# MASTER IN MANAGEMENT (MIM)

## MASTERS FINAL WORK

### DISSERTATION

# INFLUENCE OF KNOWLEDGE AND TRUST ON THE PERCEPTIONS OF ALGORITHMIC GENDER BIAS IN ARTIFICIAL INTELLIGENCE

MARLOES SUSAN HAITSMA

MARCH 2024

# MASTER IN MANAGEMENT (MIM)

## MASTERS FINAL WORK

### DISSERTATION

# INFLUENCE OF KNOWLEDGE AND TRUST ON THE PERCEPTIONS OF ALGORITHMIC GENDER BIAS IN ARTIFICIAL INTELLIGENCE

**MARLOES HAITSMA**
**SUPERVISOR: PROF. SARA FALCÃO CASACA**

**JURY:**
**PRESIDENT:** PROF. JOSÉ MANUEL CRISTÓVÃO VERÍSSIMO
**RAPPORTEUR:** PROF. Ricardo Alcobia Rodrigues
**SUPERVISOR:** PROF. SARA FALCÃO CASACA

**MARCH 2024**

# Acknowledgments

I would like to thank my brother for helping me out during my thesis with all the questions and concerns. Thank you for all the patience you had even when I struggled to understand. Thank you for helping me out and being there for me while solo travelling on the other side of the world, I appreciate the time you dedicated to picking up all my facetime calls.

I would like to thank my parents for always supporting me through this process and mentally, my mum for hearing all my problems and my dad for listening and helping me brainstorm. Thank you for being an inspiration.

I want to thank Nele, '*my big sister*', for all the support and being there when I struggled and was stuck. For helping me brainstorm and having the patience to explain. Thank you for always being there.

All thanks to my friends, especially those I met at ISEG, Lena, Erica and Anne, who made the writing of my thesis a much more pleasant process as we did it together. It was nice to not have to do this alone - thank you for all the fun times we had as distractions.

Furthermore, I want to thank my supervisor Professor Sara Falcão Casaca for all the guidance. Furthermore, I want to thank all respondents that took time out to fill in the survey.

# Abstract

Artificial Intelligence (AI) is used more frequently, automating the workplace each day, therefore becoming a part of each person's personal and work life. The AI we know of today needs data to learn from. However, there is a gender data gap, which means there is a gap of women in the data that exists today. Furthermore, when there is bias in society, and this is reflected in the data Artificial Intelligence uses, it can learn to make those same biased-decisions. It is therefore important for people to understand the systems they use can be making decisions that may not be unbiased.

This dissertation therefore aims to understand if knowledge in AI and trust in AI have any influence on how individuals in the tech sector perceive gender bias in AI. Furthermore, to understand if gender has any influence on this as well. To understand this a survey was utilized to gather perceptions on this topic.

The results show that when the knowledge in AI is higher, there is a higher perception of gender bias in AI. Furthermore, when an individual has a lower trust in AI there is also a higher perception of gender bias in AI. Gender also has an impact, as it is shown that men have on average a lower perception of gender bias in AI than women. Furthermore, the effect of trust in AI and knowledge in AI on the perception of algorithmic gender bias is weaker for men than for women.

# Table of contents

# 1. Introduction

Artificial Intelligence (AI) is a part of everyday life, and getting more accurate by the day (Goldstaub, 2020). From introducing us to new songs and TV shows by recommending algorithms to self-driving cars, Artificial Intelligence is continuously automating daily tasks and being adopted at an unprecedented rate. Therefore, becoming much more involved in our lives (DeVerter, 2023; Goldstaub, 2020; Manasi et al., 2023). In organisations AI is continuously used to improve and change our decision-making processes, as well as changing the way we are working (Bannon, 2023; Goldstaub 2020). AI is replacing and making manual, repetitive jobs more efficient - improving labour efficiency, reducing costs but also leaving room to create new jobs (Bannon, 2023; Goldstaub 2020; Zhang & Lu, 2021).

The Artificial Intelligence that exists and works today, is also called Artificial Narrow Intelligence (AI). Artificial Narrow Intelligence could be defined as “an umbrella term for a system that can perform a single intelligent task” (Goldstaub, 2020, p.10). AI therefore refers to a machine that tries to mimic and recreate human intelligence (Goldstaub, 2020). Furthermore, it gives the ability to understand, analyse and reveal patterns of big data, which would be too complex and large for humans to do so (Goldstaub, 2020; Madgavkar, 2021). In this thesis, the term AI is discussed and researched in the context of narrow AI.

When discussing AI, the term machine learning (ML) is often used as a synonym, however, it is to be noted that machine learning is a subset of AI (Fletcher et al., 2021; Kelan, 2023). While AI is an umbrella term for machines that try to replicate human intelligence, ML discusses machines that need data to create the methods and algorithms for decision-making models (Fletcher et al., 2021; Kelan, 2023;).

Artificial Intelligence has the capability to make life more efficient, happier and healthier, but it could have a negative impact such as widening the poverty gap or increasing inequality (Goldstaub, 2020). Therefore, with the rapid increase of AI in today's age, there are ethical implications that need to be considered (DeVerter, 2023).

In the book *'Invisible women; exposing data bias in a world designed for men'*, Caroline Criado Perez (2019) describes how in our history and lives men have been the default; representing humans overall. While for the other half of humanity, there is often just silence. The absence of women in our history, current lives, and future narratives, leaves according to Caroline Criado Perez 'the gender data gap' (Perez, 2019). The absence of women in our data is not just about silence - it will have consequences. Small, irritating consequences, such as not reaching the top shelf as it is set to the average male height, might not be deadly, but being a woman with an undiagnosed heart attack as the symptoms are 'atypical', is (Perez, 2019). Furthermore, according to research by Gartner (2018), in 2022 85% of AI projects would have inaccurate outcomes due to bias in the data and algorithms.

This gender data gap is probably not intentional or with any bad intent, but it is the impact of a deeply male-dominated culture (Perez, 2019). Men are thus the default and go without saying. This way of thinking has been around for centuries - *not being aware* that women are not said at all. In algorithms this gender data gap even doubles, as there is the gap in the creators of the algorithms as well as society as a whole; of the knowledge of how AI is discriminatory (Perez, 2019) Therefore, this all becomes even more dangerous with the upcoming of Big Data and AI (Perez, 2019).

The research objective is therefore to analyse what influences the perception of employees and managers in the tech sector regarding this gender bias in AI. This paper aims to understand if the factors knowledge, trust and gender influence the perception of AI users of gender bias in AI. The research is executed mainly in Portugal, the Netherlands and remaining European countries. The following questions are used as guidance in understanding the perceptions of ethical issues regarding gender bias in AI:

- Does knowledge and trust in AI influence the perception of managers and employees in the tech sector regarding algorithmic gender bias in AI?
    - Does knowledge in AI influence the perception of algorithmic gender bias in AI?
    - Does trust in AI influence the perception of algorithmic gender bias in AI?

This thesis is introduced with the existing problem of gender bias in AI, in the context of the ICT sector and the main questions are presented. The existing literature of ethical issues, gender bias and perceptions of AI are presented. This is followed by the methodology used for the data analysis. The data analysis and discussion are presented, followed by the conclusion. Lastly the limitations and future research are discussed.

# 2. Literature Review

## 2.1 What is AI

According to Razavian et al. (2020) AI “is a broad field that investigates constructing computer programs that exhibit some form of intelligent behaviour.” p.1. AI stimulates human intelligence behaviours by utilising computers to simulate behaviours such as decision-making, judging and learning (Zhang & Lu, 2021). It is a collection of disciplines such as biology, computer science, psychology and more (Zhang & Lu, 2021). Examples of AI include speech recognition, image processing, natural language processing such as translating texts and playing board games (Razavian et al., 2020; Zhang & Lu, 2021).

***Machine Learning***

The most known and dominant approach of AI is Machine learning. The main scope of ML is to use an algorithm that learns from data to improve its performance (Zhang & Lu, 2021). Machine learning learns from the data that is put into a model to perform a specific task (Razavian et al., 2020). This model has an output as well, which is what it tries to predict. The algorithm also has a training objective, that indicates how well the model predicts the output based on the input of the training data (Razavian et al., 2020). The models that contain parameters (parametric models), works by showing the learning algorithms the training data - which helps the model tweak the parameters to better fit the data (Razavien et al., 2020).

