth, and the reliability of legal systems, which is easily overlooked when discussing AI. The basis of this thesis is that AI is much more than a new, exciting technological innovation aimed at making our lives easier. It is profoundly capable of affecting multiple layers of contemporary living, so that the impact of its ethical

infringements is amplified through the power it possesses over a myriad of operations, ranging from credit approvals to student admissions.

The realization of the fact the general public does not possess sufficient literacy about the implications AI, despite widely using it and having a basic understanding of its principles underlines that the how AI works needs to be broken down and examined within the scope of its capabilities. It is not accurate to think of AI as a human-like robot with free will, making evil choices, but a firm understanding about how it relies on data is crucial in finding the source of its gender bias and also taking the matter out of the scientific realm to make it subject to ethical discussions. This requires an examination of machine learning, which enables computers to feed on data and make decisions without needing specific programming for each decision. Discussion points around machine learning provide a profound understanding into how biases of AI manifest themselves in various situations and how the effect of these biases is exacerbated. The problems of consensus and diagnosis, insufficiency of legal protections, and the false belief that scientific fields have no place for biases all contribute to the main problem in this thesis because the matter is enclosed within a scientific frame that requires technical know-how. When combined with the already existing power imbalance and lack of equality between genders, these biases are amplified.

Disproportionate gender participation within the technology industry is a widely known and widely discussed issue that is not always linked to the biases of AI. Structural injustices and historical barriers women are subject to prevent them from entering these fields, and thus, contributing equally to the development of such technologies. This, in turn, leads to male domination in the technology industry, furthering the power imbalance and lack of equal participation between genders. As AI models depend on data, the extent and effect of this imbalance appear in connection with how large datasets are created and used. In our

inter-connected world, online sharing reached massive levels and new content is being created at every second, fueling the ever-growing big data. It is no surprise that AI models inherent societal biases both on a global and local level, because users with access to the Internet possess tools which enable them to share their thoughts and the limits set concern only illegal content, not biased opinions. In addition, data from certain processes mainly with a single objective and accumulated in-house also show biases of the organizations they are used in, such as hiring processes. Therefore, data in various sizes, scopes and actualities perpetuate the gender bias of AI in multiple ways that go easily undetected.

The emergence of AI technologies and their inherent biases have not come without their economic, sociological, and legal implications. Beyond merely being a matter of ethics, gender bias of AI has profound effects in fields directly shaping our communities and the very fabric of our lives. Gendered labor market effects with lack of equal employment opportunities for women, wage disparities and gender pay gaps, marketing and consumer behavior effects, and limitations in product development and innovation are among the consequences of the gender bias of AI. These consequences hinder the development of human societies at large, disrupting the social fabric while simultaneously creating more opportunities, time, and efficiency. The negative effects are overshadowed by the positive ones, thus requiring legal intervention.

It is not fair to say that legal developments have not managed to keep up with developments in technology, but there is much room to grow regarding creating regulations specifically for AI. Although there are some regulations, these usually fail to set clear guidelines, are easy to exploit due to the complexity of the technology and are not successful in imposing clear ethical guardrails. Regulations on AI are mainly concerned with industrial developments and international competitiveness. The countries with enough resources to

dedicate to AI prioritize its development more than its ethical and inclusive use, and this hides the problems created by gender bias behind exciting innovations. Verdicts that are contradictory in interpretations demonstrate the dire need for clear regulations developed for AI, and more specifically the ethics of it.

In conclusion, the gender bias of AI is a multifaceted, complex, and difficult-to-solve problem that spans across continents, multiple industries and requires the collaboration of both professionals from various fields and the general public. The biases of society are placed within technological systems, concealing ethical issues behind their complexity. While it may be alarming that gender bias has again manifested in a new chapter of human lives, the existence of studies, scholars and reporters who dedicate their time and resources into surfacing and eliminating this problem should provide a relief that it can be overcome.

REFERENCES

Abigail Thorn, “Social Constructs (or, 'What is A Woman, Really?')”, *YouTube*, Uploaded by Philosophy Tube, 9 Jul 2021, https://youtu.be/koud7hgGyQ8, Accessed 11 Mar 2024.

Alex Garland, *Ex Machina,* DNA Films, United Kingdom, 2014.

Alice Cuddy, “IKEA to pull 'sexist' advert in China following backlash”, *Euronews*, 25 Oct 2017, https://www.euronews.com/2017/10/25/ikea-to-pull-sexist-advert-in-china-following-backlash, Accessed 11 Dec 2023.

Alicia Wallace, “Nearly 80% of women’s jobs could be disrupted, automated by AI”, *CNN*, 22 Jun 2023, https://edition.cnn.com/2023/06/21/economy/women-employment-ai-disruption/index.html, Accessed 9 Dec 2023.

Allison Master et al., “Gender stereotypes about interests start early and cause gender disparities in computer science and engineering”, *Proceedings of the National Academy of Sciences*, vol. 118, no. 48, 22 Nov 2021, https://doi.org/10.1073/pnas.2100030118, Accessed 28 Oct 2023.

Andrew Hundt et al., “Robots Enact Malignant Stereotypes”, In *2022 ACM Conference on Fairness, Accountability, and Transparency,* 23 Jul 2022, pp. 743-756, https://doi.org/10.48550/arXiv.2207.11569, Accessed 8 Dec 2023.

Andrew Van Dam, “It’s better to be born rich than gifted”, *The Washington Post,* 9 Oct 2018, https://www.washingtonpost.com/business/2018/10/09/its-better-be-born-rich-than-talented/, Accessed 12 Dec 2023.

Ann R. Bradlow et al., “Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics”, *Speech communication*, vol. 20,3, Dec 1996, pp. 255–272, https://doi.org/10.1016/S0167-6393(96)00063-5, Accessed 10 Dec 2023.

Austin Chia, “5 Unique Ways to Use AI in Data Analytics”, *DataCamp,* Oct 2023, https://www.datacamp.com/blog/unique-ways-to-use-ai-in-data-analytics, Accessed 7 Dec 2023.

Baris Cayli Messina, “Breaking the silence on femicide: How women challenge epistemic injustice and male violence”, *The British Journal of Sociology,* vol. 73, 10 Jul 2022, pp. 859-884. https://doi.org/10.1111/1468-4446.12968, Accessed 8 Mar 2024.

Brandon Griggs, “Why computer voices are mostly female”, *CNN*, 21 Oct 2011, https://edition.cnn.com/2011/10/21/tech/innovation/female-computer-voices/, Accessed 10 Dec 2023.

