International Journal of Information Management Data Insights 3 (2023) 100165

Contents lists available at ScienceDirect

International Journal of Information Management Data Insights

journal homepage: www.elsevier.com/locate/jjimei

How can we manage biases in artificial intelligence systems – A systematic literature review

Dr. Varsha P.S.

School of Commerce, Presidency University, Bangalore, India

ARTICLE INFO

Keywords:
Artificial intelligence
Bias
Vulnerabilities
Responsible Ai
AI ethics
AI systems

ABSTRACT

Artificial intelligence is similar to human intelligence, and robots in organisations always perform human tasks. However, AI encounters a variety of biases during its operational process in the online economy. The coded algorithms helps in decision-making in firms with a variety of biases and ambiguity. The study is qualitative in nature and asserts that AI biases and vulnerabilities experienced by people across industries lead to gender biases and racial discrimination. Furthermore, the study describes the different types of biases and emphasises the importance of responsible AI in firms in order to reduce the risk from AI. The implications discuss how policymakers, managers, and employees must understand biases to improve corporate fairness and societal well-being. Future research can be carryout on consumer bias, bias in job automation and bias in societal data.

1. Introduction

Artificial intelligence (AI) is becoming a much more popular and common feature in businesses of several operational processes such as customer service, marketing and sales (Brit, 2021; Verma et al., 2021). AI implementation in business and commerce has increased in the current scenario to predict better consumer choices/customization and achieve companies' competitive advantage (Teleaba et al., 2021; Waja et al., 2023). Furthermore, technology adoption in the firms is expected to enhance business growth and make decisions (Gonzales & Hargreaves, 2022; Brit, 2021). However, the debate on human cognitive bias has heated up as a result of the use of AI to forecast company sales or results (Teleaba et al., 2021). AI is not stable and data input errors may occur as a result of biased output (Huang & Rust, 2021). By considering the AI biases in various industries such as the banking industry, it was mentioned that loan decisions were biased even though no bigotry was programmed into the system (Ukanwa & Rust, 2020). Researchers discovered that gender bias can occur in their results when using the unbiased algorithm in career stem advertisement (Lambrecht & Tucker, 2019). As a result, the use of AI in decision making in many firms has raised concerns about automated choices leads to discriminatory outcomes (Sweeney, 2013) and undesirable ads (Datta et al., 2015). Hence, technological flaws are more common than human flaws in the firms (Brit, 2021).

The AI usage in sales to increase revenues has gained worldwide attention. Automation comprises of machine learning, deep learning, information retrieval and natural language processing technologies helps in business process and leverage data to bring innovative solutions, more customization, profit optimization and embrace firm transformation (Dickie, 2021; Nagwani & Suri, 2023). Apparently, bias occurs in AI tools in sales while business leads are generated to connect with customers and collect various data but fail to understand the highest lifetime value (Fatemi, 2020). Subsequently, transparency is one of the critical factors in AI deep learning systems (Sharma et al., 2021). Data transparency assorts privacy while humans train the machine learning models through an artificial neural network where decisions are not justified properly, which encounters black box problems (Selbst & Barocas, 2018). This black box problems arises risk to the firms while deploying algorithms is not easily explainable and develops biases in the AI system (Roselli et al., 2019).

Several discussion on risk of AI biases were observed like from court decisions to medicines to business (Teleaba et al., 2021). Considering the cases of Apple – gender bias (BBC, 2019) and COMPAS – African American defendant bias (Dressel & Farid, 2018), the number of biased AI systems and algorithms is expected to increase in the next five years (IBM, 2018), exploiting people were more vulnerable. Following that, people became aware of the issue of biases in order to bring fairness and equity to machine learning in certain fields such as healthcare, business and management. Furthermore, the risk of incorrect projection will have a negative impact on consumers with products or services that do not create the values. As a result of the decrease in customer satisfaction and loyalty, these biased outcomes will have a less influence on the firm's equity, revenues, and profitability (Teleaba et al., 2021). Hence, biases exist in embedded computer code even when algorithms fail to provide decisions (Edionwe, 2017). Then the data scientists and software engineers fail to understand the process and choices of larger societal scenar-

E-mail address: smiracle2@gmail.com

https://doi.org/10.1016/j.jjimei.2023.100165
Received 29 June 2022; Received in revised form 12 February 2023; Accepted 13 February 2023

Table 1
Definitions of AI.

Description	Refs.
The ability to reason, solve problems, learn, and integrate multiple human skills like perception, cognition, memory, language, or planning refers AI intelligence.	Kar et al. (2022)
AI systems use mathematical models to derive inferences from data, increases transparency and humans get answers to the questions like 'what', 'how' and 'why' to bring the benefits to the business.	Kar et al. (2022)
AI has evolved in the firms from being a just adopted technology to powering routine decision-making processes in all the domains.	Kar et al. (2022); Morande (2022)
AI techniques able to increase the knowledge of employees in the firms by allowing them to comprehend and conquer complex situations more effectively and facilitates the decision-making process by offering several alternative choices.	Malik et al. (2021)
The proficiency of a machine leveraged by AI to full fill the customer expectations and increases the operational efficiency in the organization.	Kushwaha et al. (2021)
The term machine learning is an artificial or computational intelligence technique describes a machine's capacity to learn and carry out a process given an objective and specific training tasks to accomplish the goal.	Votto et al. (2021); Garg et al. (2021)
AI is defined as systems which mimic cognitive abilities commonly associated with human characteristics such as learning, speech, and problem solving.	Dwivedi et al. (2021)

ios, bias can be introduced into firms (Akter et al., 2022; Yarger et al., 2019).

The AI driven decision making brings unfair and unequal effects in the firms that leads to algorithmic bias and there will be a paucity of studies on this topic (Kar & Dwivedi, 2020; Kumar et al., 2021; Vimalkumar et al., 2021). Such negative experience from AI bias has a great impact on firms, specifically when decisions are involved. The study also shed light to raises the automation bias on racial, gender, credit scores, face recognition, etc., which highlights the issues through virtual assistants, robotics and algorithmic recommendations in the firms. Besides that the consumer users, researchers and experts has to give a critical thinking during incorporating ai-based solutions in the firm systems. However, the outcomes obtained by AI in various sectors still have errors and our study proposed the following research questions solve this gap and more significant in addressing AI biases

- What are AI biases and how do they occur in system?
- How should AI biases and vulnerabilities be addressed in the system?

These research questions brings the novelty of the article to address the biases and its types to understand and mitigate the risks in the firms. Also the study aims to focus on addressing the biases so that vulnerabilities can be evaluated. We structured our discussions further to address these exploratory questions as follows. First, the literature review section, second research methodology section, third explains about the findings from cases and fourth will be to address AI bias in systems. Lastly we explained on discussions, implications, future research followed by conclusion.

2. Literature review

2.1. Evolution of AI systems

AI can be traced back to 1950, when Alan Turing, an English polymath, devised a test to see if a machine could mimic human cognitive functions to identify patterns (Batra et al., 2018). Then it became more popular in 1956 when John McCarthy invited academicians and industry experts in interdisciplinary fields across the globe to discuss the importance of computers that consume data and mimic human behaviour (Garg et al., 2021;Herath & Mittal, 2022). This data shares and creates the possibilities that made available by the new advancement of higher computing processing power across the globe (Akter et al., 2020). These technological advancement affects the companies and completely transform the market place (Sharma et al., 2021). Since the invention of high computing power in 1956, relevant of AI theories have been developed for several years (Cohen & Feigenbaum, 2014). Many academicians and industry experts have proposed various definitions of AI as machines with human-like cognitive abilities (McGettigan, 2017) and enable them to handle complex situations (Malik et al., 2021). Furthermore, AI helps to provide several solutions for decision making in firms (Bader & Kaiser, 2019) to evaluate analytical, intuitive, and empathetic intelligence (Kar et al., 2022). AI, which includes data, algorithms, and computing has made significant progress in recent years (Messner, 2022). As a result, AI is machine learning deploys an algorithm to feed raw data able to produce meaningful outputs via models (Sarle, 1994). It is a group of computing technologies that allow us to make rational decisions in complex situations in various contexts (Treinnick, 2017). Lastly we summarize the definition of AI with important phrases in Table 1.