An example of machine learning is if you input photos of dogs into the model, the model then learns the characteristics of the dogs, and the output is then to recognize dogs in other pictures- based on the characteristics (Zhang & Lu, 2021)

## 2.2 Ethics in Artificial Intelligence

According to Siau and Wang (2020), ethics help in understanding what is good or right by forming a system of guidelines or principles. Ethics discusses the moral obligations and responsibilities of humans. Looking at AI, it discusses this in the context of the people creating AI and Artificial Intelligence in itself (Siau & Wang, 2020). In short, ethics of AI discusses the ethical issues that can occur during the design and development stage of AI, as well as the issues that are caused by AI. Ethical issues during the designing and developing can be for example the biases that are in the data or data privacy (Siau & Wang, 2020).

Machine learning can thus have outcomes that are unfair for certain minority groups (Oneto & Chiappa, 2020). Therefore, ML fairness is increasingly becoming an area in ML to ensure that the outputs become fairer (Castaneda et al., 2022; Oneto & Chiappam, 2020). Fairness is perceived by people subjectively, relying mostly on intuition and emotive information (Narayanan et al., 2023). The decisions on how to ensure fairness can depend on the existing ethical frameworks and legal frameworks (Fletcher et al., 2021; Narayanan et al., 2023). Organizations increasingly try to promote the development of AI to be fair, however, there are still concerns about fairness of AI (Narayanan et al., 2023).

***Importance of ethical Artificial Intelligence***

AI is approaching achieving activities that are close to or even surpassing human intelligence, such as facial recognition systems, diagnosing cancer cells, and optimizing processes in organizations (Stahl et al., 2021). As Artificial Intelligence increasingly affects people's lives due to the expanding capabilities and further development, human rights such as freedom of speech and equal treatment should be on the forefront when developing (Siau & Wang, 2020; Stahl et al., 2021).

Ethical AI is relevant as Machine Learning requires human factors to function. The functioning of Narrow AI is based on training data that has been given and based on the programming; and this is influenced by big data and people (Manasi et al., 2022; Siau & Wang, 2020). With the increase of big data (large training datasets) it can especially be at risk of misuse when the data contains personal and private information. This becomes even more tricky when this data is not representative of all groups of society (Siau & Wang, 2020; Stahl et al., 2021).

Machine learning has a ‘black box.’ The black box describes the processing action of ML that creates an algorithm. However, it is unknown by the creators nor users on how this processing action exactly works (Kelan, 2023; Siau & Wang, 2020; Stahl et al., 2021). This leads to a lower understanding of the technology and a lower information symmetry among the users and experts (Siau & Wang, 2020). Furthermore, the black box can cause or even hide issues, such as unrepresentative datasets, and issues in the data quality that can cause a bias in the algorithms (Stahl et al., 2021). Therefore, the trust people have in technology lowers as well (Siau & Wang, 2020).

### *The main ethical Issues*

According to Stahl et al. (2021), reliability, control of data and lack of transparency are three main concerns related to ethical issues in Ai, and specifically ML. The concern and group of reliability covers the outputs of the system and the system itself. As ML learns from data, the issue is if the data is of quality and accurate enough. If AI is used in healthcare for example, it needs to be reliable enough as it impacts the health of people. Furthermore, if in gender-biased ML historical data is used for example, it can be that it is not representative enough for women compared to men. Therefore, the output of the system can become unreliable and inaccurate and therefore biased (Stahl et al., 2021). The control of data is about the concerns discussing confidentiality and privacy. As the issue of personal data can cause vulnerabilities of connecting data and identification of individuals. Lastly, the lack of transparency, discusses the concern of the ‘blackbox’ (Stahl et a., 2021).

Furthermore, there are issues that arise from digitalization, and therefore the increase of the use of AI. According to Stahl et al. (2021), this can include economic concerns, concerns about justice and freedom, societal concerns, and other unknown issues. Table 1 shows an overview with examples (Stahl et al., 2021).

Table 1 Ethical Issues in AI

| **Issue** | **Examples** |
|---|---|
| Economic | Changes to employment, changes in nature of work, concentration of economic power (e.g. growing dominance of big internet companies). |

| | |
|---|---|
| Justice | Impact on individual groups, access to public services (especially groups that are vulnerable), fairness such as bias and discrimination, fairness of economic activity distribution. |
| Freedom | Increase of the independence of AI may decrease the independence of humans, harm to physical well-being, impact on health |
| Social issues | Use of Ai for the military, environmental consequences, democracy concerns |
| Unknown | Lack of knowledge, criminal future uses, focusing on certain issues now may divert resources from more crucial problems. |

Source: Stahl et al., 2021, p.26-27

The issues above discuss Narrow AI. However, concerns that arise of general AI, are issues that could arise if AI would be able to change its nature or those of humans. These issues are called metaphysical (Stahl et al., 2021). Ethical values are shared cultural and geographical borders, including values such as justice, freedom, trust, solidarity, transparency, and dignity (Stahl et al., 2021)

## 2.3 Algorithmic Gender Bias in Artificial Intelligence

It is often indicated that AI improves inclusion and diversity, however, according to Kalen et al. (2023) research shows that algorithmic bias can be present in AI. Unrepresentative datasets, biases in historical data and collection bias are all issues in machine learning, as these can influence bias to arise in algorithms. These biases shown in the algorithms are repeated and amplified human biases (Castaneda et al., 2022; Kalen et al, 2023; Manasi et al., 2022).

The Oxford dictionary (as cited in Castaneda et al., 2022, p. 3) describes an algorithm as “a set of rules that must be followed when solving a particular problem”. Therefore, according to (Fletcher et al., 2021, p.6), bias in algorithms can be defined as “a systematic error or an unexpected tendency to favour one outcome over another”. To be able to combat bias in AI, it is relevant to discuss and define fairness (Ntoutsi et al., 2020). Moreover, algorithmic bias is described as the undesired dependency of specific attributes in the data that are assigned to a demographic group (Fletcher et al., 2021). There are several types of bias existing, all to be divided into implicit and explicit causes. Explicit bias can be caused by issues with data sampling, design of equipment and unexpected behaviour in data collection. Often the problems occur in unbalanced data (Fletcher et al., 2021). Implicit bias is caused by unexpected connections and links between variables (Fletcher et al., 2021).

***Societal factors in data***

Algorithmic bias can be recognized as a function of how mutually constitutive technology and society are, aside from referring to the errors in systems that lead to unfavourable outcomes for certain groups (Kalen et al., 2023). When there is bias in our society, and this enters the systems, the bias would be amplified; resulting in replicating the already existing inequalities (Ntoutsi et al., 2020). When specific social groups are continuously (dis)advantaged by institutions, institutional bias takes place. This is caused by continuously going by current norms, and as institutions often use algorithms, machine learning can reinforce those biases (Ntoutsi et al., 2020).

This means that society shapes technology, while technology can also shape society. As algorithmic bias is created by the dynamic exchange between technology and society - algorithmic inclusion would also be possible (Kalen et al., 2023).

Natural language processing (NLP) is a field of AI that refers to the linguistics in computer systems. Word embedding algorithms is an example of NLP, and this can include stereotypical bias that amplifies stereotypes (Castaneda et al., 2022). NLP learns mostly from text online, and therefore when there is a bias present in the outcomes, it is a mirror for the biases in society (Castaneda et al., 2022).

***Algorithmic design***

When data bias occurs in algorithms that are frequently used for longer periods, the outputs can impact the future learning of algorithms. The decisions made in those algorithms can include results that can be biased due to the data bias in the original input (Castaneda et al., 2022; Manasi et al., 2022). When future models and training collect these results, the same problems of bias will occur - creating a feedback loop where algorithms keep learning from biased data (Castaneda et al., 2022; Manasi et al., 2022). Therefore, sustaining the discrimination and deepening the bias (Castaneda et al., 2022; Manasi et al., 2022).

There are four known biases: aggregation, social, temporal and algorithmic-related bias (Castaneda et al., 2022). Aggregation bias happens when untrue conclusions are drawn for one subgroup, as the information is based on another one. Temporal is caused by the differences between behaviours and populations during a period (Castaneda et al., 2022). Social bias manifests when the actions or content generated by others significantly influence judgments of people (Castaneda et al., 2022). While algorithmic bias discusses the unfair outcomes, favouring a certain group of another, in computer systems due to systematic and repeating mistakes. Algorithmic gender bias refers to the discrimination of one of the genders is discriminated against (Castaneda et al., 2022).