Brian Fung, “EU approves landmark AI law, leapfrogging US to regulate critical but worrying new technology”, CNN, 13 Mar 2024, https://edition.cnn.com/2024/03/13/tech/ai-european-

union/index.html#:~:text=The%20first%2Dof%2Dits%2D,deemed%20%E2%80%9Chigh%2Drisk.%E2%80%9D, Accessed 24 Mar 2024.

Brian Kennedy et al., "Public Awareness of Artificial Intelligence in Everyday Activities", *Pew Research Center,* 15 Feb 2023, https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/, Accessed 4 Dec 2023.

Brianna O'Hara, "The Ultimate Guide to Generative AI for Email Marketing", *Persado,* 7 Nov 2023, https://www.persado.com/articles/generative-ai-for-email-marketing/#:~:text=One%20of%20the%20most%20significant,different%20segments%20of%20their%20audience, Accessed 8 Dec 2023.

Brittanica, "Turing test", https://www.britannica.com/technology/Turing-test, Accessed 3 Dec 2023.

Carmen Niethammer, "AI Bias Could Put Women's Lives At Risk – A Challenge For Regulators", *Forbes*, 2 Mar 2020, https://www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators/?sh=6d4165aa534f, Accessed 19 Jan 2022.

Carole Cadwalladr, Emma Graham-Harrison, "Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach", *The Guardian,* 17 Mar 2018,

https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election, Accessed 28 Mar 2024.

Carolina Aragao, "Gender pay gap in U.S. hasn't changed much in two decades", *Pew Research Center*, 1 Mar 2023, https://www.pewresearch.org/short-reads/2023/03/01/gender-pay-gap-facts/, Accessed 10 Dec 2023.

Catriona Gray, "Testimonial Injustice in Governmental AI Systems", *KI-Realitäten,* 2023, pp. 67-92, https://doi.org/10.1515/9783839466605-004, Accessed 18 Mar 2024.

Cevat Giray Aksoy et al., "Robots and the Gender Pay Gap in Europe", *IZA Discussion Papers,* no. 13482, Jul 2020, https://www.econstor.eu/bitstream/10419/223924/1/dp13482.pdf, Accessed 10 Dec 2023.

Chinasa T. Okolo, "Towards a Praxis for Intercultural Ethics in Explainable AI", *ArXiv*, 25 Apr 2023, p. 1, https://doi.org/10.48550/arXiv.2304.11861, Accessed 18 Mar 2024.

Chloe Ciora et al., "Examining Covert Gender Bias: A Case Study in Turkish and English Machine Translation Models", 2021, https://aclanthology.org/2021.inlg-1.7.pdf, Accessed 10 Dec 2023.

Christian Barry, Luara Ferracioli, "Young on Responsibility and Structural Injustice*",* *Criminal Justice Ethics*, vol. 32, no. 3, 11 Dec 2013, pp. 247–257, https://doi.org/10.1080/0731129X.2013.861981, Accessed 9 Dec 2023.

Christopher Pappas, “Is It Really A Battle Of The Brains? AI Versus Human Intelligence”, *eLearning Industry*, 28 Jul 2023, https://elearningindustry.com/is-it-really-a-battle-of-the-brains-ai-versus-human-intelligence#:~:text=The%20Differences%20Between%20AI%20and%20Human%20Intelligence&text=Humans%20are%20aware%20of%20their,the%20world%20the%20same%20way, Accessed 11 Dec 2023.

Christos Dogas, “AI, freedom of choice, Aristotle and Kant”, *Οικονομικός Ταχυδρόμος*, 4 Dec 2021, https://www.ot.gr/2021/12/04/english-edition/ai-freedom-of-choice-aristotle-and-kant/, Accessed 12 Dec 2023.

Claudio Novelli et al., “Accountability in artificial intelligence: what it is and how it works”, *AI & Soc,* 7 Feb 2023, https://doi.org/10.1007/s00146-023-01635-y, Accessed 12 Dec 2023.

Clementine Collett et al., “The Effects of AI on the Working Lives of Women”, *UNESCO, OECD, ID,* Mar 2022, http://dx.doi.org/10.18235/0004055, Accessed 10 Dec 2023.

Council of Europe, “Sex and gender”, https://www.coe.int/en/web/gender-matters/sex-and-gender, Accessed 8 Mar 2024.

Council of the EU, “Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the World”, 9 Dec 2023,

https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/, Accessed 12 Dec 2023.

Cristian Randieri, "Unveiling The Role Of AI Algorithms: Unmasking Societal Inequities And Cultural Prejudices", *Forbes*, 19 Jul 2023, https://www.forbes.com/sites/forbestechcouncil/2023/07/19/unveiling-the-role-of-ai-algorithms-unmasking-societal-inequities-and-cultural-prejudices/?sh=4f6c87882f60, Accessed 8 Dec 2023.

Curran McSwigan, "Why the Gender Pay Gap Persists", *Third Way,* 12 May 2023, https://www.thirdway.org/blog/why-the-gender-pay-gap-persists, Accessed 11 Dec 2023.

Dan Politz, "The Impact Of AI On Marketing (The Obvious & Not-So-Obvious)", *Credo,* 13 Jun 2023, https://www.getcredo.com/the-impact-of-ai-on-marketing/#:~:text=AI%20marketing%20tools%20operate%20based,customer%20experiences%20and%20brand%20reputation, Accessed 11 Dec 2023.

Daniel Leufer, "Computers are binary, people are not: how AI systems undermine LGBTQ identity", *Access Now,* 13 Jan 2024, https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/, Accessed 28 Mar 2024.

David Estlund, "What's Unjust About Structural Injustice?", *Ethics: An International Journal of Social, Political, and Legal Philosophy,* vol. 134, no. 3, 2024, pp. 333-448, https://doi.org/10.1086/728634, Accessed 28 Mar 2024.

Dawn McAra-Hunter, "How AI hype impacts the LGBTQ + community", *AI Ethics,* 14 Feb 2024, https://doi.org/10.1007/s43681-024-00423-8, Accessed 28 Mar 2024.

Digital Future Society, "Algorithmic gender discrimination and the platform economy", https://digitalfuturesociety.com/algorithmic-gender-discrimination-and-the-platform-economy/, Accessed 10 Dec 2023.

Domo, "Data Never Sleeps 5.0", 2017, https://www.domo.com/learn/infographic/data-never-sleeps-5, Accessed 6 Dec 2023.

Dylan Walsh, "The legal issues presented by generative AI", *MIT Sloan School of Management,* 28 Aug 2023, https://mitsloan.mit.edu/ideas-made-to-matter/legal-issues-presented-generative-ai, Accessed 12 Dec 2023.

