

DOI: [https://doi.org/10.48009/4\\_iis\\_2026\\_105](https://doi.org/10.48009/4_iis_2026_105)

## Big data Vs for detecting and preventing biases in AI algorithms: a systematic literature review

**Laurence Lawson-Body**, *University of North Dakota*, and [laurence.lawsonbody@und.edu](mailto:laurence.lawsonbody@und.edu)

**Jason Jensen**, *University of North Dakota*, and [jjensen@business.und.edu](mailto:jjensen@business.und.edu)

**Assion Lawson-Body**, *University of North Dakota*, and [assion.lawsonbody@und.edu](mailto:assion.lawsonbody@und.edu)

**Marcellin Makpotche**, *University of Quebec at Rimouski*, and [marcellin\\_makpotche@uqar.ca](mailto:marcellin_makpotche@uqar.ca)

### Abstract

The negative effects of AI algorithm biases on society have become a major concern recently. Many proposed solutions to address these biases have often failed. Our study responds to calls for using big data characteristics in designing training data to detect and prevent AI biases early. A systematic literature review examined the extent and details of algorithm discrimination, drawing on data from 8 databases and social media research sources. Results showed potential solutions for reducing discrimination bias based on big data characteristics such as volume, variety, velocity, veracity, variability, and value. Specifically, a larger volume of big data increases the likelihood of preventing AI bias. The diversity of big data allows for a wider range of training data, drawn from various formats and sources, helping to combat biases. The velocity of big data supports ongoing training, which should help eliminate bias. Veracity ensures the accuracy and truthfulness of training data, aiding early detection of biases. Variability allows for the use of changing, volatile data over time, helping to deter bias. Lastly, the high quality, or value, of big data promotes the use of superior training data to eliminate biases.

**Keywords:** AI, algorithm bias, big data Vs, theory of equality, training data, systematic literature review.

### Introduction

Artificial intelligence (AI) forms the foundation of the fourth industrial revolution (Akter et al., 2023; Sheng et al., 2026). Recently, the social impact of biases in AI algorithms has gained attention as AI technology advances (Chiarella et al., 2022; Jia et al., 2026; Kim & Ryoo, 2026). Addressing these biases is crucial in a world increasingly reliant on AI across various fields. Biases in AI algorithms that adversely affect different groups have been documented in areas like legal systems, healthcare, and business (Varsha, 2023). The prevalence of biased AI systems is growing rapidly across disciplines, resulting in more vulnerable and victimized individuals due to AI's continuous development and sophistication (Varsha, 2023).

Authors argue that AI algorithm biases may be related to numerous factors, including economic, technological, environmental, and cultural factors (Wood, 2025; Kim & Ryoo, 2026). In addition, AI algorithm biases may be linked to specific characteristics, from bias in the training data to bias in the models and algorithm deployments (Akter et al., 2023). As noted by Burton (2023), problems exist in the design of AI algorithms, the data they rely on, how they are used, and their outcomes and impacts. Because AI algorithms are being used to spread discrimination and the digital divide on the Internet, there are

opportunities for researchers in these areas to broaden and deepen research on AI algorithms (Burton, 2023). Moreover, there are concerns that AI developers are ignoring ethics, laws, rules, and common sense in favor of technical and commercial priorities (Rodgers et al., 2023; Wood, 2025; Sheng et al., 2026). Furthermore, AI algorithm design can introduce bias into decision-making without common, widely accepted considerations to guide AI developers.

Many solutions have been proposed to address AI algorithmic biases. Those solutions encompass regulations, ethics, morals, mathematical equations, measurement instruments, and assessments of the decision-making process (Rodgers et al., 2023; Akter et al., 2023; Jia et al., 2026). Most existing suggestions and solutions failed. This is why the AI algorithm bias issues increasingly create more victims in society. The advent of the generative pre-trained transformer (ChatGPT) has recently exacerbated AI algorithm discrimination worldwide. However, when using training data to educate AI algorithms, designers mistakenly treat big data as traditional data, increasing the risk of AI biases. Even though the inception of modern AI stemmed from advances in computational power and the advent of big data, existing solutions have focused on other factors rather than the big data Vs or characteristics in the training data of AI algorithms.