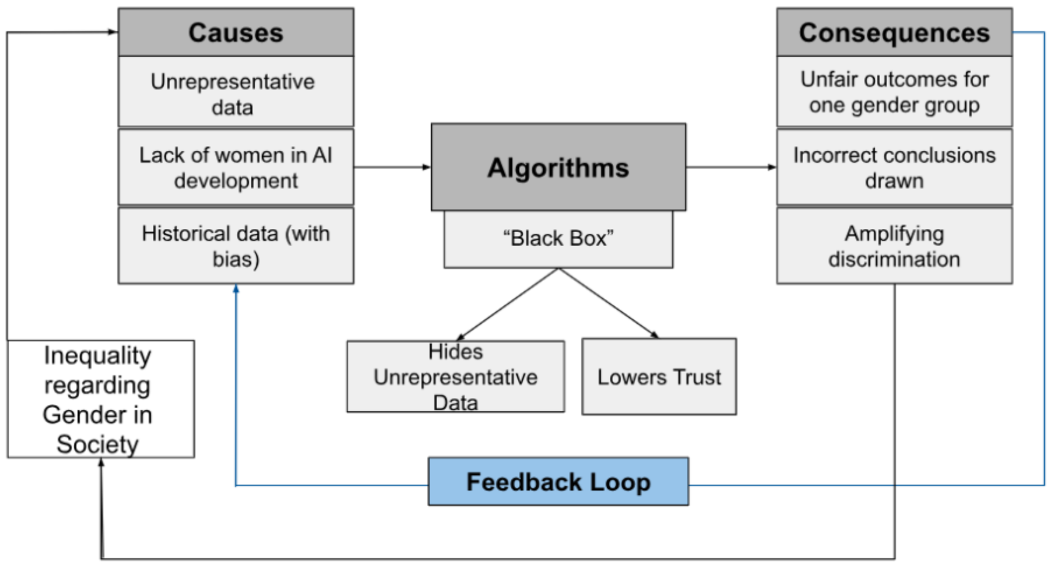
Aside from the lack of quality and representative data, the presence of gender bias in AI can also be led back to the lack of women in the workforce of AI (Castaneda et al., 2022). Women are not as present in the designing of these algorithms, as compared to men (Manasi et al., 2022). According to UNESCO (2019) (as cited in Manasi et al., 2022) 12% of researchers in AI are women and 6% of software developers are women.

Gender bias appears in different fields in AI. As described above it can be present in NLP. Furthermore, speech recognition is a field of AI that demonstrates gender bias, as Castaneda et al., (2022) describes that in audio analysis has shown to struggle more with higher pitched voices. According to Castaneda et al., (2022) this could be due to databases lacking data of women voices. The same matter applies in face recognition, as the algorithms rely on gender stereotypes, which might cause the algorithms to form untrue beliefs on the appearance of certain groups (Castaneda et al., 2022; Fletcher et al., 2021). Furthermore, datasets are commonly primarily male and Caucasian, causing the algorithms to better identify this group compared to others (Castaneda et al., 2022)

AI is employed in decision making more often, for example in regarding loans, study applications and allowing rent. When gender bias is present in decision management it can have significant social negative consequences (Castaneda et al., 2022). As the algorithms in decision management often rely partly on historical data. Furthermore, even when removing the variable of gender, the algorithm could still end up biased which has been seen in automated hiring tools (Castaneda et al., 2022; Kelan, 2023).

As described above it can be said that there are several variables impacting gender bias in AI. Figure 1 illustrates this process, including the feedback loop.

Figure 1 Gender Bias in Machine Learning




Source: Adapated from Castaneda et al., 2022; Fletcher et al., 2021; Kalen et al., 2023; Manasi et al., 2022; Ntoutsi et al., 2020; Siau & Wang, 2020; Stahl et al., 2021

## 2.4 Perceptions regarding Artificial Intelligence and Gender Bias

**Awareness regarding Ethical Artificial Intelligence**

According to Vakkuri et al. (2020) (as cited in Pant et al., 2023, p.18), the level of awareness of the ethics in AI among the practitioners of AI is lacking. The level of awareness of ethics in AI, for AI practitioners, includes workplace rules and policies, complaints from customers, news and media, university courses and personal experiences and interest (Pant et al., 2023). Furthermore, only being aware of AI ethics is not sufficient, as the understanding of the topic is essential for AI practitioners to develop AI in an ethical manner (Pant et al., 2023). According to research by Pant et al. (2023, p.8), the level of familiarity with ethics in AI is 41%, which means there is a lack of efforts needed to increase awareness of AI ethics among those working in the field of AI.

**Trust in Artificial Intelligence**

According to Lee and Rich (2021), individuals tend to perceive algorithmic decisions as lesser than human decisions and therefore have resistance towards adhering to the algorithmic decisions. Trusting algorithmic decisions less by individuals was especially present when the tasks required to have human capabilities, involve subjectivity, and need attention of people's individuality (Lee & Rich, 2021). An example can be seen in recruitment interviews and candidate selections, where people perceive humans to be fairer in decision making than AI (Narayanan et al., 2023). Therefore, in tasks that are complex and where the capabilities of humans are needed, perception of AI fairness is heavily influenced - often in which AI is perceived as less fair (Narayanan et al., 2023). Trust is the belief that someone or something will support someone in achieving the desired outcome in a challenging or uncertain situation (Lee & Rich, 2021). The trust of human decisions and algorithmic decisions change for marginalised groups; as those groups often have less trust in human decisions (Lee & Rich, 2021). Individuals with higher mistrust in human systems, perceive human and algorithmic decisions as equally fair and to be trusted - while people with a lower mistrust in human decisions, perceive algorithmic decisions to be less fair (Lee & Rich, 2021).

Individuals that are impacted by organisational decisions are becoming increasingly concerned about the fairness of AI, and if those AI systems would make decisions with care and empathy (Narayanan et al., 2023). Especially if the AI systems can make decisions that are equitable (Narayanan et al., 2023). Furthermore, according to Gupta et al., (2021) AI questionability due to gender bias is higher when a person has the cultural values of collectivism, masculinity, and uncertainty avoidance. Moreover, women have a higher AI questionability than men, making women more distrustful of algorithm decisions. Moreover, individuals that are active on the internet are also more likely to question the recommendation of AI (Gupta et al., 2021). In recruitment, screening candidates have increasingly been done with the use of AI systems. However, companies that use those AI systems can make job applicants feel more anxious, especially when the AI system is perceived to be biased against a certain race or gender (Gupta et al., 2021). Individuals that are affected by biased AI systems are often responding more negatively and are less forgiving towards corporations that use those systems (Gupta et al., 2021).

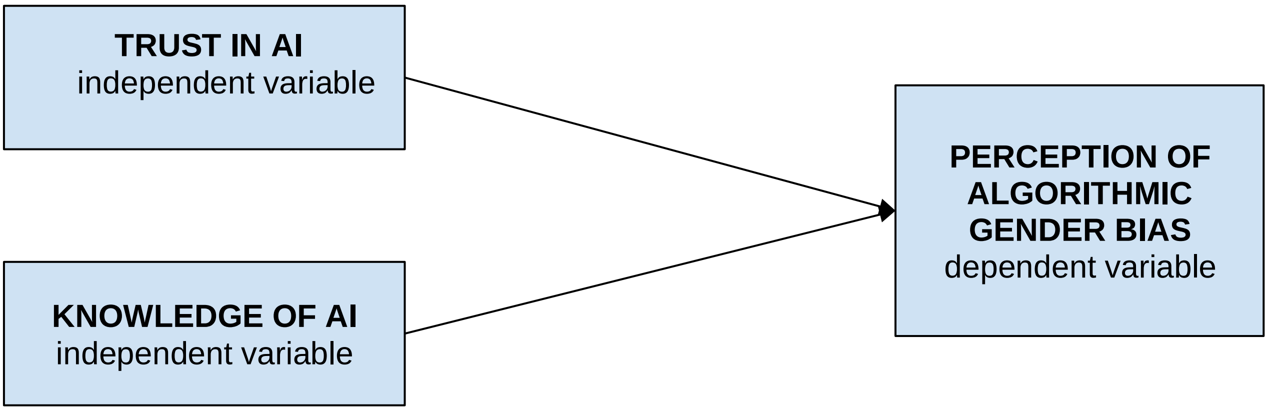
**Influence of gender on perception of gender bias**

There is a lack of literature in this area. According to research from García-González et al. (2019), men and women perceive the gender gap differently. Women perceive and have a higher awareness of the existing gender inequality compared to men in academia. Furthermore, women support increasing awareness and addressing gender inequality by implementing measures more than men. Research by Otterbacher et al. (2018), investigated the impact of sexist beliefs on recognizing gender bias in image search. This research shows that men show higher beliefs of sexism and are therefore less likely to recognize this gender bias.