Eda Eryılmaz, "The Sexism Problem of AI", *Medium*, 25 Sep 2021, https://aninjusticemag.com/the-sexism-problem-of-ai-a7594a64f7f2, Accessed 17 Jan 2022.

Elise Gould, Katherine deCourcy, "Gender wage gap widens even as low-wage workers see strong gains", *Economic Policy Institute,* 29 Mar 2023, https://www.epi.org/blog/gender-

wage-gap-widens-even-as-low-wage-workers-see-strong-gains-women-are-paid-roughly-22-less-than-men-on-average/, Accessed 11 Dec 2023.

Elizabeth Gibney, "What the EU's tough AI law means for research and ChatGPT", *Nature*, 16 Feb 2024, https://www.nature.com/articles/d41586-024-00497-8, Accessed 24 Mar 2024.

Ella Fisher, "Gender Bias in AI: Why Voice Assistants Are Female", *Adapt*, 6 Jul 2021, https://www.adaptworldwide.com/insights/2021/gender-bias-in-ai-why-voice-assistants-are-female, Accessed 10 Dec 2023.

Enterprise Garage, "Gender Stereotyping in Google Translate", 18 Mar 2021, https://www.enterprisegarage.io/2021/03/gender-stereotyping-in-google-translate/, Accessed 9 Dec 2023.

Estrella Gomez-Herrera, Sabine T. Koeszegi, "A Gender Perspective On Artificial Intelligence And Jobs: The Vicious Cycle Of Digital Inequality", *Bruegel WORKING PAPER*, 30 Aug 2022, https://www.bruegel.org/sites/default/files/2022-08/WP%202022%2015%20AI%20290822.pdf, Accessed 10 Dec 2023.

Euan Black, "AI more likely to hire women than humans are: study", *Financial Review,* 17 May 2023, https://www.afr.com/work-and-careers/workplace/ai-more-likely-to-hire-women-than-humans-are-study-20230511-p5d7ok, Accessed 11 Dec 2023.

Eugene Garver, “Point of View, Bias, and Insight” *Metaphilosophy*, vol. 24, no. 1/2, 1993, pp. 47–60. http://www.jstor.org/stable/24438954, Accessed 9 Dec 2023.

European Court of Human Rights, “Guide on Article 8 of the European Convention on Human Rights”, 31 Aug 2022, https://www.echr.coe.int/documents/d/echr/Guide_Art_8_ENG, Accessed 12 Dec 2023.

European Parliament, “EU AI Act: first regulation on artificial intelligence”, 19 Dec 2023, https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence, Accessed 24 Mar 2024.

Fabio Duarte, “Amount of Data Created Daily (2023)”, *Exploding Topics,* 13 Dec 2023, https://explodingtopics.com/blog/data-generated-per-day, Accessed 6 Dec 2023.

Frank Swain, “Is it right to use Nazi research if it can save lives?”, *BBC*, 24 Jul 2019, https://www.bbc.com/future/article/20190723-the-ethics-of-using-nazi-science, Accessed 4 Dec 2023.

Gai Sher, Ariela Benchlouch, “The privacy paradox with AI”, *Reuters*, 31 Oct 2023, https://www.reuters.com/legal/legalindustry/privacy-paradox-with-ai-2023-10-31/#:~:text=AI's%20privacy%20dilemma%20rests%20on,storage%2C%20usage%2C%20and%20access, Accessed 12 Dec 2023.

Genevieve Smith, Ishita Rustagi, "When Good Algorithms Go Sexist: Why and How to Advance AI Gender Equality", *Stanford Social Innovation Review*, 31 Mar 2021, https://ssir.org/articles/entry/when_good_algorithms_go_sexist_why_and_how_to_advance_ai_gender_equity, Accessed 19 Jan 2022.

Gil Appel et al, "Generative AI Has an Intellectual Property Problem", *Harvard Business Review*, 7 Apr 2023, https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem, Accessed 12 Dec 2023.

GovInfo, "H.R. 5628 (IH) - Algorithmic Accountability Act of 2023", https://www.govinfo.gov/app/details/BILLS-118hr5628ih/related, Accessed 12 Dec 2023.

Grand View Research, "Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning), By End-use, By Region, And Segment Forecasts, 2023 – 2030", https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market, Accessed 5 Dec 2023.

Gülriz Uygur, "Hukukta Adaletsizliği Görmek", *Türkiye Felsefe Kurumu*, Ankara, 2020.

Gülriz Uygur, "Toplumsal Cinsiyet ve Adalet: Hukuk Adaletsizdir", *Ankara Barosu Dergisi*, no. 4, 2015, pp. 121–132, https://dergipark.org.tr/tr/download/article-file/398472, Accessed 13 Mar 2023.

IBM, “How Artificial Intelligence is depicted in movies”, https://www.ibm.com/support/pages/how-artificial-intelligence-depicted-movies, Accessed 9 Dec 2023.

IBM, “What is artificial intelligence?”, https://www.ibm.com/topics/artificial-intelligence, Accessed 30 Jan 2022.

IBM, “What is machine learning?”, https://www.ibm.com/topics/machine-learning, Accessed 1 Dec 2023.

İbrahim Yıldız et al., “A Case Study On Glass Ceiling Syndrome of Female Employees in the Information Technology Sector”, *Atatürk İletişim Dergisi,* vol. 16, 2018, pp. 99-112, https://doi.org/10.32952/atauniiletisim.489417, Accessed 12 Dec 2023.

Imane El Atillah, “Nearly 80% of women’s jobs at risk from generative AI, new research finds”, *Euronews*, 29 Jun 2023, https://www.euronews.com/next/2023/06/29/womens-jobs-will-be-disproportionately-affected-by-ai-according-to-new-research, Accessed 9 Dec 2023.

International Labour Organization, “Gender Wage Gap is 15.6% in Turkey, according to the joint study by ILO Turkey Office and TURKSTAT”, 5 Oct 2020, https://www.ilo.org/ankara/news/WCMS_757055/lang--en/index.htm, Accessed 9 Dec 2023.

International Labour Organization, “The gender gap in employment: What's holding women back?”, Feb 2022, https://www.ilo.org/infostories/en-GB/Stories/Employment/barriers-women#global-gap, Accessed 10 Dec 2023.

International Women’s Day, “Gender and AI: Addressing bias in artificial intelligence”, https://www.internationalwomensday.com/Missions/14458/Gender-and-AI-Addressing-bias-in-artificial-intelligence, Accessed 1 Dec 2023.

Iris Marion Young, “Political Responsibility and Structural Injustice”, *The Lindley Lecture, University of Kansas,* 5 May 2003.