AI usage in current businesses scenarios is neither normal nor neutral and arises the various challenges in several domains (Kar et al., 2022). By considering studies on optimal control decisions on the environment (Qi et al., 2019), an approach to predict crashes (Abdel-aty & Haleem, 2011), timely identification of traffic conditions (Hossain & Muromachi, 2012), large-scale entrepreneurship (Elia et al., 2020), malware detection (Mohaisen et al., 2015) etc., uses AI driven decisions. Thus AI in the system impacts on business processes and its performance mainly in the areas technology, acceptance, social integration, job opportunities and regulations (Cao et al., 2021; Collins et al., 2021; Kumar et al., 2021). AI able to recognise patterns in customer data in marketing domain. Brinks Home Smart Security System company uses AI to provide better services with the right content to customers by recognising patterns in customer data using natural language processing (NLP). And also in Adobe Sensei firm deploys AI with machine learning has the potential to assist marketers in taming data for meaningful insights (Liesse Julie, 2021). Further AI used in human resources in firms for recruitment to predict the job description and select the right candidate (Sridevi & Suganthi, 2022; Votto et al., 2021). Subsequently, algorithms are used in healthcare, transportation, and security where decisions can be made for human life or health requires transparency and explainability (Adadi & Berrada, 2018; Chintalapudi et al., 2021; Pawar et al., 2020). Hence explainability AI is the most recent and relevant topic in Industry 4.0 which transforms operations(process) by using AI systems for decision making or predictions (Singh et al., 2022).

2.2. Biases in AI systems

In the system, the data input is biased, the output is likely to be biased (Huang & Rust, 2021). For instance, Amazon uses the AI tool to measure and rate job applicants while discriminating against female applicants (Weissman, 2018). Furthermore, AI errors occur in insurance companies when auto-calculation premiums are based on religion rather than gender (Villasenor, 2019). As a result, the automated systems have biases in dynamic pricing and targeted discounts (Miler & Hosanagar, 2020; Brit, 2021).Hence, bias can creep into algorithms and only beneficial for training data sets in the systems (Brit, 2020).

Several studies indicate that human bias results from the response of technology advancement (de Graaf & Allouch, 2017; Haring et al., 2018; Kuchenbrandt et al., 2013). While cognitive biases influence all aspects of how people make decisions through AI and robotic creations

Table 2
AI bias in various industries.

Firm	Algorithm bias	Observable/Unobservable
Adobe	The adobe software blocked the customers of a specific demographics while purchasing the software (Jared Council, 2021)	Unobservable
Lyft	The bias exists in dynamic ridesharing prices recommends higher surge prices have to be paid by the customers (Wiggers, 2020)	Observable
Facebook(FB)	The AI bias occurs in FB that allows the advertisers to target the marketing ads/job ads to the specific gender race and religious with the minority backgrounds (Dilmegani, 2022)	Observable
Banker	The error that occurs in fintech starts up in the payment process and revenue sharing partnerships (Annie Brown, 2021)	Observable
Microsoft	When customers started chatting with the Tay chatbot regarding racism comments and voice assistant started to retype the phrases rather than address the customer queries (Brit, 2021)	Observable
Nikon	The data-driven bias arises from the new Nikon product about Asian faces and HP media smart computers has skin tone problems in their face recognition (Hammond, 2016)	Observable
Pulse oximeter	It's one of the important new products (device) in clinical management to monitor oxygen levels during pandemics the bias encountered that less accuracy in darker skin than lighter skin (racial bias) (The Conversation, 2022)	Observable

(Letheren et al., 2020). As a result of the proliferation of humanised technology amongst consumers and marketers, it is crucial to comprehend the accidental transfer of human biases into the design of artificial intelligence (Jobin et al., 2019). In this context, AI biases can be explained that information can be passed from human to AI while programming and coding the data process develops the racism and discrimination issue (Penny, 2017). Hence, bias is an anomaly from machine learning algorithms caused by preconception assumptions made during the algorithm development phase or determined training data sets (Dilmegani, 2022). Thus, all of these pitfalls suggest that the problems caused by algorithm bias are not trivial unless marketers and consumers are educated (Knight, 2017). It is concluded as a beautiful mess is evolved and influences the firms when programming bias in an AI or robots. Subsequently, the study discovered that bias in psychology and behavioural economics to describe the AI risk in predicting consumer choice were classified into two types - observable (in an e-commerce website, marketers like to know their customers' biases by analysing large customer data sets from several years, which includes the most frequently purchased products, product features, and product availability) and unobservable (impossible to detect big data related to purchase or pricing data that requires additional research skills to identify bias) (Teleaba et al., 2021). AI biases in various industries discusses both unobservable and observable risks in Table 2.

The impact of cognitive bias and improper datasets create bias that affects firms in various parameters like dynamic pricing, e-business, hiring, and healthcare (Dilmegani, 2022). Further the research provided several significant contributions and we discovered and related three theories from literature includes social theory (Joyce et al., 2021; Zajko, 2022), stimulus-organism-response theory(Mehrabian & Russell, 1974) and organisational justice theory (Colquitt & Rodell, 2015). In align with these theories, the experts from global tech companies were implemented various measures to reduce AI bias (Weyerer & Langer, 2020). For example, Google has implemented a Testing with Concept Activation Vectors (TCAV) programme in which developers test decision-making algorithms to reduce bias and gender discrimination (Google, 2019). Accenture introduced 'Teach and Test' AI testing services to assist businesses to minimise biases and discriminatory content (Accenture, 2018). IBM's AI Fairness 360 toolkit is a holistic and comprehensive approach that includes 70 parameter fairness metrics to reduce biases in AI systems (Bellamy et al., 2019). In healthcare systems to decrease the bias by creating fairness standards, regulating algorithms, tools for clinical decision making and fostering relationship between public and business (Panch et al, 2019).

3. Methodology

Systematic review is widely used in the multidisciplinary domains, but currently we are advancing in business, management, and accounting to examine the vast amount of data scattered throughout the internet in order to provide quantifiable, reproducible, systematic way to articulate and thorough specific domains (Weed, 2006). To conduct the literature review, we followed the guidelines provided by seminal review studies (Cranfield et al., 2003; Durach et al., 2017). We identified the various sources of AI bias in firms in this study. To carry out the literature review, we used the Scopus database to look for publications for this systematic review, since it gives a broader range of scholarly information to gain a deeper understanding of the research we intend to conduct (Kar et al., 2022). The inclusion of Scopus-indexed research papers in the database was contingent on stringent selection criteria, so we may rely on them for academic research (Kumar et al., 2022; Tiwary et al., 2021). To conduct our research, we employed a list of keywords in combination with a database search of article titles and keywords.

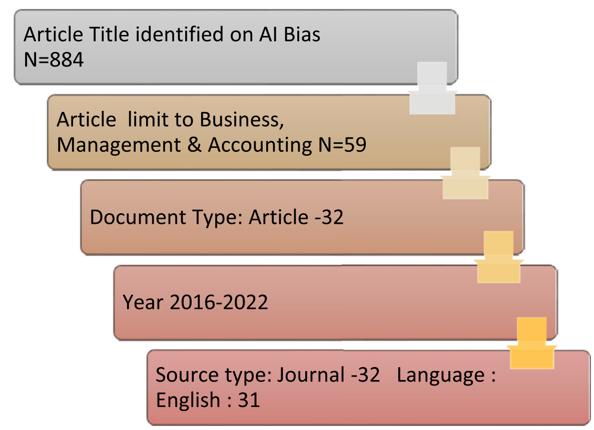

Fig. 1. Flow chart of relevant papers(inclusion & exclusion).

To extract the literature we used the key words search ALL (("AI Bias*" OR "artificial intelligence bias" OR "Algorithm bias*") AND ("Bias*" OR "Risk*")) by using Boolean logic (AND/ OR) and found 884 documents. In Fig. 1 represents the flowchart for relevant papers inclusion and exclusion criteria.