## 2.5 Conceptual Model

Figure 2 shows the conceptual model with the relation of the variables.

Figure 2 Conceptual Model




**Research questions and Hypotheses**:

***Research Question 1: Does knowledge in AI influence the perception of algorithmic gender bias in AI?***

*H0: Knowledge in AI is not significantly associated with the perception of algorithmic gender bias in AI.*
*H1: Knowledge in AI is associated with the perception of algorithmic gender bias in AI.*

***Research Question 2: Does trust in AI influence the perception of algorithmic gender bias in AI?***

*H0: Trust in AI is not significantly associated with the perception of algorithmic gender bias in AI.*
*H1: Trust in AI is associated with the perception of algorithmic gender bias in AI*

The conceptual model is grounded on the literature review. Due to the lack of literature on the influence of gender on the individuals' perceptions of gender bias in AI, gender has not been included in the model and in the outlined hypotheses. However, the following research question has been raised:

***Is there any difference in the way in which men and women perceive gender bias in AI?***

# 3. Methodology

## 3.1 Research purpose and strategy and data collection

The focus of this paper is to explore people's perceptions of the topic of understanding managers' and employees’ perception of gender bias in AI in the ICT sector. As there is not much known about the perception of individuals of gender bias, the research purpose is exploratory research (Sekaran & Bougie, 2016). This research focuses on understanding if trust in AI and knowledge influences the perception of gender bias. Furthermore, quantitative research is used using a survey, therefore primary data was collected. As this paper is focused on the perception of individual employees and managers, the unit of analysis is individual. The data is derived at once, meaning the study is cross- sectional (Sekaran & Bougie, 2016). The survey was done in an environment without having interference in the answers, meaning this research was conducted in a non-contrived setting (Sekaran & Bougie, 2016).

The survey was distributed using social media, such as LinkedIn and Instagram. Moreover, the survey was distributed using personal contacts of the researcher. The target group was focused on people working in the tech sector. There were 107 respondents in total, whereas 22 are not in the tech sector, meaning those are not included. The survey was online for the month of February 2024. The total respondents that were used for the analysis was therefore 85. The aim for this research was to collect a total of 100 perspectives of people working in the tech sector, however, there are 15 respondents lacking and this has not been achieved.

According to Investopedia the tech sector can be defined as a sector of the economy that consists of businesses focusing on electronics, software, computers, social media, and other industries related to information technology (Frankenfield, 2022). The survey had three main sections: demographic characteristics, knowledge of AI, trust in AI, and the perception of ethics regarding gender bias in AI. The survey is based on the conceptual model (See Figure 2). The questions can be found in the appendix. The research of Gupta et al., (2021) was used as a basis of this survey regarding the scenarios asked. Gupta et al., (2021) researched if the cultural values of individuals impact the questionability of AI recommendations. However, instead of asking the participants about the questionability, it was asked how likely the participants would question the outcome. The other questions were inspired from sources such as Castagno and Khalifa (2020) and Gupta et al., (2021).

The questions that asked can be found in Appendix. 1. Table 2 shows which questions are used to analyse which variable.

Table 2, Survey Questions used for regression analysis

| Knowledge | Trust | Perception of Bias |
|---|---|---|
| Q9 How would you rate your overall knowledge of AI? | Q14 How much do you trust the decisions made by AI systems? | Q18 You are both applying for the same financial product (such as a credit card or home mortgage/loan) on the same bank app/website using your own devices. You notice the products that are offered to your friend charge higher interest rates than those offered to you.<br>How likely are you to question this outcome? |
| Q11 How often do you read about AI-related topics? | Q15 In your opinion, how likely is it that AI is more fair in its decisions than humans? | Q19 You are both looking for similar jobs on the same employment app/website using your own devices. You notice the jobs that are offered to your friend usually have lower-paying salaries than those offered to you.<br>How likely are you to question this outcome? |
| Q13 How confident are you in explaining the basic concepts of AI to others? | Q16 To what extent do you believe AI systems are biased in their decision-making processes? | Q20 You both have the same nationality and are at the airport going through the same automated immigration kiosk that uses face recognition technology to verify travelers' identity. The automated immigration kiosk directs your friend to see an immigration officer while you are cleared to go through.<br>How likely are you to question this outcome? |
| | Q17 To what extent do ethical considerations influence your trust in AI? | Q21 You are both booking a similar hotel room using the same hotel booking app/website using your own devices. Hotel rooms offered to your friend have higher prices than those offered to you.<br>How likely are you to question this outcome? |
| | | Q22 How familiar are you with the concept of gender bias in AI? |
| | | Q23 In your opinion, how likely is it that AI would discriminate towards one gender? |
| | | Q24 To what extent do you believe that AI systems exhibit gender bias in their decision-making processes? |
| | | Q25 How concerned are you about the potential impact of gender bias in AI systems on society? |

| | | |
|---|---|---|
| | | Q27 How much do you think societal stereotypes contribute to gender bias in AI? |
| | | Q28 How likely are you to advocate for measures to address and mitigate gender bias in AI? |

Questions adapted from: Gupta et al. (2021) and Castagno and Khalifa (2020).

Knowledge and trust in AI are the independent variables, as to research if those factors have any influence on the perception of algorithmic gender bias (dependent variable). All questions considering all the three variables were taken together and an average was created using scales which were reverted to 1-5. A regression analysis was performed in STATA.

Furthermore, as there has been a lack of research on the impact of gender, the researcher had found interest in understanding if gender has any impact on the variables (trust in AI, knowledge in AI, and perception of algorithmic gender bias). Therefore, during the data analysis gender had been included as a slope dummy to understand and learn if there is any correlation regarding trust and knowledge in AI. As well as a regression including gender aside from trust and knowledge in AI.

## 3.2 Validity, reliability, ethics, and data management

To ensure validity, the same survey was sent out to all participants to ensure the data collection to be consistent. Moreover, the survey's answers have been analysed using STATA. The reliability of the research was ensured through data cleaning, using established scales, pilot testing the survey, and a sufficient sample size.

Voluntary participation, consent and confidentiality guaranteed the ethics in the research. Survey respondents had autonomy to decide to participate and at any time were allowed to stop the survey in question. The survey was distributed non-intrusively. All information necessary to consent was shared such as the content purpose and scope of the research.

To ensure confidentiality only general descriptions of the survey participants are shared and asked in this paper, but no data to be directed to an individual is shared. Therefore, the participants of this research remain anonymous. All data is referenced correctly using APA referencing to avoid plagiarism. Moreover, if any sensitive data was shared by the participants, permission was asked.

# 4. Data Analysis and Discussion

## 4.1 Descriptive Statistics

### 4.1.1. Sociodemographic description

There are a total of 85 observations (N= 85), which included 40 women (47.06%) and 45 men (52.94%). 10.59% of the respondents are senior managers, 14.12% are middle managers and 75.29% of the respondents could be classified as employees. 51.76% of the respondents are 25 or below, 20% are 26-35, 11.76%, 36-45 are between 36 and 45 and both the group of 45 until 55, and 56 and above are 8.24% of the respondents. The majority of respondents are from the Netherlands with 54.12%, followed by Portugal with 29.41% of the respondents. Remaining respondents are from Germany, Israel, Finland, Belgium, Italy, Luxembourg, and the United Kingdom. Most respondents are from the cyber security industry (28.24%), E-commerce, social media, networking (15.29%) and Artificial Intelligence (12.94%). Most respondents have less than 1 year (31.76%) or 1 to 2 years' experience working with AI (34.12%).

### 4.1.2 Description of responses

40% of respondents classifies the overall knowledge of AI as average. While 30.59% considers their knowledge limited, and 27.06% as above average. 47.06% of respondents reported having taken courses or training, while 45.88% stated that they read about AI-related topics occasionally.