Iris Marion Young, “Justice and the Politics of Difference”, *Princeton University Press,* New Jersey, 1990.

Isobel Asher Hamilton, “Why it's totally unsurprising that Amazon's recruitment AI was biased against women”, *Business Insider,* 13 Oct 2018, https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10, Accessed 10 Dec 2023.

IWDA, “What Does Intersectional Feminism Actually Mean?”, 11 May 2018, https://iwda.org.au/what-does-intersectional-feminism-actually-mean/, Accessed 30 Nov 2023.

Jason Osler, "Consumers are pushing back against gendered products", *Canadian Broadcasting Corporation*, 25 Oct 2018, https://www.cbc.ca/news/canada/consumers-are-pushing-back-against-gendered-products-1.4855023, Accessed 11 Dec 2023.

Jeffrey Dastin, "Insight - Amazon scraps secret AI recruiting tool that showed bias against women", *Reuters*, 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-airecruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/, Accessed 24 Jan 2022.

Jeffrey Reiman, "The Structure of Structural Injustice: Thoughts on Iris Marion Young's 'Responsibility for Justice.'", *Social Theory and Practice*, vol. 38, no. 4, Oct 2012, pp. 738–751, https://www.jstor.org/stable/23558769. Accessed 9 Dec 2023.

Jess Romeo, "How Computer Science Became a Boys' Club", *JSTOR Daily,* 29 Aug 2021, https://daily.jstor.org/how-computer-science-became-a-boys-club/, Accessed 3 Dec 2023.

Jill Rosen, "Flawed AI Makes Robots Racist, Sexist", *Georgia Tech,* 23 Jun 2022, https://research.gatech.edu/flawed-ai-makes-robots-racist-sexist, Accessed 9 Dec 2023.

Joe Neel, "Poll: Discrimination Against Women Is Common Across Races, Ethnicities, Identities", *National Public Radio*, 11 Dec 2017, https://www.npr.org/2017/12/11/569251267/poll-discrimination-against-women-is-common-across-races-ethnicities-identities, Accessed 2 Dec 2023.

John Harris, "'There was all sorts of toxic behaviour': Timnit Gebru on her sacking by Google, AI's dangers and big tech's biases", *The Guardian,* 22 May 2023, https://www.theguardian.com/lifeandstyle/2023/may/22/there-was-all-sorts-of-toxic-behaviour-timnit-gebru-on-her-sacking-by-google-ais-dangers-and-big-techs-biases, Accessed 26 Mar 2024.

Jonathan Dupont et al., "What does the public think about AI?", *PublicFirst,* 2023, https://publicfirst.co.uk/ai/, Accessed 8 Dec 2023

Jose Ramon Saura et al., "From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets", *International Journal of Information Management,* vol. 60, Oct 2021, https://doi.org/10.1016/j.ijinfomgt.2021.102331, Accessed 7 Dec 2023.

Jules Gleeson, "Judith Butler: 'We need to rethink the category of woman'", *The Guardian,* 7 Sep 2021, https://www.theguardian.com/lifeandstyle/2021/sep/07/judith-butler-interview-gender, Accessed 8 Mar 2024.

Julianne McHugh, Karen Miki Behr, "Artificial Intelligence: A Buoy Or Barrier For Women In The Workplace?", *Newton Investment Management*, 18 Sep 2023, https://www.newtonim.com/uk-institutional/insights/blog/artificial-intelligence-a-buoy-or-barrier-for-women-in-the-workplace-2/, Accessed 10 Dec 2023.

Karen Frost-Arnold, "Epistemic Injustice and AI Ethics: Learning from Google's Treatment of Timnit Gebru", *Feminism, Social Justice, and AI*, https://feminism-social-justice-ai.org/epistemic-injustice-and-ai-ethics-learning-from-googles-treatment-of-timnit-gebru/, Accessed 20 Mar 2024.

Karen Gilchrist, "World's first major act to regulate AI passed by European lawmakers", *CNBC*, 13 Mar 2024, https://www.cnbc.com/2024/03/13/european-lawmakers-endorse-worlds-first-major-act-to-regulate-ai.html, Accessed 24 Mar 2024.

Karina Mochetti, "The Impact of Women in Computer Science History: A Post-War American History", *Transversal International Journal for the Historiography of Science,* no. 6, 30 Jun 2019, https://doi.org/10.24117/2526-2270.2019.i6.07, Accessed 28 Oct 2023.

Kashmir Hill, "Wrongfully Accused by an Algorithm", *The New York Times,* 24 Jun 2020, https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html, Accessed 18 Jan 2022.

Kim Parker, "What's behind the growing gap between men and women in college completion?", *Pew Research Center*, 8 Nov 2021, https://www.pewresearch.org/short-reads/2021/11/08/whats-behind-the-growing-gap-between-men-and-women-in-college-completion/, Accessed 5 Dec 2023.

Kişisel Verileri Koruma Kurumu, "Kurumsal Tarihçe", https://www.kvkk.gov.tr/Icerik/2075/Kurumsal-Tarihce, Accessed 14 Jan 2022.

Klaudyna Horniczak et al., “Is Epistemic Status Gender-Biased? Gender As a Predictor of Testimonial Reliability Assessments in Violent Crimes”, *International Journal for the Semiotics of Law,* 17 Jan 2024, https://doi.org/10.1007/s11196-023-10055-6, Accessed 21 Feb 2024.

Krishna Nathwani, “Why Girls, Why Literacy, Why Now?”, *World Literacy Foundation,* 1 Mar 2022, https://worldliteracyfoundation.org/why-girls-why-literacy-why-now/, Accessed 11 Dec 2023.

LANGLOIS, “Ethics in artificial intelligence: Discriminatory biases”, 12 Sep 2023, https://langlois.ca/ethics-in-artificial-intelligence-discriminatory-biases/#:~:text=Discriminatory%20AI%20has%20many%20underlying%20causes&text=The%20biases%20found%20in%20AI,not%20neutral%2C%20fair%20or%20equitable, Accessed 1 Dec 2023.

Lauren Balser, " How Computing Became Masculine in the Cold War”, *Confluence*, 24 Sep 2021, https://confluence.gallatin.nyu.edu/sections/research/how-computing-became-masculine-in-the-cold-war#easy-footnote-bottom-22-20468, Accessed 9 Dec 2023.

LGBT Foundation, “What Is A Woman?”, 8 Aug 2023, https://lgbt.foundation/whatisawoman, Accessed 30 Nov 2023.