Further developed a conceptual model Fig. 2 to address the research question.

4. Findings from cases

4.1. AI bias in E-commerce

Amazon is a US based global online retailer deploys AI to improve work efficiency and customer personalization of its services and products. The company's workforce capacity is 60 percent male whereas 74 percent of the firm's managerial positions were identified as gender distorted while using an AI recruiting tool (Hamilton, 2018).Then the experts used the data to create the algorithm through programme to search for resumes from the previous ten years and algorithms were looking only for white males. Furthermore, these algorithms are trained

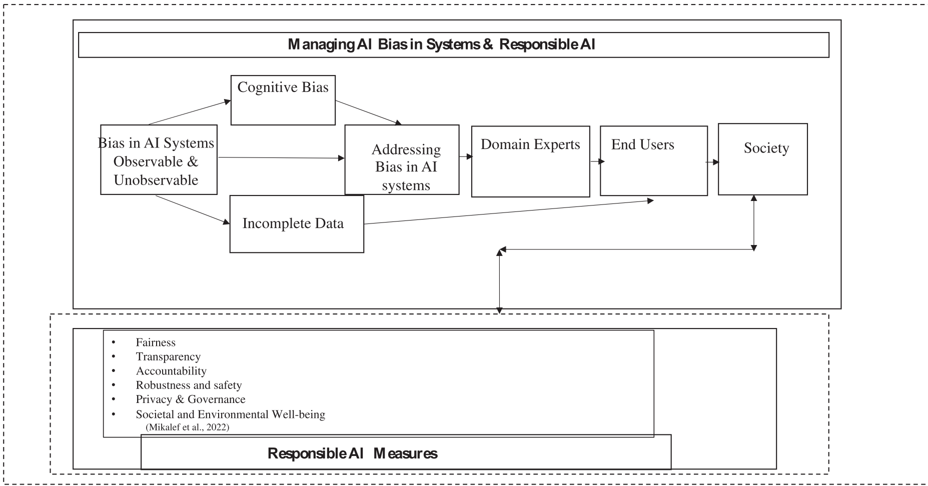

Fig. 2. Conceptual Model of Managing AI Bias in systems & Responsible AI.

to recognise word patterns in resumes rather than specific skill sets and company that developed this AI tool has a best practises while ignore the word woman (James Vincent, 2018). Thus, gender bias exists during recruitment indicates that gender inequality in the workplace during recruitment process in the firms (Lindsay, 2019).

4.2. AI bias in online ads

This case explains about the ads and compared for white names on websites. The Federal Trade Commission (FTC) report discovered that online search inquiries for African-American names were mentioned in advertisements from various services develops the arrest records for black people (Sweeney & Zang, 2013). A similar incident occurred in the micro-targeting of high interest credit cards and other financial products in the website's advertisements, which constantly recommended to black customers to purchase the high interest credit card offerings, which exploits the innocent these customers and they lose trust in brands (FTC Hearing, 2018). This study revealed that big data analytics is used incorrectly to track online users based on their profiles, digital activities and behaviour (Ramirez et al., 2016). As a matter of fact, the FTC discovered that online users are denied access to their credit cards while browsing the internet. Additionally, predictive analytics used to compile web browsing history suggested incorrect way and defines individuals data for specific jobs, personal credit or educational opportunities. Besides that, marketers will use online proxies that include zip codes to gather information about individuals' socioeconomic status based on their neighbourhood results in inaccurate assumptions about individual lifestyles or preferences (Noyes, 2015).Thus, mis appropriate big data analysis develops disparities for genuine people based on race, gender, age, skills, religion and sexual orientation. As an outcome, algorithms are implemented for vulnerable populations, replicating and causing explicit discrimination or creating a new type of error, consciously or unconsciously to develop societal bias, fostering stereotypes, and unfair profiling of online users.

4.3. Customer's discriminations in sharing economy

Airbnb is a global online marketplace and sharing economy that caters to lodging, homestays, and tourism businesses. In 2017, Homesharing companies discovered that a small number of hosts were rejecting renters (customers) based on race, age, gender, and other factors (Murphy, 2016). Customers were rejected primarily because of their online public profiles on social media websites (Murphy, 2016). The case concludes that AI underestimates or judges' customers, resulting in incorrect predictions on profits and an impact on people through racial discrimination.

4.4. Bias in digital advertising

In 2014, Google facilitated users with four gender options: "male", "female", "rather not to say" and "custom" (Bennet, 2014). Male and female options are more conventional gender demographics. The choice of "rather not to say" is mentioned in the group of users who do not wish to reveal their identity, and "custom" is meant for the nonconventional gender groups. The policies were made and provided to the customers or users on a larger scale. As a result, each customer will have two segments to collect information in Google. Then customers or users register their Google account, and Google then collects the customer's data from web searches, advertisements, geolocation, history, payment transactions, etc. (Shekhawat et al., 2019). Then, ad customization page of Google Ad Settings facilitates customers that the chance to choose how they want to be grouped. Google provides customers with options that are not ideal or relevant for categorising the male or female orientation in order to create an individual preference on websites in systems that encounter bias, privacy and data transparency concerns in Google Ad personalised page (Noble, 2018). Further, Google encountered errors through Google Images with cultural bias, where the search for CEO photos of white people is more dominant (Kay et al., 2015). As a result, big data will be a major concern in AI bias, which impacts customer preferences and causes cultural bias in the online environment (Schroeder, 2021).

4.5. Bias in hiring

The organisations introduces AI in human resource management and considered as novel business practices (Votto et al., 2021). However, there is a under-representation in the workplace has been investigated using demographic factors such as race and gender (Dovidio & Gartner, 1996). Woodruff et al. (2018) reported in their study that half of the respondents saw that the hiring process as a problem with black respondents and they have a less rate of hiring possibilities which is linked to national concerns about racial justice and economic inequality. Furthermore, in the IT sector, technology workers face screening bias based on physical appearance, gender and ethnicity during the recruitment process (Beattie & Johnson, 2012). As a result of racial and gender bias, women will be treated worse than male associates in the technology job market (Wachter-Boettcher, 2017).

4.6. Bias in facial recognition

Joy Buolamwini, MIT scholar discovered in her study that algorithms power facial recognition software devices failed to identify dark-skinned complexions (Hardesty, 2018). In facial recognition software has a capability to train the data sets and evaluates higher than 75 percent male and more than 80 percent white. When the individual in the photo was a white man the software was precise and exact 99 percent of the time while recognizing the human as male. According to the study that the product error exists which were lower than a percent as a whole population, however, it increased to greater than 20 percent in one product and 34 percent in the other two recognized as darker-skinned women as a female (Lee et al., 2019).

5. Addressing bias in AI systems

The fact that marketers deploys of AI to capture emotional data for analysis in order to understand the customer's emotions. Customers interacting with AI results in high level of customer disengagement on social media platforms such as Facebook (likes and dislikes) (Srinivasan et al., 2016). As a result, customers are unable to comprehend or interact with AI (Luo et al., 2019). Customers will interact with voice assistants to solve their problems in the future, which will introduce biases. As a result, various measures have been considered to avoid these biases. While AI is being used in marketing analysis, particularly targeting and customised marketing actions, marketers must be aware of AI biases and improve their competency in order to minimise AI biases. These AI biases are unpredictable (Fuchs, 2018) and increases societal vulnerabilities has a major concern will dicuss on three levels—domain experts, end-users and society (Lockey et al., 2021).