The respondents answered the question of “how familiar are you with the concept of gender bias in AI” evenly distributed, with “None at all familiar” (25.88%), “Slightly familiar” (23.53%), “Somewhat familiar” (23.53%) and “Moderately familiar” (18.82%), while only 8.24% of respondents stated to be extremely familiar. Furthermore, the findings show that a majority of respondents (85.53%) have not personally observed instances of gender bias in AI systems. When asked if the respondents would advocate for measures to address and mitigate gender bias in AI, 40% responded with “Somewhat likely” to advocate for measures, while 9.41% responded “Extremely unlikely” and 16.47% responded with completely. Moreover, opinions on the likelihood of AI systems discriminating towards one gender varied, as 38.82% considered it “Somewhat likely”, 21.18% “Somewhat unlikely” and 7.06% “Extremely unlikely”,

When asked about the concerns of the respondents regarding the potential impact, 7.06% is “Not concerned at all”, while 28.24% answered “Slightly concerned” and 12.94% answered with “Extremely concerned”. Additionally, for the question if AI is fairer in its decisions than humans, the responses were mostly between “Somewhat unlikely” (31.76%) and “Somewhat likely” (24.71%), with only one respondent (1.18%) answering “Extremely likely” and 8.24% answering “Extremely unlikely. When directly asked if the respondents trust AI, the majority of people answered a “moderate amount” (60%), while 1.18% answered “none at all”, and 1.18% answered a great deal.

When cross tabulating the years of experience and overall knowledge, most respondents with less than 1 year experience rate their knowledge as “Limited” and “Average”, while respondents with more experience tend to rate their knowledge as “Average” and “Above average”. Women rate their overall knowledge as “Limited” or “Average”, while men have a more balanced distribution, also including “Above average”.

## 4.2 Regression Analysis

### 4.3.1 Regression Perception, Trust, and Knowledge

A multiple regression analysis was performed, to understand the relationship between the dependent variable of the perception of algorithmic gender bias and the independent variables of knowledge and trust in AI. Table 3 shows the results. In the appendix 2.2, the output in STATA can be found.

The model shows an overall statistical significance at a 0.05 significance level ($F(2,82) = 18.56$, $p = 0.0000$), meaning that at least one of the independent variables is impacting the dependent variable ceteris paribus. Furthermore, the r-squared value is 0.3116, showing that 31.16% of the variability of the dependent variable can be explained by the model. For each increase of knowledge, the perception of algorithmic gender bias increases by 0.3774, holding all others constant. For each increase of trust, the perception of algorithmic gender bias decreases by 0.588 ceteris paribus. Therefore, when an individual's knowledge increases, the higher the perception of algorithmic gender bias becomes. Furthermore, higher levels of trust are associated with a lower perception of gender perception.

The variable trust has a p-value of 0.000, and as $p < 0.05$ the variable is statistically significant, ceteris paribus. The variable knowledge has a p-value of 0.000, and as $0.000 < 0.05$, the variable can be said to be statistically significant, ceteris paribus.

Overall, the regression analysis shows that there is a statistically significant relationship, between knowledge in AI and the perception of algorithmic gender bias, ceteris paribus, therefore rejecting the null hypotheses of “Knowledge in AI is not significantly associated with the perception of algorithmic gender bias in AI”. Higher levels of knowledge in AI are thus associated with an increase of perception of algorithmic gender bias. Therefore, individuals are more aware and critical regarding gender bias in AI, when there is more knowledge in AI.

Furthermore, the regression analysis shows that there is a statistically significant relationship between Trust in AI, and the perception of algorithmic gender bias, ceteris paribus, therefore rejecting the null hypotheses of Trust in AI is not significantly associated with the perception of algorithmic gender bias in AI. Here it is shown that higher levels of trust in AI, are linked with a lower perception of algorithmic gender bias - therefore have a less critical perception of gender bias in AI when there is a higher trust in AI.

Table 3 Regression Results

| **Variable** | **Coefficient** | **Std. Error** | **t-value** | **p-value** |
|---|---|---|---|---|
| *Intercept* | *4.0183* | *0.4118* | *9.76* | *0.000* |
| Knowledge in AI | 0.3774 | 0.0967 | 3.90 | 0.000 |
| Trust in AI | -0.5883 | 0.1135 | -5.19 | 0.000 |

### 4.3.2 Regression including gender

As of the interest to see if gender can be associated with the dependent variable of perception of algorithmic gender bias, a multiple regression analysis was performed. Table 4 shows the results. In the appendix, the output in STATA can be found.

The model shows an overall statistical significance at a 0.05 significance level ($F(3, 81) = 14.31$, $p = 0.0000$), meaning that at least one of the independent variables is impacting the dependent variable ceteris paribus. Furthermore, the r-squared value is 0.3464, showing that 34.65% of the variability of the dependent variable can be explained by the model. In this model knowledge in AI has a p-value of 0.000 and trust in AI has a p-value of 0.000. As $p < 0.05$, both variables are statistically significant, when all others constant. Gender in this model has a p-value of 0.041, and is therefore statistically significant, all others constant as $p < 0.05$. As the coefficient is 0.2960 (women are coded as 0, and men as 1), men tend to have a lower perception score compared to women, even when accounting for the variables of trust and knowledge. This is consistent with the limited literature found regarding this topic. As García-González et al. (2019), women have a higher perception on gender inequality compared to men. Furthermore, research by Otterbacher et al. (2018), shows that men are less likely to recognize gender bias.

Table 4 Regression Results Gender

| **Variable** | **Coefficient** | **Std. Error** | **t-value** | **p-value** |
|---|---|---|---|---|
| *Intercept* | *3.887* | *0.4086* | *9.51* | *0.000* |
| Knowledge in AI | 0.4313 | 0.0983 | 4.39 | 0.000 |
| Trust in AI | -0.5442 | 0.1132 | -4.81 | 0.000 |
| Gender | -0.2960 | 0.1426 | -2.81 | 0.041 |

The R-squared value increased from 0.3116 to 0.3464 when adding gender to the model. Therefore, the overall model including gender explained more variance on the perception of algorithmic gender bias, compared to without. This can thus suggest that gender plays a specific role regarding the dependent variable perception of algorithmic gender bias.

As there was interest in understanding if gender made any impact, gender was included in a regression analysis. Here it is shown that in the overall model, the r-squared increased compared to the first model, suggesting that gender itself is a statistically significant factor and therefore has a specific role in shaping the perception. This shows that women tend to have a higher perception of algorithmic gender bias than men, even when accounting knowledge and trust.

## 4.3 Influence of Gender

To understand if gender has an impact on trust and knowledge, a regression analysis is performed using gender as slope dummy.

### 4.3.1 Slope dummy: Gender on Trust

The model shows an overall statistical significance at a 0.05 significance level ($F(3, 81) = 15.11$, $p = 0.0000$), meaning that at least one of the independent variables is impacting the dependent variable ceteris paribus. The R-square is 0.3588, therefore 35.88% of the variability of the dependent variable can be explained by the model.

All P-values of the variables are lower than 0.05 ($p < 0.05$), meaning that when all other constant, the independent variables are statistically significant. When trust increases, the expected change of the perception of algorithmic gender bias is different between genders. Looking at the coefficient, on average, for each increase in trust, the predicted perception decreases by an additional 0.116 for men compared to women, holding other variables constant. The slope dummy shows that trust has a weaker effect on gender perception for men than for women.

Table 5 Regression with slope dummy gender on trust

| Variable | Coefficient | Std. Error | t-value | p-value |
|---|---|---|---|---|
| *Intercept* | *3.6673* | *0.4249* | *8.63* | *0.000* |
| Knowledge in AI | 0.5022 | 0.1069 | 4.70 | 0.000 |
| Trust in AI | -0.5280 | 0.1129 | -4.68 | 0.000 |
| Slope Dummy: GenderTrust | -0.1163 | 0.0476 | -2.44 | 0.017 |

### 4.3.2 Gender on Knowledge

Table 6 Regression with slope dummy gender on knowledge

| Variable | Coefficient | Std. Error | t-value | p-value |
|---|---|---|---|---|
| *Intercept* | *3.6673* | *0.4248* | *8.63* | *0.000* |
| Knowledge in AI | 0.5022 | 0.1069 | 4.70 | 0.000 |
| Trust in AI | -0.5280 | 0.1129 | -4.68 | 0.000 |
| Slope Dummy: GenderKnowledge | -0.1163 | 0.4762 | -2.44 | 0.017 |

The model shows an overall statistical significance at a 0.05 significance level ($F(3, 81) = 15.11$, $p = 0.0000$), meaning that at least one of the independent variables is impacting the dependent variable ceteris paribus. The R-square is 0.3588 therefore 35.88% of the variability of the dependent variable can be explained by the model.

All P-values of the variables are lower than 0.05 ($p < 0.05$), meaning that when all other constant, the independent variables are statistically significant. When trust increases, the expected change of the perception of algorithmic gender bias is different between genders. Looking at the coefficient, on average, for each increase in knowledge, the predicted perception decreases by an additional 0.116 for men compared to women, holding other variables constant. The slope dummy shows that trust has a weaker effect on gender perception for men than for women.