Liz Simmons, “Most Influential Women in Computer Science”, *Computer Science.org*, 26 Jul 2023, https://www.computerscience.org/resources/most-influential-women-computer-science/, Accessed 8 Dec 2023.

Local Government Association, “General Data Protection Regulation (GDPR)”, https://www.local.gov.uk/our-support/research-and-data/data-and-transparency/general-data-protection-regulation-gdpr#:~:text=The%20GDPR%20is%20Europe's%20new,started%20on%2025%20May%202018, Accessed 01 Dec 2023.

Maeve McKeown, “Structural injustice”, *Philosophy Compass,* vol. 16, 18 Jun 2021, https://doi.org/10.1111/phc3.12757, Accessed 19 Oct 2023.

Margaret Hu, “Cambridge Analytica’s black box”, *Big Data & Society,* 24 Aug 2020, https://doi.org/10.1177/2053951720938091, Accessed 7 Dec 2023.

Martina Sardelli, “Epistemic Injustice in the Age of AI”, *Aporia*, vol. 22, 17 Aug 2022, pp. 44-53, https://ojs.st-andrews.ac.uk/index.php/aporia/article/view/2455/1871, Accessed 24 Mar 2024.

Mary Lou O’Neil et al., “The State of Gender (In)Equality in Academia in Turkey 1984 – 2018”, *KHAS Institutional Research,* Apr 2019, https://www.khas.edu.tr/wp-content/uploads/2022/05/The-State-of-Gender-Inequality-in-Academia-in-Turkey_0.pdf, Accessed 9 Dec 2023.

Merriam-Webster, "social construct", https://www.merriam-webster.com/dictionary/social%20construct, Accessed 8 Mar 2024.

MGI Research, "The Global Tech Market Is Bigger Than You Think", 28 Feb 2023, https://mgiresearch.com/research/the-global-tech-market-is-bigger-than-you-think/, Accessed 5 Dec 2023.

Miranda Fricker, *Epistemik Adaletsizlik - İktidar ve Bilmenin Etiği,* Translated by Kadir Gülen, Fol Kitap, Ankara, 2023.

Morten Fibieger Byskov, "What Makes Epistemic Injustice an "Injustice"?", *Journal of Social Philosophy,* vol. 52, 18 May 2020, pp. 114-131, https://doi.org/10.1111/josp.12348, Accessed 3 Mar 2024.

Nadire Özdemir, "Iris Marion Young Teorisine Genel Bir Bakış: Yapısal Adaletsizlik ve Adaletsizlikten Sorumluluğa İlişkin Sosyal Bağlantı Modeli", *Türkiye Barolar Birliği Dergisi*, vol. 32, no. 146, Jan 2020, pp. 247-267, http://tbbdergisi.barobirlik.org.tr/m2020-146-1899, Accessed 9 Dec 2023.

Naomi Cahn, "Do Women And Men Have A Confidence Gap?", *Forbes*, 26 Feb 2020, https://www.forbes.com/sites/naomicahn/2020/02/26/do-women-and-men-have-a-confidence-gap/?sh=3d7bdc527bd2, Accessed 26 Oct 2023.

Natalie Huet, "Gender bias in recruitment: How AI hiring tools are hindering women's careers", *Euronews*, 8 Mar 2022, https://www.euronews.com/next/2022/03/08/gender-bias-in-recruitment-how-ai-hiring-tools-are-hindering-women-s-careers, Accessed 10 Dec 2023.

Nima Sanandaji, "The Nordic Glass Ceiling", *CATO Institute,* 8 Mar 2018, https://www.cato.org/policy-analysis/nordic-glass-ceiling, Accessed 1 Dec 2023.

Nitin Kumar Singh et al., "Artificial intelligence and machine learning-based monitoring and design of biological wastewater treatment systems", *Bioresource Technology,* vol. 369, Feb 2023, https://doi.org/10.1016/j.biortech.2022.128486, Accessed 12 Dec 2023.

NVIDIA, "What is Generative AI?", https://www.nvidia.com/en-us/glossary/data-science/generative-ai/, Accessed 08 Dec 2023.

Office for National Statistics, "Gender pay gap in the UK: 2023", 1 Nov 2023, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/genderpaygapintheuk/2023#:~:text=In%202023%2C%20the%20gap%20among,seen%20in%202019%20(17.4%25), Accessed 10 Dec 2023.

OHCHR, "Gender stereotyping", https://www.ohchr.org/en/women/gender-stereotyping#:~:text=A%20gender%20stereotype%20is%20a,personal%20abilities%2C%20pursue%20their%20professional, Accessed 9 Dec 2023.

Ömer Ersin Kahraman, "Surveillance Capitalism as a Form of Biopower in Historical Perspective", *İzmir Journal of Economics,* vol. 37, no. 3, 2022, pp. 778-792, https://doi.org/10.24988/ije.1047476, Accessed 7 Dec 2023.

Pauline Kim, "AI and Inequality", *Washington University in St. Louis Legal Studies Research Paper,* no. 21-09-03, 11 Oct 2021, https://ssrn.com/abstract=3938578, Accessed 11 Dec 2023.

Peter Dockrill, "Robots With Flawed AI Make Sexist And Racist Decisions, Experiment Shows", *ScienceAlert,* 27 Jun 2022, https://www.sciencealert.com/robots-with-flawed-ai-make-sexist-racist-and-toxic-decisions-experiment-shows, Accessed 9 Dec 2023.

Prathamesh Patalay, "COMPAS : Unfair Algorithm?", *Medium*, 22 Nov 2023, https://medium.com/@lamdaa/compas-unfair-algorithm-812702ed6a6a, Accessed 24 Mar 2024.

Qian Li et al., "AI-enabled price discrimination as an abuse of dominance: a law and economics analysis", *China-EU Law Journal,* vol. 9, 27 Apr 2023, pp. 51-72, https://doi.org/10.1007/s12689-023-00099-z, Accessed 11 Dec 2023.

Rachel Goodman, "Why Amazon's Automated Hiring Tool Discriminated Against Women ", *American Civil Liberties Union*, 12 Oct 2018, https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against, Accessed 24 Jan 2022.

Rhaina Cohen, “What Programming's Past Reveals About Today's Gender-Pay Gap”, *The Atlantic,* 7 Sep 2016, https://www.theatlantic.com/business/archive/2016/09/what-programmings-past-reveals-about-todays-gender-pay-gap/498797/, Accessed 10 Dec 2023.

Rick Merritt, “What Is Retrieval-Augmented Generation, aka RAG?”, *NVIDIA Blog,* 15 Nov 2023, https://blogs.nvidia.com/blog/what-is-retrieval-augmented-generation/, Accessed 24 Mar 2024.