5.1. Domain expert vulnerabilities

The experts from AI domain in the firms deploy the AI system for operational processes. For an instance, in healthcare doctors use AI-enabled medical diagnosis applications. This domain expert knowledge can be incorporated into the development of codified information to train AI systems, and they work with the outputs for service delivery (Lockey et al., 2021). The main vulnerabilities faced by the domain experts specifically professional knowledge, skills, identity and reputation and automation leads to deskilling (Rinta-Kahila et al., 2018; Sutton et al., 2018). An additional vulnerability in healthcare is where domain experts understand the problems to make clinical decisions, and anthropomorphism may threaten professional identity and reputation (Lockey et al., 2021). Similarly, the firms recruitment process deploys the AI to hire people. The vulnerability will be managers' knowledge in using AI tools to select candidates without any gender or racial bias in the software (Black & van Esch, 2021). However, algorithmic bias has increased as a public scrutiny of AI-powered HR solutions (Drage & Mackereth, 2022). Amazon reported in 2018 that they were abandoning the development of an AI-powered recruitment engine because it identified gender proxies on candidates' CVs and discriminated against female applicants (Dastin, 2018). An impartial assessment of Facebook's job advertisement algorithm in 2021 revealed that it provided different advertisements to male and female users based on the gender distribution of women and men in certain fields (Hao, 2021). Further, an online photo-editing app, FaceApp, was later discovered that racial bias to be lightening the darker skin tones of African-Americans because European faces ruled the training data, thereby defining the algorithm's standard of beauty (Morse, 2017). Subsequently, Airbnb implicit the racial in African-Americans with unusual names are less likely to receive a successful booking than visitors with more common names while discriminating the African people (Edelman et al., 2017). Lastly, racial bias has been found in many areas of financial services, such as mortgage lending, other personal lending, and business lending, and credit scores in the insurance industry, where white households can claim a higher amount of insurance than black homemakers (Casualty Actuarial Society (2022)).

To address this vulnerabilities includes racial biases and discrimination caused by machine learning and AI where domain expert must embrace technological change. Further domain experts or managers and data scientists come up with the agreement to bring the principles and values in the operational process as an ‘ AI bias-free’ zone with continuous innovation in the firms. Lastlty, the managers from various firms must understand responsible AI to protect the customer's data usage for a specific purpose.

5.2. End users

End users are directly impacted by the AI systems. There will be several vulnerable problems, inaccuracies, or biases in the system. Many end-users face susceptible problems while understating AI-based decisions (Lockey et al., 2021). Considering an example of AI in personal insurance companies accumulates the thousands of data points to judge the bias when someone claims automobile insurance (Lockey et al., 2021). By understanding this context, customers loses the data privacy and vulnerability arises the loss of human dignity (Lockey et al., 2021). Subsequently, in the healthcare domain there is a mismatch between the data or environment in the system is trained from machine learning occurs bias on the patients' health records data (Challen et al., 2019). Apparently AI has a capability to understand the consumers and voice assistants able to describe the consumer relationship through customers voice results the privacy concerns (Cheng et al., 2022; Grewal et al., 2021). Additionally, payments through facial recognition will have major privacy issues like human face has a individual information on appearance, age, gender, etc. (Dantcheva & Brémond, 2016; Dibeklioğlu et al., 2015; Liu et al., 2021). Furthermore chatbots creates a major problems that unable to understand and the customers leads to distrust sellers or buyers (Yen & Chiang, 2021). Also by leveraging AI in TripAdvisor - as sharing economy company has a negative implications in terms of privacy and security concerns of customers data and social interactions between the customers and virtual assistants in online creates less customer satisfaction (Grunder & Neuhofer, 2021). Hence AI based chatbots provides the extensible solutions to the end-users (customers) but there will be a major risks and challenges occurs in human and machine interaction, automatic detection and biases (Kushwaha et al., 2021).

In the present world AI is a driving force for all the fields to attain sustainability. To reduce these biases while interacting with the end-users, the marketers must aware of the five core functions includes: recognising the ethical concerns with AI such as fairness, transparency, parity, kindness, and benefits for society; increasing human awareness of AI by helping people comprehend how individual products' AI systems function and how businesses create their algorithms; working together with AI through dialogue, listening, and comprehension between humans and AI; ensuring the accountability of AI—confirming that both the creators and users of AI systems adhere to ethical standards; AI system integrity—keeping it constrained to the purposes for which the technology was designed to decrease bias (Schrader & Ghosh, 2018).

5.3. Society

Societal vulnerabilities include knowledge asymmetry, power centralization, and the ability of AI to cascade failures (Lockey et al., 2021). Knowledge asymmetry means by consider an example between two IT companies, policymakers, and citizens is constantly changing (Nemitz, 2018). Digital disruption is driving factor of AI development for data extraction for various operational decisions (Nemitz, 2018). Moreover, this data will be inaccurate and biased and privacy concerns in AI systems will have a negative impact on citizens' encounters with errors and inequality as well as undermine human rights such as the right to privacy (Lockey et al., 2021). Perhaps we are still in the early stages of understanding the AI biases and vulnerabilities caused by the technologies in today's world.

To address these issues, industry experts, policymakers, and academics can anticipate how to develop and use AI (Barredo Arrieta et al.,

Table 3
Responsible AI factors.

Factors	Explanation	Measures are taken to address AI Bias
Fairness	AI devices validate the diversity inclusion and reduce the biases	IBM has taken the initiative to minimize bias by providing the open-source toolkit to evaluate, create a report and alleviate discrimination and bias through machine learning models (IBM AI Fairness, 360)
Transparency	AI systems must be more transparent with respect to processes and results in the firms	Accenture practices the responsible AI has a transparency in four pillars organizational, operational, technical and reputational to create company values and ethics. By understanding the sources of bias and developing the Accenture Algorithmic toolkit used to investigate errors and bring the fairness and transparency decisions by creating a new model (Accenture, 2021).
Accountability	AI systems must bring the accountability of their results with ethics	In Microsoft, people are accountable for the AI system's impact on the world due to a variety of models, data sets and new technology disruption. The company follows principles and guidelines to understand the customers through facial recognition and understanding and monitoring the errors in each stage to minimize the bias in the AI life cycle (Microsoft, 2022)
Robustness & Safety	AI systems should be created with precautionary measures taken to reduce the errors	Google is in adversarial learning while using the neural network by creating adversarial illustrations to fool a system in the network to detect the frauds (Google)
Privacy & Governance	Personal data is used to make decisions and privacy controls must be created to support technology ensure that personal data will be used for specific and fair purposes. AI Systems must follow the regulations related to international data privacy laws and standards	Cisco developed a privacy engineering practices into the Cisco Secure Development Lifecycle (CSDL). These practices help to assures that data privacy in the service offerings. Further, the company sets and follows the principles of the Global Personal Data Protection and Privacy Policy (Cisco, 2022).
Societal & Environmental Wellbeing	AI Systems bring the ethical and equitable AI as a comprehensive approach to society and environment well being	Intel is designed the AI lifecycle to reduce the risks by bringing the ethical principles and to maximize the benefits the society by using the right tools and enabling an inclusive and sustainable environment (Intel)

Source: (Mikalef et al., 2022).

2020). As a result, AI raises new ethical, legal, and governance issues, such as racial discrimination, gender bias, and issues related to customer awareness and knowledge of AI entailed in decision outcomes (Singapore Government, 2021). Previous research discussed about distant factors associated with responsible principles such as bias removal (Brighton & Gigerenzer, 2015), explain ability of AI results (Gunning et al., 2019), and safety and security (Srivastava et al., 2017). In recent years, we have learned more about responsible AI in businesses (Dignum, 2019). As a result, the growing awareness of responsible AI addresses biases. Fairness, transparency, accountability, robustness, safety, privacy, governance, and societal and environmental well-being are key areas for increasing responsible AI (Mikalef et al., 2022). By mentioning the numerous causes of bias in the system, we cannot think that a single approach to mitigate all of them (Roselli et al., 2019). The study posit to provide a combination of quantitative assessments, business processes, monitoring, data review, evaluations, and experimental studies to minimise the AI bias (Roselli et al., 2019).Thus, the study implemented responsible AI (Mikalef et al., 2022) factors and mentioned in Table 3.