Overall, the regressions including the slope dummy gender thus shows that the effect of trust and knowledge on the perception of algorithmic gender bias is weaker for men than for women. These findings show that gender moderates the relationship between the variables.

# 5. Conclusion

This research answered the research question: ***Does knowledge and trust in AI influence the perception of managers and employees in the tech sector regarding algorithmic gender bias in AI?*** With the respective sub questions:

- Does knowledge in AI influence the perception of algorithmic gender bias in AI?
- Does trust in AI influence the perception of algorithmic gender bias in AI?

To answer the main research question, a multiple regression analysis was performed. Here it was found that both trust in AI and knowledge in AI have a significant relationship with the perception of algorithmic gender bias in AI. The research shows that when individuals have a higher knowledge in AI, there is a higher perception of gender bias in AI. Therefore, individuals with a lower knowledge, have less awareness of algorithmic gender bias. When individuals have a higher trust in AI, there is a lower perception of gender bias in AI. Meaning that those individuals are less aware of gender bias in AI. It can thus be said that individuals that trust AI less, and have more knowledge in AI, have a higher perception of the algorithmic gender bias in AI.

To answer the question: Is there any difference in the way in which men and women perceive gender bias in AI?, a multiple regression analysis was performed. This regression shows that gender has an impact on the perception of algorithmic gender bias in AI. As women have a higher perception of gender bias compared in AI to men. As when gender is added to the model the variability (R-square) of the model increases. Furthermore, gender shows as a moderator variable for the variables trust and knowledge. As it is shown that the effect of trust and knowledge on the perception of gender bias in AI is less for men compared to women. In conclusion, there is a difference in how women perceive gender bias in AI compared to men.

## 5.2. Limitations and future research

The research limitations include the low sample size. As the sample size of 85 respondents can be seen as not representative enough, this could be a limitation. Furthermore, the age of the respondents is mainly 25 and below (and an employee) and there is a lack of diversity of country of residence, which can be seen as a limitation due to not being representative of the population. For future research it would be interesting to continue this research with a larger sample size. Furthermore, to include in gender in the conceptual model would be interesting for future research.

## 5.3 Policy and practical considerations

This thesis can add to policy and practical considerations, by helping to understand how knowledge and trust of AI has an impact on the understanding of gender bias. Regarding policy considerations, investing in AI education can make sure people are more aware regarding gender bias in AI. As well as it is important to address the gender gap, as men perceive gender bias less than women, according to the research above. Regarding practical considerations, organizations should educate its employees and managers regarding the gender bias in AI and its consequences that can be attached to it, keeping in mind the bridge between the perception between male and women.

For organizations to embrace a non-bias and ethical approach in the use of AI, it is important to educate the AI users on the possibility of gender bias within the systems and to use those systems with that in mind. So not to further amplify the bias in organizations. Furthermore, organizations should be transparent in the AI systems that are being used, and prioritize explainability for its employees and managers to have a better understanding of the systems that are used and how gender bias can manifest in those. The systems that are used should have continuous monitoring and evaluation to check on the biases in the AI systems. Furthermore, AI ethical frameworks should be deployed for the right use of AI systems within companies.

# REFERENCES


Baker-Brunnbauer, J. (2020). Management perspective of ethics in artificial intelligence. *AI And Ethics*, *1*(2), 173–181. https://doi.org/10.1007/s43681-020-00022-3

Bannon, M. T. (2023, June 22). How AI is changing the future of work. Forbes. https://www.forbes.com/sites/marenbannon/2023/06/22/how-ai-is-changing-the-future-of-work/

Castagno, S., & Khalifa, M. (2020). Perceptions of artificial intelligence among healthcare staff: A Qualitative survey study. *Frontiers in Artificial Intelligence*, *3*. https://doi.org/10.3389/frai.2020.578983

Castaneda, J., Jover, A., Calvet, L., Yanes, S., Juan, A. A., & Sainz, M. (2022). Dealing with gender bias issues in data-algorithmic processes: a social-statistical perspective. Algorithms, 15(9), 303. https:// doi.org/10.3390/a15090303

Choung, H., David, P., & Ross, A. (2022). Trust and ethics in AI. AI & SOCIETY, 38(2), 733–745. https://doi.org/10.1007/s00146-022-01473-4

DeVerter, J. (2023, August 9). From Curation To Creation: How Ethical AI Can Shape A Responsible Future. *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2023/08/09/from-curation-to-creation-how-ethical-ai-can-shape-a-responsible-future/

Fletcher, R., Nakeshimana, A., & Olubeko, O. (2021). Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Frontiers in Artificial Intelligence*, *3*. https://doi.org/10.3389/frai.2020.561802

Frankenfield, J. (2022, January 2). *Technology Sector: Definition, 4 major sectors, Investing in tech*. Investopedia. https://www.investopedia.com/terms/t/technology_sector.asp

García-González, J., Forcén, P., & Jimenez-Sanchez, M. (2019). Men and women differ in their perception of gender bias in research institutions. PLOS ONE, 14(12), e0225763. https://doi.org/10.1371/journal.pone.0225763

Gartner. (2018, February 13). *Gartner says nearly half of CIOs are planning to deploy artificial intelligence*. https://www.gartner.com/en/newsroom/press-releases/2018-02-13-gartner-says-nearly-half-of-cios-are-planning-to-deploy-artificial-intelligence

Goldstaub, T. (2020). *How To Talk To Robots: A Girls' Guide To a Future Dominated by AI*. HarperCollins UK.

Gupta, M., Parra, C. M., & Dennehy, D. (2021). Questioning racial and gender bias in AI-based recommendations: Do espoused national cultural values matter? *Information Systems Frontiers*, *24*(5), 1465–1481. https://doi.org/10.1007/s10796-021-10156-2

Kelan, E. K. (2023). Algorithmic inclusion: Shaping the predictive algorithms of artificial intelligence in hiring. Human Resource Management Journal, 1–14. https://doi. org/10.1111/1748-8583.12511

Lee, M. K., & Rich, K. R. (2021). Who Is Included in Human Perceptions of AI?: Trust and Perceived Fairness around Healthcare AI and Cultural Mistrust. *CHI Conference on Human Factors in Computing Systems (CHI '21)*. https://doi.org/10.1145/3411764.34455700

Madgavkar, A. (2021, April 7). *A conversation on artificial intelligence and gender bias*. McKinsey & Company. https://www.mckinsey.com/featured-insights/asia-pacific/a-conversation-on-artificial-intelligence-and-gender-bias

Manasi, A., Panchanadeswaran, S., Sours, E., & Lee, S. (2022). Mirroring the bias: gender and artificial intelligence. *Gender, Technology and Development*, *26*(3), 295–305. https://doi.org/10.1080/09718524.2022.2128254

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, *54*(6), 1–35. https://doi.org/10.1145/3457607

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernández, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., . . . Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *WIREs Data Mining and Knowledge Discovery*, *10*(3). https://doi.org/10.1002/widm.1356

Narayanan, D., Nagpal, M., McGuire, J., Schweitzer, S., & De Cremer, D. (2023). Fairness Perceptions of Artificial Intelligence: A review and Path forward. *International Journal of Human-Computer Interaction*, *40*(1), 4–23. https://doi.org/10.1080/10447318.2023.2210890

Sekaran, U., & Bougie, R. (2016). Research methods for business: a skill-building approach (7 ed.). Chichester: John Wiley & Sons

Siau, K., & Wang, W. (2020). Artificial Intelligence (AI) Ethics. *Journal of Database Management*, *31*(2), 74–87. https://doi.org/10.4018/jdm.2020040105

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2021). Organisational responses to the ethical issues of artificial intelligence. *AI & SOCIETY*, *37*(1), 23–37. https://doi.org/10.1007/s00146-021-01148-6

Oneto, L., & Chiappa, S. (2020). Fairness in machine learning. In *Studies in computational intelligence* (pp. 155–196). https://doi.org/10.1007/978-3-030-43883-8_7

Otterbacher, J., Checco, A., Demartini, G., & Clough, P. (2018). Investigating User Perception of Gender Bias in Image Search. SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. https://doi.org/10.1145/3209978.3210094

Pant, A., Hoda, R., Spiegler, S. V., Tantithamthavorn, C., & Turhan, B. (2023). Ethics in the Age of AI: An analysis of AI practitioners' awareness and challenges. *ACM Transactions on Software Engineering and Methodology*. https://doi.org/10.1145/3635715

Perez, C. C. (2020). *Invisible Women : Exposing data bias in a world designed for men*. https://lib.ugent.be/en/catalog/rug01:002787216

Razavian, N., Knoll, F., & Geras, K. J. (2020). Artificial intelligence explained for nonexperts. Seminars in Musculoskeletal Radiology, 24(01), 003–011. https://doi.org/10.1055/s-0039-3401041

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

# APPENDIX

## Appendix 1. Survey Questions

**Start of Block: Demographic Questions**

Dear Participant,

My name is Marloes Haitsma, and I am a student at ISEG, currently completing my Master's in Management. I am working on my dissertation which is focused on the perceptions of AI and gender bias. The aim of this survey is to gain insights from employees and managers currently working in the tech sector.