Toffler first coined the term big data in the 1980s (Wang & Wang, 2021). However, it was only in 2008 that the challenges and opportunities of big data from different perspectives, such as the Internet of Things, network economy, and bio-medicine, were discussed (Wang & Wang, 2021). Since that time, big data has gradually attracted the attention of organizations and individuals. Big data acquisition is related to the characteristics of big data: variety, volume, value, velocity, veracity, and variability (Ghasemaghaei & Calic, 2020; Youssef et al., 2022; Ye & Meiyu, 2022; Zhang et al., 2022; Chou et al., 2022), where volume refers to the size of data, velocity is the speed at which data are generated, variety is the different types of data, variability refers to how the structure and meaning of data constantly change, veracity concerns the accuracy and precision of data, and value concerns the potential benefits that big data can offer (Zhang et al., 2022).

The advent of big data as the source of new AI insights was ignored and neglected when the research community was called to investigate AI algorithmic bias. This study fills this gap in the literature by examining the nature of big data Vs or characteristics to solve the AI algorithm bias. The approach suggested in this study focuses on the early detection and prevention of algorithm biases. This problem allows us to propose the following overarching research question, with several sub-questions that address specific big data characteristics.

- *RQ: How can big data Vs or characteristics be used for early detection and prevention of AI algorithm biases?*
  - *How can big data volume be used to solve AI algorithm biases?*
  - *How does big data variety in designing training data reduce or minimize AI algorithm biases?*
  - *How can big data veracity be used to address AI algorithm biases?*
  - *How can big data velocity be used to solve AI algorithm biases?*
  - *How can big data value be used to fix AI algorithm biases?*
  - *How can big data variability be used to overcome AI algorithm biases?*

The objective of this study is to use big data Vs or characteristics for early detection and prevention of AI Algorithm biases. We then proposed big data Vs or characteristics awareness and fairness to overcome these types of bias. To accomplish this objective, we conducted a systematic literature review.

As a result, this study makes the following significant contributions to the literature. First, it proposes big data Vs or characteristics for early anomaly detection and prevention in AI algorithms. Second, it extends the list of solutions to AI algorithm bias in society. Third, it focuses on training data as the primary driver of AI algorithm biases. Data used to train AI algorithms often contains biases that can lead to discrimination and unfairness (Gonzalez-Sendino et al., 2024). Fourth, it is one of the first studies to use a systematic literature review to establish a bridge between big data and AI algorithm biases. Fifth, it extends the theory of equality using big data and AI.

This study is organized as follows. Section 2 outlines the theoretical background. Section 3 proposes the literature review. Section 4 presents the research methodology. Section 5 formulates the findings. Section 6 details the analysis and discussions. Finally, section 7 provides the conclusion.

## **Theoretical Background: Theory of Equality and AI**

The theory of equality has been used worldwide because stereotypes of inequality exist across race, gender, religion, age, ethnicity, political affiliation, education, healthcare, and the workplace (Peng et al., 2024; Acheampong et al., 2024). These inequalities continue to be used across many domains and fields for reasons no one can explain. Even though gender inequality is significantly lower than in the last century, many countries struggle to accept and support gender equality, with many believing that women should have equal rights with men in the workplace and in political leadership (Peng et al., 2024). Gender and educational inequalities are ethical and moral concerns that policymakers must address (Acheampong et al., 2024). Peng et al. (2024) found that digital technology can affect stereotypes about inequality. Recently, AI algorithm biases have exacerbated the inequality stereotypes. This study aims to make significant progress towards solutions that prevent the negative impacts of AI algorithmic biases on the theory of equality. To accomplish this goal, this study borrowed big data Vs or characteristics to improve the training data of AI algorithms and diminish the discrimination society is currently facing.

First, bias in algorithms often stems from unrepresentative or biased training data, which can restrict fair access to accurate, culturally relevant AI-generated information. This may result in users from diverse demographic, cultural, or linguistic backgrounds not having equal opportunities to benefit from generative AI. Additionally, ongoing disparities in AI outcomes can reinforce social and educational inequalities. Second, this study sees AI bias not just as a technical issue, but as a reflection of societal inequalities reproduced or worsened by biased training data. Therefore, fairness in AI involves both algorithmic performance and equitable data representation, ensuring diverse groups are adequately represented.