6. Discussion

6.1. Implications for literature

This paper outlines the several AI biases in the firms. In addition, we developed a framework to explain the AI biases in detail. In the AI literature, a stream of studies highlights the importance of AI for decision-making in firms (Akter et al., 2020; Garg et al., 2021; Kar et al., 2022; Sharma et al., 2021). Due to the rapid growth of technology, it impacts both firms and markets (Sharma et al., 2021). Thus, the AI has the cognitive ability to perform the work and mimic like a human (Dwivedi et al., 2021). Therefore the various studies proved that the AI usage will have major challenges in various domain Abdel-aty and Haleem (2011); Elia et al. (2020); Hossain and Muromachi (2012); Kar et al. (2022); Qi et al. (2019) . Also, the study discusses the types of biases, including cognitive bias and incomplete data, by explaining various instances (Panch et al., 2019; Shrestha et al., 2019; Weyerer & Langer, 2020. However, this gamut of literature on AI bias is limited, and our study focuses on the novel findings to contribute to methodological advancement and the implications of vulnerabilities in firms to measure bias and optimise the risk by establishing policies. Further research brings out the importance of AI community engagement. Managers and employees must educate themselves and invest the time to understand the bias and to find solution for this.

6.2. Implications for managerial and business practice

Our findings also help in numerous ways to practice. The study analysed from literature-based conclusions were verified and suggests that experts, managers and scientists have two opportunities to identify and reduce biases in organizations.The first is the possibility of employing AI to identify and mitigate the impact of human biases. The second potential is to improve AI systems how they exploit data used to build and deployed the models so that they do not perpetuate human and society prejudices or create bias and related issues of their own. Consequently, collaboration across domains helps to develop and implement technical innovations, operational methods and ethical standards is necessary procedures to be taken to reduce bias in businesses. Also, practitioners and business policy leaders in the firms can minimise or reduce the AI risk by considering the following suggestions: first, understand the situations where bias can be corrected by AI as well as those in which there is a significant chance that bias could be made worse by AI in the firms; second, establish policies and techniques to detect and mitigate bias in artificial intelligence systems; and third, engage in fact-based discussions with respect to human decision-making biases (Silberg & Manyika, 2019). Further, our research highlights on operational procedures may involve to enhance data through more sampling, employing internal teams or outside entities to audit data and models and engaging proactively with the groups. Lastly, clarity regarding methods and metrics in the systems that enables to comprehend the measures in fairness in business forcasting. Subsequently, the managers and employees invest more time in research on bias as a multidisciplinary approach to work on ethical issues and data privacy concerns. Managers' motivates the employees in the firms to do more research for advancement will require interdisciplinary participation (de Almeida et al., 2021). By fostering AI education and access to tools and opportunities, managers and employees in firms should devote more time to the development of the AI community in order to eliminate unfair biases and increase people's well-being.

6.3. Implications for theory

A major theoretical contribution of this reserach is that understanding the concept of AI bias by providing in the current IS literature systematically identified, explained and synthesised the theoretical relations that have been conceptually or empirically investigated in previous studies. The social theory which was analysed as bias in data science able to develop social or gender or racial bias in firms (Joyce et al., 2021; Zajko, 2022). Social theory emphasises algorithmic bias based on class and economic inequality (Grusky, 2019; Costanza-Chock, 2020), gender disparity (Costanza-Chock, 2020), and racism (Wong, 2020). The stimulus-organism-response theory (Mehrabian & Russell, 1974) explains about biases that develops in algorithm outputs will impact the consumer behaviour through perceived fairness. This theory posits that external stimuli influence the internal(psychological) stimuli of individuals, results in behavioural reactions (Mehrabian & Russell, 1974) based on algorithmic bias. According to the organisational justice theory, justice from the employee's perspective shows that firms or top management is seen to operate consistently, equitably, respectfully, and transparently in decision contexts through fairness (Colquitt & Rodell, 2015). This study followed a more theoretical approach, so the research results will be useful to behavioural and organisational researchers.

6.4. Future research work

AI bias in research is still in early stages. This study has a number of limitations, which open up exciting avenues for future investigation. The primary limitation was methodology where articles are limit to business, management and accounting to develop the literature review. Further research can be explored for AI bias in computer and medical field. Also, study can be carried out on consumer bias or pricing bias while purchasing products through e-commerce. Further research can be explored on job automation bias may be conducted in order to address and reduce gender and racial bias in recruitment of AI systems. In addition, the study looks into social data bias to address data security and ethical concerns. Additionally the research can explore on ethical bias in tourism and hospitality sector. Lastly, the research can be conducted on AI bias in the health insurance sector in terms of product development leads to customer bias,and risk evaluation based on limited training data sets .

7. Conclusion

There is massive digital disruption in the industries and deploying AI for decision-making to achieve firm success. Hence, there are numerous flaws that have been identified in various domains. We discussed how to be responsible AI for firms, customers, and stakeholders because we are still in the e stages where vulnerabilities can be minimised. Moving forward in time, firms demonstrate that artificial intelligence cannot compete with human or emotional intelligence in online economy. Since AI bias research is in its infancy phase in all the management domain and there will be lots of new opportunities for the firms for innovation, research or policies can be introduced to minimise the bias.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdel-Aty, M., & Haleem, K. (2011). Analyzing angle crashes at unsignalized intersections using machine learning techniques. *Accident Analysis & Prevention, 43*(1), 461–470. https://doi.org/10.1016/j.aap.2010.10.002.

Accenture (2021). “Responsible AI: From principles to practice” https://www.accenture.com/us-en/insights/artificial-intelligence/responsible-ai-principles-practice Retrieved from 14h August 2022.

Accenture (2018). Accenture launches new artificial intelligence testing services, https://newsroom.accenture.com/news/accenture-launches-new-artificial-intelligence-testing-services.htm}:~:Text=Accenture's%20%E2%80%9CTeach%20and%20Test%E2%80%9D%20methodology,used%20to%20train%20machine%20learning Retrieved from Nov 26th, 2022.

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access: Practical Innovations, Open solutions, 6*, 52138–52160.

Akter, S., Michael, K., Uddin, M. R., McCarthy, G., & Rahman, M. (2020). Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics. *Annals of Operations Research*, 1–33. https://doi.org/10.1007/s10479-020-03620-w.

Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research, 144*, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al., (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion, 58*, 82–115.

Bader, V., & Kaiser, S. (2019). Algorithmic decision-making? The user interface and its role for human involvement in decisions supported by artificial intelligence. *Organization, 26*(5), 655–672. https://doi.org/10.1177/1350508419855714.

Batra, G., Queirolo, A., & Santhanam, N. (2018). *Artificial intelligence: The time to act is now*. McKinsey Retrieved February 28, 2022 from https://www.mckinsey.com/industries/advanced-electronics/our-insights/artificial-intelligence-the-time-to-act-is-now.

BBC, (2019). Apple's 'sexist' credit card investigated by US regulator", https://www.bbc.com/news/business-50365609 Retrieved from 28th February 2022.

Beattie, G., & Johnson, P. (2012). Possible unconscious bias in recruitment and promotion and the need to promote equality. *Perspectives: Policy and Practice in Higher Education, 16*(1), 7–13. https://doi.org/10.1080/13603108.2011.611833.

Bellamy, R. K., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., et al., (2019). AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM Journal of Research and Development, 63*(4/5), 4 -1.

Bennet, R.(2014). Google-Announcement for Addtion of Rather Not Say and Custom gender options. https://plus.google.com/118279113645730324236/posts/FKK2trDERAC Retrieved from 2nd March 2022.

Brit, P. (2020). “Overcoming Bias Requires an AI Reboot”. Speech Technology; Medford Vol. 25, Iss. 2, (Spring 2020): 14–17.

Black, J. S., & van Esch, P. (2021). AI-enabled recruiting in the war for talent. *Business Horizons, 64*(4), 513–524. https://doi.org/10.1016/j.bushor.2021.02.015.

Brighton, H., & Gigerenzer, G. (2015). The bias bias. *Journal of Business Research, 68*(8), 1772–1784. https://doi.org/10.1016/j.jbusres.2015.01.061.

Brit, P. (2021). Tips for battling bias in ai-based personalization https://www.destinationcrm.com/Articles/Editorial/Magazine-Features/Tips-for-Battling-Bias-in-AI-Based-Personalization-147143.aspx Retrieved from 21st February 2022.