Robin Allen, Dee Masters, “French Parcoursup Decision”, *AI Law Hub,* 16 Apr 2020, https://ai-lawhub.com/2020/04/16/french-parcoursup-decision/, Accessed 12 Dec 2023.

Robin Allen, Dee Masters, “SyRI: Think Twice Before Risk Profiling”, *AI Law Hub,* 30 Mar 2020, https://ai-lawhub.com/2020/03/30/syri-think-twice-before-risk-profiling/, Accessed 12 Dec 2023.

Safe Computing - University of Michigan, “History of Privacy Timeline”, https://safecomputing.umich.edu/privacy/history-of-privacy-timeline, Accessed 8 Dec 2023.

Sapna Cheryan et al., “There Are Too Few Women in Computer Science and Engineering”, *Scientific American,* 27 Jul 2022, https://www.scientificamerican.com/article/there-are-too-few-women-in-computer-science-and-engineering/, Accessed 28 Oct 2023.

Simone De Beauvoir, *Kadın “İkinci Cins”,* Translated by Bertan Onaran, Payel Yayınları, İstanbul, 1993.

Sourajit Saha, "Should Artificial Intelligence Be Capitalized?", *Medium*, 16 Jul 2023, https://medium.com/@sourajitsaha17/should-artificial-intelligence-be-capitalized-98e63485ba0f#:~:text=There%20is%20no%20universally%20agreed,the%20general%20concept%20or%20technology, Accessed 03 Dec 2023.

Statista, "Social Media & User-Generated Content", https://www.statista.com/markets/424/topic/540/social-media-user-generated-content/#overview, Accessed 6 Dec 2023.

Stephen John, "Why science isn't objective", *iai news,* 26 Jul 2021, https://iai.tv/articles/why-science-isnt-objective-auid-1846, Accessed 19 Jan 2022.

Şule Güner, "İlk yerli yapay zekâ anketi", *Para Dergisi,* 29 Apr 2018, https://itucekirdek.com/haberler/ilk-yerli-yapay-zeka-anketi/, Accessed 20 Jan 2022.

Sven Nyholm, "Artificial Intelligence, Ethics of", *Encyclopedia of the Philosophy of Law and Social Philosophy,* 12 Nov 2022, pp. 1-8, https://doi.org/10.1007/978-94-007-6730-0_1093-1, Accessed 7 Dec 2023.

Thomas Wahl, "Commission: White Paper on AI", 1 Apr 2020, https://eucrim.eu/news/commission-white-paper-ai/, Accessed 3 Dec 2023.

Tim Radford, “What has medicine learned from the Nazis?”, *The Guardian,* 14 Oct 2004, *https://www.theguardian.com/news/2004/oct/14/thisweekssciencequestions.cancer,* Accessed 4 Dec 2023.

Timnit Gebru, “Race and Gender: Data-Driven Claims Race and Gender Perpetuate the Negative Biases of the Day”, *The Oxford Handbook of Ethics of AI* (Compiled by: Markus D. Dubber, Frank Pasquale, Sunit Das), Oxford University Press, New York, 2020, p. 258.

TOPPAN, “Grammatical Genders in Different Languages”, https://toppandigital.com/us/blog-usa/grammatical-genders-in-different-languages/#:~:text=Languages%20such%20as%20Finnish%2C%20Hungarian,both%20sexes)%20and%20inanimate%20objects, Accessed 9 Dec 2023.

TRT Haber, “Ulusal Yapay Zekâ Stratejisi Yönlendirme Kurulu Toplanacak”, 13 Jan 2023, https://www.trthaber.com/haber/gundem/ulusal-yapay-zeka-stratejisi-yonlendirme-kurulu-toplanacak-644569.html, Accessed 13 Jan 2022.

Umut Özdemir, “Facebook-Cambridge Analytica Skandalının Katılımcı Kültür, Dijital Emek Sömürüsü ve Mahremiyetin İhlali Çerçevesinde İncelenmesi”, *Egemia Ege Üniversitesi İletişim Fakültesi Medya Ve İletişim Araştırmaları Hakemli E-Dergisi,* no. 10, 2022, pp. 22-34, https://doi.org/10.56075/egemiadergisi.999935, Accessed 8 Dec 2023.

UNDP, “Women and Girls in science and technology: Bridging the gender gap”, 10 Feb 2022, https://www.undp.org/ghana/blog/women-and-girls-science-and-technology-bridging-

gender-gap?gad_source=1&gclid=CjwKCAiAjrarBhAWEiwA2qWdCMTCi5Cf66EDleK_qbRiRivhynG-2QNIa1fF6tgtmM7Qn3yafv70XRoCquQQAvD_BwE, Accessed 1 Dec 2023.

UNESCO, "AI and the Rule of Law: Capacity Building for Judicial Systems", https://www.unesco.org/en/artificial-intelligence/rule-law/mooc-judges, Accessed 3 Dec 2023.

UNESCO, "Artificial Intelligence: examples of ethical dilemmas", 21 Apr 2023, https://www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases, Accessed 30 Nov 2023.

UNESCO, "Does Artificial Intelligence advance gender equality?", 8 Mar 2023, https://www.unesco.org/en/articles/does-artificial-intelligence-advance-gender-equality, Accessed 1 Dec 2023.

UNESCO, "Ethics of Artificial Intelligence" https://www.unesco.org/en/artificial-intelligence/recommendation-ethics, Accessed 12 Dec 2023.

UNESCO, "Recommendation on the Ethics of Artificial Intelligence", 2022, https://unesdoc.unesco.org/ark:/48223/pf0000381137, Accessed 12 Dec 2023.

Upwork, "How AI Is Used in Decision-Making Processes", 16 Aug 2023, https://www.upwork.com/resources/ai-in-decision-making, Accessed 6 Dec 2023.

Valentina Bellini et al., “Understanding basic principles of Artificial Intelligence: a practical guide for intensivists.”, *Acta Biomed,* vol. 93,5, 26 Oct 2022, https://doi.org/10.23750/abm.v93i5.13626, Accessed 5 Dec 2023.

Volker Roelcke, “Nazi medicine and research on human beings”, *The Lancet,* vol. 364. Dec 2004, pp. 6-8, https://doi.org/10.1016/S0140-6736(04)17619-8, Accessed 4 Dec 2023.

William A. Carter, William D. Crumpler, “II. China’s National Strategy for AI”, *Smart Money on Chinese Advances in AI,* 1 Sep 2019, pp. 4–6, https://www.jstor.org/stable/resrep22599.7?seq=1, Accessed 12 Dec 2023.