## **Literature Review**

### **Artificial Intelligence (AI)**

AI is part of the electronic ecosystem, including big data, cloud computing, the Internet of Things, quantum computing, and high-speed computer networks (5G/6G) (Burton, 2023). It affects people's daily lives through interconnected intelligent objects (Rosca et al., 2024; Wood, 2025). AI aims to enable machines to learn, think, reason, and make decisions like humans and to help develop applications such as self-driving and automatic identification (Li, 2022). Most AI systems exhibit intelligence commensurate with or exceeding human intelligence (Burton, 2023; Sheng et al., 2026).

There is no unanimity in the definition of AI. A non-technical definition of AI states that a computer system can perform tasks typically performed by human intelligence, including learning, reasoning, and self-correction or feedback (Marler, 2024). In other words, AI is the theory and development of computer

systems capable of performing tasks requiring human intelligence (Varsha, 2023). In this sense, it is a transformative technology that can revolutionize industries, enhance efficiency, and augment human capabilities (Varsha, 2023). Similarly, AI is a technology that attempts to simulate human reasoning in computers and other machines (Rodgers et al., 2023). AI also encompasses intelligent behavior and refers to algorithms that enable a computer system to behave like a human, including supervised and unsupervised machine learning, artificial neural networks, computer vision, deep learning, natural language processing, and fuzzy logic (Fountain, 2022). Additionally, there is an interaction between AI and social relations among the non-technical definitions of AI. Burton (2023) emphasizes this by considering AI as a registry of power that reflects and generates social relations and competitions worldwide (Burton, 2023). Moreover, AI brings convenience and comfort to work and society (Li, 2022) and can be described as a system's ability to interpret external data accurately and learn from it to achieve flexible adaptation (Marler, 2024).

However, AI models' inability to identify when a problem is outside the model's scope has been a problem for decades (Fountain, 2022). AI's problematic characteristics make it different from human beings, who can understand when a problem is outside the scope of a phenomenon. This reinforces AI's objective nature compared with human beings' subjective nature. The absence of a subjective nature in AI is evident in many AI algorithms, thereby generating unacceptable biases.

## AI Algorithms

AI algorithms may be associated with process-driven or data-driven logic. The AI algorithm-oriented process-driven logic may be as simple as a decision rule, as complex as an advanced optimization problem, or as integrated as an autonomous car algorithm (Alter, 2022). Specifically, AI algorithms are rules or procedures inferred from detected patterns composed of positive and negative correlations (Fountain, 2022; Kim & Ryoo, 2026). Also, it can perform complex computations much faster and make more objective decisions than humans (Pessach & Shmueli, 2021). Most importantly, AI algorithms perform activities including 1) freeing up time to devote to tasks that require more judgment and intuition, 2) reducing errors and increasing fairness, 3) increasing efficiency, and 4) increasing accessibility to more information (Marler, 2024). In this context, engineers, scientists, and programmers use algorithms to design a learning machine (Rodgers et al., 2023; Jia et al., 2026). Therefore, AI algorithms are rules formalized in AI applications (Fountain, 2022). AI algorithms can also lower prediction costs and increase value propositions (Cecere et al., 2024). Furthermore, AI algorithms have empowered individuals and organizations to make decisions and take action on society's behalf across many domains due to their efficiency (Rodgers et al., 2023; Kim & Ryoo, 2026).

The AI algorithm-oriented data-driven logic is an abstract procedure for converting data inputs into data outputs (Alter, 2022). Similarly, AI algorithms are automated procedures that transform data into desired information in ways that are often more opaque than those of previous technological systems (Marler, 2024). Equally important, Algorithms used in AI are specifications for performing calculations, data processing, automated reasoning, and other tasks (Rodgers et al., 2023). The algorithms are the mechanism by which the machine processes data (Rodgers et al., 2023). Therefore, AI algorithms can process datasets of unlimited size, generating valuable insights for businesses (Cecere et al., 2024). Additionally, AI algorithms may integrate large-scale datasets under complex uncertainties (Pessach & Shmueli, 2021).