Brow, A. (2021). “The AI-bias problem and how Fintech's should be fighting it: A deep- dive with Sam Farao”, https://www.forbes.com/sites/anniebrown/2021/09/29/the-ai-bias-problem-and-how-fintechs-should-be-fighting-it-a-deep-dive-with-sam-farao/?sh=38b226492129 Retrieved from 28th February 2022.

Cao, G., Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2021). Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making. *Technovation, 106*, Article 102312. https://doi.org/10.1016/j.technovation.2021.102312.

Casualty Actuarial Society (2022)., “Approaches to Address Racial Bias in Financial Services: Lessons for the Insurance Industry”, https://www.casact.org/sites/default/files/2022-03/Research-Paper_Approaches-to-Address-Racial-Bias_0.pdf. Retrived from 29 December, 2022.

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety, 28*(3), 231–237.

Cheng, X., Lin, X., Shen, X. L., Zarifis, A., & Mou, J. (2022). The dark sides of AI. *Electronic Markets, 32*(1), 11–15. https://doi.org/10.1007/s12525-022-00531-5.

Chintalapudi, N., Battineni, G., & Amenta, F. (2021). Sentimental analysis of COVID-19 tweets using deep learning models. *Infectious Disease Reports, 13*(2), 329–339. https://doi.org/10.3390/idr13020032.

Cisco (2022). “Cisco Principles for Responsible Artificial Intelligence” https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/cisco-responsible-artificial-intelligence-principles.pdf Retrieved from 14th April 2022.

Cohen, P. R., & Feigenbaum, E. A. (2014). The handbook of artificial intelligence: Volume 3 *(Vol. 3)*. Butterworth-Heinemann.

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, Article 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383.

Colquitt, J. A., & Rodell, J. B. (2015). Measuring justice and fairness. In *The oxford handbook of justice in the workplace: 1* (pp. 187–202). Oxford University Press.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the world's we need*. The MIT Press.

Cranfield, J. A., Eales, J. S., Hertel, T. W., & Preckel, P. V. (2003). Model selection when estimating and predicting consumer demands using international, cross section data. *Empirical economics, 28*(2), 353–364. https://doi.org/10.1007/s001810200135.

Dantcheva, A., & Brémond, F. (2016). Gender estimation based on smile–dynamics. *IEEE Transactions on Information Forensics and Security, 12*(3), 719–729. https://doi.org/10.1109/TIFS.2016.2632070.

Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. In *Ethics of data and analytics* (pp. 296–299). Auerbach Publications.

Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated experiments on ad privacy settings. *Proceedings on Privacy Enhancing Technologies, 2015*(1), 92–112.

de Almeida, P. G. R., dos Santos, C. D., & Farias, J. S. (2021). Artificial intelligence regulation: A framework for governance. *Ethics and Information Technology, 23*(3), 505–525. https://doi.org/10.1007/s10676-021-09593-z.

de Graaf, M. M. A., & Allouch, S. B. (2017). The influence of prior expectations of a robot's lifelikeness on users' intentions to treat a zoomorphic robot as a companion. *International Journal of Social Robotics, 9*(1), 17–32. https://doi.org/10.1007/s12369-016-0340-4.

Dibeklioğlu, H., Alnajar, F., Salah, A. A., & Gevers, T. (2015). Combining facial dynamics with appearance for age estimation. *IEEE Transactions on Image Processing, 24*(6), 1928–1943. https://doi.org/10.1109/TIP.2015.2412377.

Dicki, J. (2021). "The ethics dilemma of AI for sales" https://www.destinationcrm.com/Articles/Columns-Departments/Reality-Check/The-Ethics-Dilemma-of-AI-for-Sales-149181.aspx.

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer Nature.

Dilmegan, C. (2022). Bias in AI: What it is, Types, Examples & 6 Ways to Fix it in 2022. AIMultiple.

Dovidio, J. F., & Gaertner, S. L. (1996). Affirmative action, unintentional racial biases, and intergroup relations. *Journal of Social Issues, 52*(4), 51–75.

Drage, E., & Mackereth, K. (2022). Does AI debias recruitment? race, gender, and AI's "eradication of difference". *Philosophy & Technology, 35*(4), 1–25. https://doi.org/10.1007/s13347-022-00543-1.

Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances, 4*(1), eaao5580. https://doi.org/10.1126/sciadv.aao5580.

Durach, C. F., Kembro, J., & Wieland, A. (2017). A new paradigm for systematic literaturereviews in supply chain management. *Journal of Supply Chain Management, 53*(4), 67–85. https://doi.org/10.1111/jscm.12145.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., et al., (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*, Article 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002.

Edelman, B., Luca, M., & Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American economic journal: Applied economics, 9*(2), 1–22.

Edionwe, T. (2017). The fight against racist algorithms: Can we teach our machines to unlearn racism. *The Outline*.

Elia, G., Margherita, A., & Passiante, G. (2020). Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. *Technological Forecasting and Social Change, 150*, Article 119791. https://doi.org/10.1016/j.techfore.2019.119791.

Fatem, F. (2020). "Three Platforms Where AI Bias Lives", https://www.forbes.com/sites/falonfatemi/2020/04/15/three-platforms-where-ai-bias-lives/?sh=68393df3b0c1 Retrieved from 28th December 2022.

Federal Trade Commission (FTC, 2018), "FTC Hearing #7: The Competition and Consumer Protection Issues of Algorithms, Artificial Intelligence, and Predictive Analytics", https://www.ftc.gov/news-events/events/2018/11/ftc-hearing-7-competition-consumer-protection-issues-algorithms-artificial-intelligence-predictive Retrieved from 28 December, 2022.

Fuchs, D. J. (2018). The dangers of human-like bias in machine-learning algorithms. *Missouri S&T's Peer to Peer, 2*(1), 1.

Garg, S., Sinha, S., Kar, A. K., & Mani, M. (2021). A review of machine learning applications in human resource management. *International Journal of Productivity and Performance Management*. https://doi.org/10.1108/IJPPM-08-2020-0427.

Gonzales, R. M. D., & Hargreaves, C. A. (2022). How can we use artificial intelligence for stock recommendation and risk management? A proposed decision support system. *International Journal of Information Management Data Insights, 2*(2), Article 100130. https://doi.org/10.1016/j.jjimei.2022.100130.

Google (2019). https://blog.google/technology/developers/io19-helpful-google-everyone/ Retrieved from Nov 26th, 2022.

Grewal, D., Guha, A., Satornino, C. B., & Schweiger, E. B. (2021). Artificial intelligence: The light and the darkness. *Journal of Business Research, 136*, 229–236. https://doi.org/10.1016/j.jbusres.2021.07.043.

Grundner, L., & Neuhofer, B. (2021). The bright and dark sides of artificial intelligence: A futures perspective on tourist destination experiences. *Journal of Destination Marketing & Management, 19*, Article 100511.

Grusky, D. (2019). *Social stratification, class, race, and gender in sociological perspective*. Routledge.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI – Explainable artificial intelligence. *Science Robotics, 4*(37), eaay7120.

Hamilton, I. A. (2018). *Why it's totally unsurprising that amazon's recruitment AI was biased against women*. Business Insider October 13Available at https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10 from 28th February 2022.

Hammon, K.(2016). 5 unexpected sources bias in artificial intelligence https://techcrunch.com/2016/12/10/5-unexpected-sources-of-bias-in-artificial-intelligence/Retrieved from 1st March 2022

Hao, K. (2021). Facebook's ad algorithms are still excluding women from seeing jobs. *MIT Technology Review, 21*, 2022 *Retrieved January*.

Hardesty, L. (2018). Study finds gender and skin-type bias in commercial artificial-intelligence systems. *Retrieved April*, 3, 2019.

Haring, K. S., Watanabe, K., Velonaki, M., Tossell, C. C., & Finomore, V. (2018). FFAB—The form function attribution bias in human–robot interaction. *IEEE Transactions on Cognitive and Developmental Systems, 10*(4), 843–851. https://doi.org/10.1109/TCDS.2018.2851569.