The survey should take approximately 5 minutes to complete. Please answer all questions honestly, as your input is invaluable to the success of this research. Rest assured that your responses in this survey are entirely anonymous and will be handled with strict confidentiality. The data collected will be used solely for the purpose of my research.

If you have any questions or concerns, please feel free to reach out via email at marloeshaitsma@aln.iseg.ulisboa.pt.

Thank you so much for your collaboration!

*By starting this survey you confirm that you have understood the purpose of this enquiry and that your participation is entirely voluntary, giving consent to the use of data collected in this survey. Furthermore, You are free to stop the survey at any time.*

Q1 The Tech Sector can be defined as the following: **"a sector of the economy that consists of businesses focusing on electronics, software, computers, social media, and other industries related to information technology."**

Would you say you work in this sector?

Yes (1)

No (2)

Don't know (3)

Q2 How would you define the industry you are currently working in?

Cyber Security

E-commerce, Social media and networking

Gaming and entertainment industry

Telecommunications

Hardware/Software

Internet and web services

Artificial Intelligence

Biotechnology and health tech

Not in the Tech Industry

Other

Q3 Are you currently in the position of manager or employee without managerial responsibilities?

Senior Manager

Middle manager (or equivalent)

Employee (without managerial responsibilities)

Q4 What is your gender?

Men

Women

Non-binary / third gender

Prefer not to say

Other

Q5 How old are you?

25 or below

26-35

36-45

46-55

Above 56

Q6 What is your country of residence?

________________________________________________________________

Q7 What is the highest degree or level of education you have completed?

Less than high school

High school graduate

Bachelor's degree

Master's degree

Ph.D or Higher

Prefer not to say

Other ________________________________________________

Q8 How many years of experience do you have in AI related work?

No experience (1)

Less than 1 year (2)

Between 1 to 2 years (3)

Between 3 to 5 years (4)

Between 6 to 10 years (5)

Over 10 years (6)

**Start of Block: Knowledge in AI**

Q9 How would you rate your overall knowledge of AI?

None at all (1)

Limited (2)

Average (3)

Above average (4)

Expert (5)

Q10 Have you taken any formal courses or training related to AI?

No

Yes

Q11 How often do you read about AI-related topics?

Never (1)

Rarely (2)

Occasionally (3)

Frequently (4)

Always (5)

Q12 How do you stay informed about advancements in AI?

News articles

Online courses

Academic journals

Workshops/seminars

Through work related activities

Social media

N/A
Other

Q13 How confident are you in explaining the basic concepts of AI to others?

Not confident at all (1)

Slightly confident (2)

Moderately confident (3)

Very confident (4)

Completely confident (5)

**Start of Block: Trust in AI**

Q14 How much do you trust the decisions made by AI systems?

None at all (1)

A little (2)

A moderate amount (3)

A lot (4)

A great deal (5)

Q15 In your opinion, how likely is it that AI is more fair in its decisions than humans?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

Q16 To what extent do you believe AI systems are biased in their decision-making processes?

Not biased at all (5)

Slightly biased (4)

Moderately biased (3)

Very biased (2)

Extremely biased (1)

Q17 To what extent do ethical considerations influence your trust in AI?

Not at all (5)

Slightly (4)

Moderately (3)

Very much (2)

Completely (1)

**Start of block: Gender Perception**

**Let us imagine that you and a friend have the same race, age, as well as practically identical educational and professional achievements but have a different gender. How likely are you to question the following outcomes of AI to happen?**

Q18 You are both applying for the same financial product (such as a credit card or home mortgage/loan) on the same bank app/website using your own devices. You notice the products that are offered to your friend charge higher interest rates than those offered to you.

How likely are you to question this outcome?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

Q19 You are both looking for similar jobs on the same employment app/website using your own devices. You notice the jobs that are offered to your friend usually have lower-paying salaries than those offered to you.

How likely are you to question this outcome?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

Q20 You both have the same nationality and are at the airport going through the same automated immigration kiosk that uses face recognition technology to verify travelers' identity. The automated immigration kiosk directs your friend to see an immigration officer while you are cleared to go through.

How likely are you to question this outcome?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

Q21 You are both booking a similar hotel room using the same hotel booking app/website using your own devices. Hotel rooms offered to your friend have higher prices than those offered to you.

How likely are you to question this outcome?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

**End of Block: Trust in AI**

**Start of Block: Perception of gender bias**

Q22 How familiar are you with the concept of gender bias in AI?

Not at all familiar (1)

Slightly familiar (2)

Somewhat familiar (3)

Moderately familiar (4)

Extremely familiar (5)

Q23 In your opinion, how likely is it that AI would discriminate towards one gender?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

Q24 To what extent do you believe that AI systems exhibit gender bias in their decision-making processes?

Not at all (1)

Slightly (2)

Moderately (3)

Very much (4)

Completely (5)

Q25 How concerned are you about the potential impact of gender bias in AI systems on society?

Not concerned at all (1)

Slightly concerned (2)

Moderately concerned (3)

Very concerned (4)

Extremely concerned (5)

Q26 Have you personally observed instances of gender bias in AI systems?

No (1)

Yes (2)

Q27 How much do you think societal stereotypes contribute to gender bias in AI?

Not at all (1)

Slightly (2)

Moderately (3)

Very much (4)

Completely (5)

Q28 How likely are you to advocate for measures to address and mitigate gender bias in AI?

Extremely unlikely (1)

Somewhat unlikely (2)

Neither likely nor unlikely (3)

Somewhat likely (4)

Extremely likely (5)

# Appendix 2. Output STATA

# Appendix 2.1 Descriptive

## 2.2.1 Summarize of demographics

**. summarize Position Gender Age**

| Variable | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| Position | 85 | 2.647059 | .6673666 | 1 | 3 |
| Gender | 85 | .5294118 | .5020964 | 0 | 1 |
| Age | 85 | 2.011765 | 1.313792 | 1 | 5 |

**. tab Position**

| Are you currently in the position of manager or employee without managerial resp | Freq. | Percent | Cum. |
|---|---|---|---|
| 1 | 9 | 10.59 | 10.59 |
| 2 | 12 | 14.12 | 24.71 |
| 3 | 64 | 75.29 | 100.00 |
| Total | 85 | 100.00 | |

**. tab Gender**

| What is your gender? - Selected Choice | Freq. | Percent | Cum. |
|---|---|---|---|
| 0 | 40 | 47.06 | 47.06 |
| 1 | 45 | 52.94 | 100.00 |
| Total | 85 | 100.00 | |

. tab CountryResidence

| What is your country of residence? | Freq. | Percent | Cum. |
|---:|---:|---:|---:|
| Belgium | 1 | 1.18 | 1.18 |
| Finland | 1 | 1.18 | 2.35 |
| Germany | 5 | 5.88 | 8.24 |
| Israel | 3 | 3.53 | 11.76 |
| Italy | 1 | 1.18 | 12.94 |
| Luxembourg | 1 | 1.18 | 14.12 |
| Portugal | 25 | 29.41 | 43.53 |
| The Netherlands | 46 | 54.12 | 97.65 |
| United Kingdom | 2 | 2.35 | 100.00 |
| Total | 85 | 100.00 | |

. tab Age

| How old are you? | Freq. | Percent | Cum. |
|---:|---:|---:|---:|
| 1 | 44 | 51.76 | 51.76 |
| 2 | 17 | 20.00 | 71.76 |
| 3 | 10 | 11.76 | 83.53 |
| 4 | 7 | 8.24 | 91.76 |
| 5 | 7 | 8.24 | 100.00 |
| Total | 85 | 100.00 | |