However, AI algorithms must be inclusive, non-discriminatory, and fair by design (Delecraz et al., 2022). Their outcomes must also be ethical and provide equal opportunities. AI algorithms face two challenges: algorithm uncertainty and data uncertainty (Li et al., 2022). Algorithm uncertainty concerns the system's lack of transparency, leading to an unknown reliability of its results and preventing users from knowing when the algorithm might fail. Data uncertainty deals with naturally heterogeneous datasets in real life. The opacity of AI algorithms prevents users from knowing when they produce wrong results (Li et al., 2022).

The natural gap between training datasets and real-world data may lead to unexpected AI system malfunction (Li et al., 2022).

## AI Algorithm biases

There is no unanimity on the definition of AI algorithm bias. For this study, we define it as the inconvenience and benefits that AI can bring to a specific group of people or communities more than to others, without any objective justification for such discrimination (Cecere et al., 2024; Jia et al., 2026; Kim & Ryoo, 2026). In simpler terms, AI algorithm bias refers to the unfair advantages or disadvantages that AI systems can create for certain groups without any valid reason for these disparities. AI algorithm biases arise when AI applications for social problem-solving create paths that differ across demographic groups and produce anomalous results (Fountain, 2022). That generates discrimination against traditionally marginalized groups embedded in data, procedures, processes, and policies (Fountain, 2022). In this sense, AI algorithm biases can also generate and perpetuate discrimination against minorities, women, and children, which raises ethical and fairness concerns about AI (Cecere et al., 2024; Kim & Ryoo, 2026). Discrimination can be based on gender, race, age, disability, political affiliation, or other groups that disadvantage them (Fountain, 2022). Fairness is being reasonable, right, and just (Delecraz et al., 2022). It is also defined as the absence of prejudice or favoritism toward an individual or a group based on their inherent or acquired characteristics (Delecraz et al., 2022). One suspected reason for these AI algorithm discriminations is that most of the recorded historical big data were about white people. Therefore, the data an AI algorithm needs to produce expected outcomes is generally unavailable across all communities (Marler, 2024). AI algorithms trained on such data can lead to unfairness and discrimination against specific groups, reinforcing privileged and underprivileged classes in society (Gonzalez-Sendino et al., 2024). AI algorithm biases and vulnerabilities experienced by people across industries lead to gender and racial discrimination (Varsha, 2023). Also, AI algorithms based on biased data encoded those biases into algorithms, models, inputs, and outputs (Fountain, 2022).

AI algorithmic bias may come from the viewpoint of the algorithm creator or maybe the intended result of omissions (Alter, 2022). Unfortunately, AI algorithm biases can be intentional or not, recognized and unknown (Fountain, 2022; Jia et al., 2026). Rodgers et al. (2023) presented a throughput framework that depicts how perceptions, judgments, and the use of information affect the strategic selection of specific ethical decision-making algorithmic pathways. On their side, Akter et al. (2023) designed a framework of AI algorithmic bias management capabilities for industrial markets. They found that AI algorithmic bias management capabilities consist of three pillars: data, model, and deployment. In other words, the sources of AI algorithm bias are embedded in poor training data, weak mathematical models, and historical and social contexts (Akter et al., 2023). Cecere et al. (2024) presented an algorithmic audit framework to assess algorithmic decision-making to detect biases. Rodgers et al. (2023) designed and developed ethical human resource management systems to eliminate AI design bias. Rodgers et al. (2023) proposed a more systematic approach to depicting algorithms by emphasizing ethical behavior in AI systems. They applied the six dominant ethical theories to algorithmic modeling, including ethical egoism (preference-based), deontology (rule-based), utilitarianism, relativism, virtue ethics, and the ethics of care (Rodgers et al., 2023).