Herath, H. M. K. K. M. B., & Mittal, M. (2022). Adoption of artificial intelligence in smart cities: A comprehensive review. *International Journal of Information Management Data Insights, 2*(1), Article 100076. https://doi.org/10.1016/j.jjimei.2022.100076.

Hossain, M., & Muromachi, Y. (2012). A Bayesian network based framework for real-time crash prediction on the basic freeway segments of urban expressways. *Accident Analysis & Prevention, 45*, 373–381. https://doi.org/10.1016/j.aap.2011.08.004.

Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science, 49*(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9/.

IBM Research, (2018). "AI bias will explode. But only the unbiased AI will survive," [online] Available at: https://newsroom.ibm.com/IBM-research?item=30305 Retrieved from 28th December 2022.

Jared Council (2021). "How adobe's ethics committee helps manage AI bias; A diverse selection of voices can help companies spot potential problems", https://www.wsj.com/articles/how-adobes-ethics-committee-helps-manage-ai-bias-11620261997 Retrieved from 28th February 2022.

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

Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., et al., (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius, 7*, Article 2378023121999581. https://doi.org/10.1177/2378023121999581.

Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research–Moving away from the "What" towards the "Why". *International Journal of Information Management, 54*, Article 102205. https://doi.org/10.1016/j.ijinfomgt.2020.102205.

Kar, A. K., & Kushwaha, A. K. (2021). Facilitators and barriers of artificial intelligence adoption in business–insights from opinions using big data analytics. *Information Systems Frontiers*, 1–24. https://doi.org/10.1007/s10796-021-10219-4.

Kar, A. K., Choudhary, S. K., & Singh, V. K. (2022). How can artificial intelligence impact sustainability: A systematic literature review. *Journal of Cleaner Production*, Article 134120. https://doi.org/10.1016/j.jclepro.2022.134120.

Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 3819–3828). https://doi.org/10.1145/2702123.2702520.

Knight, W. (2017). Forget killer robots bias is the real AI danger. https://www.technologyreview.com/2017/10/03/241956/forget-killer-robotsbias-is-the-real-ai-danger/ Retrieved February 28, 2022.

Kuchenbrandt, D., Eyssel, F., Bobinger, S., & Neufeld, M. (2013). When a robot's group membership matters. *International Journal of Social Robotics, 5*(3), 409–417. https://doi.org/10.1007/s12369-013-0197-8.

Kumar, P., Dwivedi, Y. K., & Anand, A. (2021). Responsible artificial intelligence (AI) for value formation and market performance in healthcare: The mediating role of Patient's cognitive engagement. *Information Systems Frontiers*, 1–24. https://doi.org/10.1007/s10796-021-10136-6.

Kumar, P., Hollebeek, L. D., Kar, A. K., & Kukk, J. (2022). Charting the intellectual structure of customer experience research. *Marketing Intelligence & Planning* ahead-of-print. https://doi.org/10.1108/MIP-05-2022-0185.

Kushwaha, A. K., Kumar, P., & Kar, A. K. (2021). What impacts customer experience for B2B enterprises on using AI-enabled chatbots? Insights from Big data analytics. *Industrial Marketing Management, 98*, 207–221. https://doi.org/10.1016/j.indmarman.2021.08.011.

Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management science, 65*(7), 2966–2981. https://doi.org/10.1287/mnsc.2018.3093.

Lee, N.T., Resnick, P. and Barton, G. (2019). https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/}footnote-16 Retrieved 14th February 2022.

Letheren, K., Russell-Bennett, R., & Whittaker, L. (2020). Black, white or grey magic? Our future with artificial intelligence. *Journal of Marketing Management, 36*(3–4), 216–232. https://doi.org/10.1080/0267257X.2019.1706306.

Liesse, J. (2021). Advertising age. *Chicago, 92*(13), 8. IssSep 20 https://www.proquest.com/trade-journals/ai-marketing-what-brands-need-know/docview/2575515238/se-2?accountid=177896.

Lindsa, P. (2019). Quality progress. *Milwauke, 10*(9), 6–8. Oc https://www.proquest.com/docview/2312156914/fulltextPDF/A02089CE539F4E3FPQ/3?accountid=177896. Retrieved from March 2nd, 2022.

Liu, Y., Yan, W., & Hu, B. (2021). Resistance to facial recognition payment in China: The infuence of privacy–related factors. *Telecommunications Policy, 45*(5), Article 1021155. https://doi.org/10.1016/j.telpol.2021.102155.

Lockey, S., Gillespie, N., Holm, D., & Someh, I. A. (2021). A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions. http://hdl.handle.net/10125/71284. Retrieved 14 Dec 2022.

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Machines versus humans: the impact of AI chatbot disclosure on customer purchases. Luo, X, Tong S, Fang Z, Qu, (2019), 20–33.

Malik, N., Tripathi, S. N., Kar, A. K., & Gupta, S. (2021). Impact of artificial intelligence on employees working in industry 4.0 led organizations. *International Journal of Manpower*. https://doi.org/10.1108/IJM-03-2021-0173.

McGettigan, T. (2017). Artificial Intelligence: Is Watson the Real Thing?. The IUP Journal of Information Technology, 13(2), 44–69.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.

Messner, W. (2022). Improving the cross-cultural functioning of deep artificial neural networks through machine enculturation. *International Journal of Information Management Data Insights, 2*(2), Article 100118. https://doi.org/10.1016/j.jjimei.2022.100118.

Microsoft (2022). "Responsible AI", https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6 Retrieved 14 April 2022.

Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side'of AI. *European Journal of Information Systems*, 1–12. https://doi.org/10.1080/0960085X.2022.2026621.

Miller, Alex, & Hosanagar, Kartik (2020). Personalized discount targeting with causal machine learning. In *Proceedingss of the ICIS 2020* https://aisel.aisnet.org/icis2020/digital_commerce/digital_commerce/7.

Morande, S. (2022). Enhancing psychosomatic health using artificial intelligence-based treatment protocol: A data science-driven approach. *International Journal of Information Management Data Insights, 2*(2), Article 100124. https://doi.org/10.1016/j.jjimei.2022.100124.

Morse, J. (2017). App creator apologizes for 'racist'filter that lightens skin tones. *Mashable*. available at: https://mashable.com/2017/04/24/faceapp-racism-selfie/}zeUItoQB5iqI. *accessed 5 March 2018*.

Murphy, L. W. (2016). Airbnb's work to fight discrimination and build inclusion. *Report submitted to Airbnb, 8*, 2016.

Nagwani, N. K., & Suri, J. S. (2023). An artificial intelligence framework on software bug triaging, technological evolution, and future challenges: A review. *International Journal of Information Management Data Insights, 3*(1), Article 100153. https://doi.org/10.1016/j.jjimei.2022.100153.

Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*(2133), Article 20180089. https://doi.org/10.1098/rsta.2018.0089.

Noble, S. U. (2018). *Algorithms of oppression*. New York University Press.

Noyes, K. (2015). *Will big data help end discrimination—Or make it worse?*. Fortune.

Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: Implications for health systems. *Journal of Global Health, 9*(2). https://doi.org/10.7189/jogh.09.020318.

Pawar, D. S., Yadav, A. K., Akolekar, N., & Velaga, N. R. (2020). Impact of physical distancing due to novel coronavirus (SARS-CoV-2) on daily travel for work during transition to lockdown. *Transportation Research Interdisciplinary Perspectives, 7*, Article 100203. https://doi.org/10.1016/j.trip.2020.100203.

Penny, L. (2017). Robots are racist and sexist. *Just like the people who created them*. The Guardian.

Qi, X., Luo, Y., Wu, G., Boriboonsomsin, K., & Barth, M. (2019). Deep reinforcement learning enabled self-learning control for energy efficient driving. *Transportation Research Part C: Emerging Technologies, 99*, 67–81. https://doi.org/10.1016/j.trc.2018.12.018.

Ramirez, E., Brill, J., Ohlhausen, M. K., & McSweeny, T. (2016). Big data: A tool for inclusion or exclusion? Understanding the issues. *Report, Federal Trade Commission, Washington (DC)*. https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf. *accessed August, 20*, 2018.