. tab Industry

| How would you define the industry you are currently working in? - Selected Choic | Freq. | Percent | Cum. |
|---:|---:|---:|---:|
| 4 | 24 | 28.24 | 28.24 |
| 5 | 13 | 15.29 | 43.53 |
| 7 | 2 | 2.35 | 45.88 |
| 8 | 3 | 3.53 | 49.41 |
| 9 | 7 | 8.24 | 57.65 |
| 11 | 4 | 4.71 | 62.35 |
| 12 | 11 | 12.94 | 75.29 |
| 13 | 3 | 3.53 | 78.82 |
| 14 | 17 | 20.00 | 98.82 |
| 15 | 1 | 1.18 | 100.00 |
| Total | 85 | 100.00 | |

### 2.2.2 Summarize of descriptive questions

| Variable | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| Q8Expierence | 85 | 2.470588 | 1.160749 | 1 | 6 |
| Q10Course | 85 | 1.470588 | .5020964 | 1 | 2 |
| Q9OverallK~e | 85 | 2.964706 | .822989 | 1 | 5 |
| Q11ReadAI | 85 | 3.070588 | .8699786 | 1 | 5 |
| Q12HowDoYo~d | 0 | | | | |
| Q13Confide~I | 85 | 2.764706 | 1.019419 | 1 | 5 |
| Q14Trust | 85 | 2.882353 | .6798418 | 1 | 5 |
| Q15AIbiased | 85 | 2.788235 | .9523669 | 1 | 5 |
| Q16Fairer | 85 | 3.094118 | .9079481 | 1 | 5 |
| Q17Ethical | 85 | 3 | 1.112697 | 1 | 5 |
| Q18Situati~1 | 85 | 4.129412 | 1.121099 | 1 | 5 |
| Q19Situati~2 | 85 | 3.941176 | 1.127079 | 1 | 5 |
| Q20Situati~3 | 84 | 3.333333 | 1.399943 | 1 | 5 |
| Q21Situati~4 | 85 | 3.952941 | 1.15373 | 1 | 5 |
| Q22Familia~s | 85 | 2.6 | 1.283596 | 1 | 5 |
| Q23AIDiscr~e | 85 | 3.247059 | 1.122347 | 1 | 5 |
| Q24AIsystems | 85 | 2.705882 | 1.067078 | 1 | 5 |
| Q25ImpactG~s | 85 | 3.011765 | 1.12857 | 1 | 5 |
| Q26Personal | 85 | 1.164706 | .3731162 | 1 | 2 |
| Q27Social | 85 | 3.564706 | 1.138577 | 1 | 5 |
| Q28Measures | 85 | 3.447059 | 1.159904 | 1 | 5 |

**. tabulate Gender Q9OverallKnowledge**

| What is your gender? - Selected Choice | How would you rate your overall knowledge of AI? None at a | Limited | Average | Above Ave | Expert | Total |
|---|---|---|---|---|---|---|
| Woman | 0 | 14 | 19 | 7 | 0 | 40 |
| Man | 1 | 12 | 15 | 16 | 1 | 45 |
| Total | 1 | 26 | 34 | 23 | 1 | 85 |

**. tabulate Q8Expierence Q9OverallKnowledge**

| How many years of experience do you have in AI related work? | How would you rate your overall knowledge of AI? None at a | Limited | Average | Above Ave | Expert | Total |
|---|---|---|---|---|---|---|
| No Expierence | 0 | 9 | 7 | 2 | 0 | 18 |
| Less than 1 year | 0 | 12 | 10 | 5 | 0 | 27 |
| 1 to 2 years | 1 | 4 | 14 | 10 | 0 | 29 |
| 3 to 5 years | 0 | 0 | 2 | 5 | 0 | 7 |
| 6 to 10 years | 0 | 0 | 0 | 1 | 0 | 1 |
| >10 years | 0 | 1 | 1 | 0 | 1 | 3 |
| Total | 1 | 26 | 34 | 23 | 1 | 85 |

# Appendix 2.2 Regression

## 2.2.1 Regression Conceptual model

**. reg AVERAGEGENDERPERCEPTION AVERAGEKNOWLEDGE AVERAGETRUST**

| Source | SS | df | MS | | | |
|---|---|---|---|---|---|---|
| | | | | Number of obs | = | 85 |
| | | | | F(2, 82) | = | 18.56 |
| Model | 14.7229572 | 2 | 7.36147859 | Prob > F | = | 0.0000 |
| Residual | 32.5292781 | 82 | .396698514 | R-squared | = | 0.3116 |
| | | | | Adj R-squared | = | 0.2948 |
| Total | 47.2522353 | 84 | .562526611 | Root MSE | = | .62984 |

| AVERAGEGENDERP~N | Coefficient | Std. err. | t | P>\|t\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| AVERAGEKNOWLEDGE | .3774109 | .0967125 | 3.90 | 0.000 | .185019 | .5698028 |
| AVERAGETRUST | -.5883446 | .113454 | -5.19 | 0.000 | -.8140407 | -.3626484 |
| _cons | 4.018255 | .4117751 | 9.76 | 0.000 | 3.199103 | 4.837407 |

## 2.2.2 Regression Gender

**. reg AVERAGEGENDERPERCEPTION AVERAGEKNOWLEDGE AVERAGETRUST Gender**

| Source | SS | df | MS | | | |
|---|---|---|---|---|---|---|
| | | | | Number of obs | = | 85 |
| | | | | F(3, 81) | = | 14.31 |
| Model | 16.3666545 | 3 | 5.45555151 | Prob > F | = | 0.0000 |
| Residual | 30.8855808 | 81 | .381303466 | R-squared | = | 0.3464 |
| | | | | Adj R-squared | = | 0.3222 |
| Total | 47.2522353 | 84 | .562526611 | Root MSE | = | .6175 |

| AVERAGEGENDERP~N | Coefficient | Std. err. | t | P>\|t\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| AVERAGEKNOWLEDGE | .4313244 | .0983087 | 4.39 | 0.000 | .2357208 | .6269279 |
| AVERAGETRUST | -.5441978 | .1132448 | -4.81 | 0.000 | -.7695196 | -.3188761 |
| Gender | -.2960337 | .1425823 | -2.08 | 0.041 | -.5797277 | -.0123397 |
| _cons | 3.886989 | .4086266 | 9.51 | 0.000 | 3.073951 | 4.700028 |

## 2.2.3 Slope Dummy

```
. reg  AVERAGEGENDERPERCEPTION AVERAGEKNOWLEDGE AVERAGETRUST GENDERTRUST

      Source |       SS           df       MS      Number of obs   =        85
-------------+----------------------------------   F(3, 81)        =     15.11
       Model |  16.9543378         3  5.65144592   Prob > F        =    0.0000
    Residual |  30.2978975        81  .374048118   R-squared       =    0.3588
-------------+----------------------------------   Adj R-squared   =    0.3351
       Total |  47.2522353        84  .562526611   Root MSE        =    .61159

------------------------------------------------------------------------------
AVERAGEGENDERP~N | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-----------------+------------------------------------------------------------
AVERAGEKNOWLEDGE |   .5021923   .1069081     4.70   0.000     .2894788    .7149059
    AVERAGETRUST |  -.5280272   .1129014    -4.68   0.000    -.7526656   -.3033887
     GENDERTRUST |   -.116331   .0476291    -2.44   0.017    -.211098    -.021564
           _cons |   3.667342   .4248758     8.63   0.000     2.821972    4.512712
------------------------------------------------------------------------------
```

```
. reg  AVERAGEGENDERPERCEPTION AVERAGEKNOWLEDGE AVERAGETRUST GENDERKNOWLEDGE

      Source |       SS           df       MS      Number of obs   =        85
-------------+----------------------------------   F(3, 81)        =     15.11
       Model |  16.9543378         3  5.65144592   Prob > F        =    0.0000
    Residual |  30.2978975        81  .374048118   R-squared       =    0.3588
-------------+----------------------------------   Adj R-squared   =    0.3351
       Total |  47.2522353        84  .562526611   Root MSE        =    .61159

------------------------------------------------------------------------------
AVERAGEGENDERP~N | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-----------------+------------------------------------------------------------
AVERAGEKNOWLEDGE |   .5021923   .1069081     4.70   0.000     .2894788    .7149059
    AVERAGETRUST |  -.5280272   .1129014    -4.68   0.000    -.7526656   -.3033887
 GENDERKNOWLEDGE |   -.116331   .0476291    -2.44   0.017    -.211098    -.021564
           _cons |   3.667342   .4248758     8.63   0.000     2.821972    4.512712
------------------------------------------------------------------------------
```