Winnie Ma, Vincent Valton, “Toward an Ethics of AI Belief” *ArXiv*, 18 Sep 2023, p. 5, https://doi.org/10.48550/arXiv.2304.14577, Accessed 18 Mar 2024.

Women in Tech Network, “Women in Tech and STEM Workforce”, https://www.womentech.net/en-us/women-technology-statistics, Accessed 4 Dec 2023.

World Economic Forum, “Assessing Gender Gaps in Artificial Intelligence”, https://reports.weforum.org/global-gender-gap-report-2018/assessing-gender-gaps-in-artificial-intelligence/, Accessed 13 Jan 2022.

World Health Organization, “Female genital mutilation”, https://www.who.int/news-room/fact-sheets/detail/female-genital-mutilation, Accessed 2 Dec 2023.

WorldData, “Average height and weight by country”, https://www.worlddata.info/average-bodyheight.php, Accessed 10 Mar 2024.

Zeynep Tufekci, “Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls”, *ICWSM '14: Proceedings of the 8th International AAAI Conference on Weblogs and Social Media,* 2014, https://www.academia.edu/6613059/Big_Questions_for_Social_Media_Big_Data_Representativeness_Validity_and_Other_Methodological_Pitfalls, Accessed 6 Dec 2023.

Zhisheng Chen, “Ethics and discrimination in artificial intelligence-enabled recruitment practices”, *Humanities and Social Sciences Communications*, vol. 10, 13 Sep 2023, https://doi.org/10.1057/s41599-023-02079-x, Accessed 11 Dec 2023.

ABSTRACT

This study deals with the gender bias of artificial intelligence. The recent developments in artificial intelligence and its increasing use create ethical problems, particularly those that concern gender. An understanding about gender biases, how artificial intelligence works and how reliant it is on data is important to understand its impact on society at large. Therefore, the first section of this study focuses on addressing the gender problem while discussing the emergence of biases in technology with various manifestations. The second section focuses on laying out the economic, social implications of artificial intelligence and discussing the questions existing within the legal landscape. Thus, it has been aimed to explore and analyze ways in which the gender bias comes into existence, how it penetrates artificial intelligence, and how the intricate problems the gender bias of artificial intelligence causes require complex solutions with the collaboration of multiple stakeholders.

Keywords: Ethics of Artificial Intelligence, Feminist Artificial Intelligence, Gender Inequality, Epistemic Injustice, Structural Injustice

ÖZET

Çalışma, yapay zekânın cinsiyet ön yargısını ele almaktadır. Yapay zekâdaki son gelişmeler ve bu teknolojinin artan kullanımı, özellikle toplumsal cinsiyeti ilgilendiren etik sorunlar yaratmaktadır. Toplumun genelini etkileyen bu konuyu ele almak için cinsiyet ön yargıları, yapay zekânın nasıl çalıştığı ve veriye ne kadar bağımlı olduğu konularında bir anlayışa sahip olmak önemlidir. Bu nedenle, bu çalışmanın ilk bölümü, teknolojideki ön yargıların ortaya çıkışını ve farklı tezahürlerini tartışırken toplumsal cinsiyet sorununu ele almaya odaklanmaktadır. İkinci bölüm, yapay zekânın ekonomik ve sosyolojik açılardan sonuçlarını ortaya koymakta ve mevcut hukuki düzenlemeler içindeki soruları tartışmaktadır. Böylece, yapay zekâdaki cinsiyet ön yargısının nasıl oluştuğunu, yapay zekâya nasıl nüfuz ettiğini ve yapay zekânın cinsiyet ön yargısının ortaya çıkardığı karmaşık sorunların farklı paydaşların iş birliğiyle karmaşık çözümler gerektirdiğini keşfetmek ve analiz etmek amaçlanmıştır.

Çalışmanın ilk bölümü kuramsal çerçevede genel olarak yapay zekâdaki cinsiyet ön yargısının ortaya koyulmasına odaklanmış, cinsiyet ön yargısının geniş şekilde ele alınması ve teknolojide ön yargıların ortaya çıkışı üzerinde şekillenmiştir. Öncelikle, epistemik adaletsizlikten yola çıkılarak ön yargı ve toplumsal cinsiyet kavramları üzerinden mevcut sorun anlamlandırmaya çalışılmıştır. Kişilerin bilen özne olma kapasitesine yönelik adaletsizlik, ön yargıları görünmez kılmakta ve ön yargılara tabi tutulan grupların bunu bilmesini engelleyerek bir iktidar meselesi olmaktadır. Toplumsal cinsiyet eşitsizliğine dair ön yargıların küresel ölçekte varlığı, yapay zekâya da tezahür etmekte ve yetersiz temsil edilen farklı gruplar arasında kadınlar üzerindeki ön yargıların yaygınlığı ile ön plana çıkmaktadır. Fakat, ön yargı ve toplumsal cinsiyet ekseninde bir değerlendirme yapabilmek için yapay zekânın genel çalışma prensiplerinin de anlaşılması gereklidir. Yapay zekâ

sistemleri temel olarak veri toplayan, verileri işleyen ve algoritmalar kullanarak belirli karar verme mekanizmalarını öğrenen sistemlerdir. Bu tanım doğrultusunda ön yargılarının varlığının verilerle, verilerin işlenme amaç ve biçimleriyle ve nihai olarak da algoritmaların belirlenmesinde rol oynayan bilim insanları ve sektörel paydaşlarla ilişkili olduğu açıktır.

Yapay zekânın bir alt dalı olan makine öğrenmesi ise sistemlerin belirli görevleri yerine getirmek için analiz yapmasına, deneyim yoluyla öğrenmesine dayalı bir sistemdir; dolayısıyla makine öğrenmesi etrafında şekillenen çeşitli tartışma noktalarının da yapay zekânın cinsiyet ön yargısına etkisinin değerlendirilmesi yapılmıştır. Bu bağlamda konuya dair oluşmuş mutabakat ve teşhis sorunu, sorunların çözümü bakımından hukuki korumaların yetersizliği ve bilimsel alanlarda ön yargı veya ayrımcılık olmayacağı yönündeki yaygın toplumsal inanış ele alınmış; yapay zekânın cinsiyet ayrımcılığı sorununun derinliğinin ortaya koyulması amaçlanmıştır.