Rinta-Kahila, T., Penttinen, E., Salovaara, A., & Soliman, W. (2018, January). Consequences of discontinuing knowledge work automation-surfacing of deskilling effects and methods of recovery. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. (Vol. 2018, pp. 5244–5253). Hawaii International Conference on System Sciences.

Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in AI. In *Proceedings of The Companion 2019 World Wide Web Conference* (pp. 539–544).

Sarle, W. S. (1994). Artificial neural networks and statistical models. In *Proceedings of the Nineteenth Annual SAS Users Group International Conference* (pp. 1538–1550). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.27.699.

Schrader, D. E., & Ghosh, D. (2018). Proactively protecting against the singularity: Ethical decision making in AI. *IEEE Security & Privacy, 16*(3), 56–63.

Schroeder, J. E. (2021). Reinscribing gender: Social media, algorithms, bias. *Journal of Marketing Management, 37*(3–4), 376–378.

Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham L. Rev., 87*, 1085.

Sharma, R., Kumar, A., & Chuah, C. (2021). Turning the blackbox into a glassbox: An explainable machine learning approach for understanding hospitality customer. *International Journal of Information Management Data Insights, 1*(2), Article 100050. https://doi.org/10.1016/j.jjimei.2021.100050.

Shekhawat, N., Chauhan, A., & Muthiah, S. B. (2019). Algorithmic privacy and gender bias Issues in google ad settings. *In Proceedings of the 10th ACM Conference on Web Science*, 281–285.

Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review, 61*(4), 66–83. https://doi.org/10.1177/0008125619862257.

Silberg, J., & Manyika, J. (2019). Notes from the AI frontier: Tackling bias in AI (and in humans). *McKinsey Global Institute*, 1–6.

Singapore Government. (2021). *AI Singapore* https://www.nrf.gov.sg/programmes Retrieved from 14th April 2022.

Singh, V., Konovalova, I., & Kar, A. K. (2022). When to choose ranked area integrals versus integrated gradient for explainable artificial intelligence–a comparison of algorithms. *Benchmarking: An International Journal* ahead-of-print. https://doi.org/10.1108/BIJ-02-2022-0112.

Sridevi, G. M., & Suganthi, S. K. (2022). AI based suitability measurement and prediction between job description and job seeker profiles. *International Journal of Information Management Data Insights, 2*(2), Article 100109. https://doi.org/10.1016/j.jjimei.2022.100109.

Srinivasan, S., Rutz, O. J., & Pauwels, K. (2016). Paths to and off purchase: Quantifying the impact of traditional marketing and online consumer activity. *Journal of the Academy of Marketing Science, 44*(4), 440–453. https://doi.org/10.1007/s11747-015-0431-z.

Srivastava, S., Bisht, A., & Narayan, N. (2017). Safety and security in smart cities using artificial intelligence – A review. In *Proceedings of the 2017 7th International conference on cloud computing, data science & engineering-confluence* (pp. 130–133). IEEE.

Sutton, S. G., Arnold, V., & Holt, M. (2018). How much automation is too much? Keeping the human-relevant in knowledge work. *Journal of Emerging Technologies in Accounting, 15*(2), 15–25.

Sweeney, L. (2013). Discrimination in online ad delivery. *Communications of the ACM, 56*(5), 44–54.

Teleaba, F., Popescu, S., Olaru, M., & Pitic, D. (2021). Risks of observable and unobservable biases in artificial intelligence used for predicting consumer choice. *Amfiteatru Economic, 23*(56), 102–119.

The Conversation (2022). "Artificial Intelligence can discriminate on the basis of race, gender and also age". https://theconversation.com/artificial-intelligence-can-discriminate-on-the-basis-of-race-and-gender-and-also-age-173617 Retrieved from 1st March 2022.

Tiwary, N. K., Kumar, R. K., Sarraf, S., Kumar, P., & Rana, N. P. (2021). Impact assessment of social media usage in B2B marketing: A review of the literature and a way forward. *Journal of Business Research, 131*, 121–139. https://doi.org/10.1016/j.jbusres.2021.03.028.

Tredinnick, L. (2017). Artificial intelligence and professional roles. *Business Information Review, 34*(1), 37–41. https://doi.org/10.1177/0266382117692621.

Ukanwa, K., & Rust, R. T. (2020). Discrimination in service. *Marketing Science Institute Working Paper Series Report*, 18–121.

Verma, S., Sharma, R., Deb, S., & Maitra, D. (2021). Artificial intelligence in marketing: Systematic review and future research direction. *International Journal of Information Management Data Insights, 1*(1), Article 100002. https://doi.org/10.1016/j.jjimei.2020.100002.

Villasenor(2019), John (February). Artificial intelligence and bias: Four key challenges, 28, 2022–2023. https://www.brookings.edu/blog/techtank/2019/01/03/artificial-intelligence-and-bias-four-key-challenges/Retrieved.

Vimalkumar, M., Gupta, A., Sharma, D., & Dwivedi, Y. (2021). Understanding the effect that task complexity has on automation potential and opacity: Implications for algorithmic fairness. *AIS Transactions on Human-Computer Interaction, 13*(1), 104–129. https://doi.org/10.17705/1thci.00144.

Vincent, J. (2018). *Amazon reportedly scraps internal ai recruiting tool that was biased against women*. The Verge October 10Available at https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report last accessed April 20, 2019.

Votto, A. M., Valecha, R., Najafirad, P., & Rao, H. R. (2021). Artificial intelligence in tactical human resource management: A systematic literature review. *International Journal of Information Management Data Insights, 1*(2), Article 100047. https://doi.org/10.1016/j.jjimei.2021.100047.

Wachter-Boettcher, S. (2017). *Ai recruiting tools do not eliminate bias*. Time Magazine.

Waja, G., Patil, G., Mehta, C., & Patil, S. (2023). How AI can be used for governance of messaging services: A study on spam classification leveraging multi-channel convolutional neural network. *International Journal of Information Management Data Insights, 3*(1), Article 100147. https://doi.org/10.1016/j.jjimei.2022.100147.

Weed, M. (2006). Sports tourism research 2000–2004: A systematic review of knowledge and a meta-evaluation of methods. *Journal of Sport & Tourism, 11*(1), 5–30. https://doi.org/10.1080/14775080600985150.

Weissman, J. (2018). "Amazon created a hiring tool using a.I. it immedi- ately started discriminating against women". Retrieved October 10, 2022 from https://slate.com/business/2018/10/amazon-artificial-intelligence-hiring-discrimination-women.html.

Weyerer, J. C., & Langer, P. F. (2020). Bias and Discrimination in artificial intelligence: Emergence and Impact in E-business. In *Interdisciplinary approaches to digital transformation and innovation* (pp. 256–283). IGI Global. https://doi.org/10.4018/978-1-7998-1879-3.ch011.

Wigger, K. (2020). "Researchers find racial discrimination in 'dynamic pricing' algorithms used by Uber, Lyft, and others" https://venturebeat.com/2020/06/12/researchers-find-racial-discrimination-in-dynamic-pricing-algorithms-used-by-uber-lyft-and-others/ Retrieved February 28, 2022.

Wong, P. H. (2020). Democratizing algorithmic fairness. *Philosophy & Technology, 33*(2), 225–244. https://doi.org/10.1007/s13347-019-00355-w.

Woodruff, A., Fox, S. E., Rousso-Schindler, S., & Warshaw, J. (2018). A qualitative exploration of perceptions of algorithmic fairness. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1–14).

Yarger, L., Payton, F. C., & Neupane, B. (2019). Algorithmic equity in the hiring of underrepresented IT job candidates. *Online Information Review, 44*(2), 383–395. https://doi.org/10.1108/OIR-10-2018-0334.

Yen, C., & Chiang, M. C. (2021). Trust me, if you can: A study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments. *Behaviour & Information Technology, 40*(11), 1177–1194.

Zajko, M. (2022). Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates. *Sociology Compass, 16*(3), e12962. https://doi.org/10.1111/soc4.12962.