İlk bölümün devamında yapay zekânın teknoloji sektörünün ayrılmaz bir parçası olduğundan yola çıkılarak teknoloji sektörünün genel sorunlarının ön plana çıkarılmasına odaklanılmıştır. Elbette cinsiyete dair ön yargılar kadınları dezavantajlı konuma düşüren ve ayrımcılığı sürdüren endüstriyel koşullardan ve yapay zekâyı besleyen verilerden ayrıştırılamayacaktır. Bu itibarla teknoloji sektöründeki orantısız cinsiyet dağılımı yapısal adaletsizlikler ve bilgisayar bilimindeki erkek baskınlığı ekseninde ele alınmıştır. Kadınların bu alanlarda niceliksel bağlamda erkekler kadar etkinlik gösterememesi sadece ilgisizlik veya basmakalıp inanışlarla açıklanamamaktadır. Katmanlı bir şekilde çeşitli yapısal adaletsizlikler, ekonomik ve eğitimsel dezavantajlar, engelleyici toplumsal baskılar ve yetersiz desteklerle karşılaşan kadınlar; dolaylı olarak yapay zekâ teknolojilerinin geliştirilmesine katılamamakta, bu teknolojilerin eşitlikçi ve kapsayıcı olması doğrultusundaki süreçlere müdahil olamamaktadır.

Yapay zekâ ile veri, belirli amaçlar doğrultusunda kullanılmaları halinde birbirinden ayrılamayacak iki kavramdır. Veri ile beslenen yapay zekâ, büyüyen veri ve kavramlar havuzuna yeni içerikler katmakta; hayatı kolaylaştırırken beraberinde karmaşık sorunlar getirmektedir. Fakat toplumda yapay zekâya dair yeterli düzeyde farkındalık veya okuryazarlığın olmaması yaygın inanışları da beraberinde getirmekte, yapay zekânın tam olarak ne olduğunun toplum tarafından anlaşılamaması ortaya çıkan sorunlara dikkat çekilmesini zorlaştırmaktadır. Günümüzde internete erişimi olan herkesin giderek büyüttüğü veri havuzu, beraberinde getirdiği çok sayıda farklı tartışmanın gölgesinde kalan cinsiyet ön yargısı sorununu da kötüleştirmekte, küresel veri paylaşımı ve kullanıcılar tarafından oluşturulan içerikler üreticilerinin ön yargılı düşüncelerini bu havuz aracılığıyla yapay zekânın oyun alanına sokmaktadır. Bunun yanı sıra işe alım gibi belirli süreçlere dair biriktirilen ve yapay zekâya sunulan veri kümeleri ise ön yargıların daha belirli durumlarda etkisini göstermesine yol açmakta, kadınların toplum nezdinde oluşan ve yaygınlaşan basmakalıp düşüncelerin de ötesinde aktif olarak ayrımcılığa uğramasına neden olmaktadır.

Çalışmanın ikinci bölümü kavramsal çerçevede yapay zekânın cinsiyet ön yargısının ekonomik ve sosyolojik sonuçlarının ortaya koyulmasına odaklanmış ve bu eksende şekillenen hukuki tartışmalara değinmiştir. Cinsiyet dağılımı orantılı olmayan iş pazarları, kadınlar için adil olmayan istihdam fırsatları ve gelir eşitsizliği kapsamında açıklanmıştır. Bununla birlikte yapay zekânın pazarlama ve tüketici davranışı üzerindeki geniş etkisi, ürün geliştirme ve yenilikçilik süreçlerinde yarattığı kısıtlamalar da ekonomik etkisini daha büyük ölçekte göstermektedir. Dolayısıyla zaman kazanımı gibi pek çok avantajıyla ön plana çıkan yapay zekânın toplumdaki kadınlar ve genel olarak da şirketlerin kârlılık amacıyla yönlendirdiği kitleler üzerinde ön yargılarla ilişkili olarak sebep olduğu olumsuzluklar ekonomik sonuçları arasındadır.

Yapay zekâdaki cinsiyet ön yargısının sosyolojik açıdan etkileri önce sosyal normlar ve tavırların, sonra da eşitsizlik ve dışlanmışlığı daha kötü bir duruma sokmasının tartışılması ile ortaya konulmuştur. Belirtmek gerekir ki yapay zekâ bir insan aklı gibi çalışarak, insanlar arasındaki davranışsal nüansları veya kültürel durumları bir insan kadar derinlikle anlayarak hayatımıza yepyeni basmakalıp düşünceler getirmemekte, ancak var olan olumsuz algıların, cinsiyet ön yargılarının ve toplumsal cinsiyete dair rollerin güçlenmesine yardımcı olmaktadır. Karar verme süreçlerinde pek çok örnekte doğrudan veya dolaylı olarak ön yargıları artıran yapay zekâ, cinsiyetlerin yapay zekâ teknolojilerine erişimlerinin de eşit olmaması nedeniyle etik bağlamda isabetli düzeyde değerlendirilememektedir.

Hukuki tartışmalar bakımından yapay zekânın gündeme henüz girmediğini iddia etmek tutarlı olmayacaktır, nitekim yapay zekânın etik şekilde kullanımı, yapay zekâ tarafından üretilen neticelerin sorumluluğu ve doğrudan yapay zekâyı ilgilendiren regülasyonların geliştirilmesi hukukun gündemindedir. Ancak bunların cinsiyet ön yargısını tespit etmek ve çözüme kavuşturmak konusundaki etkisi tartışmaya açıktır. Bu etkinin yetersizliği, durumun karmaşıklığı ve derinliği ile bağlantılıdır; ne var ki ortak bir kurallar bütününün eksikliği gerek çelişkili yargı kararları ve hukuki yorumlara, gerekse sorumluluğun atfı konusunda mutabakata varılamamasına yol açmaktadır. Ek olarak yapay zekâ için çıkarılan hukuki kurallar da ağırlıklı olarak uluslararası rekabete odaklandığından etik tartışmalar arka planda kalmaktadır.

Sonuç bölümünde yapay zekânın cinsiyet ön yargısının kıtaları, birçok endüstriyi etkileyen ve çözümü için pek çok farklı alandan profesyoneller ile halkın iş birliğini gerektiren karmaşık bir problem olduğu vurgulanmıştır. Cinsiyet ön yargılarının insan hayatını kolaylaştıran bu yeni ve heyecan verici teknolojide tezahür etmesi her ne kadar endişe verici olsa da bu sorunu ortaya çıkarmak için zaman ve kaynaklarını harcayan

paydaşların varlığı ve sorunun farklı düzlemlerde giderek daha çok tartışılır hale gelmesi, üstesinden gelinebileceğine dair kanı oluşturmaktadır.

Anahtar Kelimeler: Yapay Zekâ Etiği, Feminist Yapay Zekâ, Toplumsal Cinsiyet Eşitsizliği, Epistemik Adaletsizlik, Toplumsal Adaletsizlik