In this study, we argue that these AI algorithm biases are learned from training data. Our work reduces algorithmic bias in training data by leveraging big data characteristics. Biases in data may arise from measurement errors and human decisions (Pessach & Shmueli, 2021; Kim & Ryoo, 2026). Additionally, biases may arise from missing values, leading to datasets that are not representative of the target population (Pessach & Shmueli, 2021). However, the emergence of AI algorithm bias in the literature and its widespread recognition as a significant issue has sparked concern among individuals, leaders, governments, organizations, and researchers. We must address this issue, as most of the problems, anxieties, and fears related to the use of AI in society have AI algorithm bias as a common denominator.

## **Big Data Vs or Characteristics**

Big data characteristics help distinguish it from traditional data (Johnson et al., 2017; Chen et al., 2026). Meanwhile, big data, involving enormous, exhaustive, and varied datasets, is beyond the capabilities of conventional database tools to capture, store, manage, and analyze (Lyu et al., 2022). Additionally, the evolution created by big data was due to the advent of unstructured data, which could not be directly processed using traditional tools (Zhang et al., 2021). Regarding this, big data requires specialized data handling and information processing tools, such as machine learning and natural language processing (NLP). This is why big data enabled the explosion of sensor technologies, artificial intelligence, and machine learning techniques (Onyeji-Nwogu et al., 2020). Moreover, big data are derived from three leading sources: user-generated content (UGC), operations (web searching), and the Internet of Things (Lyu et al., 2022). Therefore, the emergence of abundant computing power has enabled the generation and storage of big data (Onyeji-Nwogu et al., 2020).

The literature has highlighted various aspects of big data, such as data in diverse formats, unprecedented scale, and real-time digital data flows (Lin & Kunnathur, 2019). Big data consists of large volumes of diverse data flowing at high speed, collected and stored across multiple locations (Lin & Kunnathur, 2019). These aspects lead to large-scale datasets characterized by volume, variety, and velocity, commonly referred to as the 3Vs (Mishra & Singh, 2020; Hassan et al., 2021). Besides 3Vs, value, veracity, and variability are the other three well-known Vs that describe the characteristics of big data (Mishra & Singh, 2020; Hassan et al., 2021). Overall, big data is an information asset characterized by high volume, velocity, variety, veracity, and variability that requires specific technologies and analytical methods to transform it into value (Hassan et al., 2021). In conclusion, the big data Vs or characteristics herein are volume, variety, velocity, veracity, variability, and value.

Big data volume refers to a collection of large, massive, and complex data from different instruments at all stages of the process, which go from acquisition, storage, and sharing to analysis and visualization (Johnson et al., 2017; Oliveira et al., 2019; Ghasemaghahi & Calic, 2020; Lawson-body et al., 2024; Ahmed et al., 2022; Lawson-body et al., 2024; Chen et al., 2026).

Big data variety refers to the types and formats of data generated (Ghasemaghahi & Calic, 2020; Lawson-body et al., 2024; Chen et al., 2026) that provide a complete picture of the organization's products and services.

Big data veracity refers to the truthfulness and trustworthiness of the data used in competitive advantage decision-making (Ghasemaghahi & Calic, 2020; Chen et al., 2026).

Big data velocity refers to the speed at which big data is produced and analyzed (Johnson et al., 2017; Ghasemaghahi & Calic, 2020).

Big data value deals with the potential benefits that big data can offer (Zhang et al., 2022; Ahmed et al., 2022; Ghasemaghahi & Calic, 2021; Lawson-body et al., 2024).

Big data variability refers to fluctuations in data storage and processing over time (Zhang et al., 2022; Lawson-Body et al., 2024).

## **Integration of Big Data Vs and AI Algorithm Biases**

We highlight that the volume of big data does not automatically remove bias. While larger datasets may improve sample representation, they can also magnify existing biases if underrepresented groups remain insufficiently covered. Additionally, the variety of big data enhances algorithmic fairness by incorporating

diverse sources, demographic groups, languages, and formats, helping address systematic underrepresentation and fostering more inclusive AI models. However, the velocity of big data can introduce bias when rapidly streaming data is used without enough quality checks, potentially reinforcing societal biases before they are identified or fixed. The veracity of big data also affects fairness; inaccurate, incomplete, or biased data can embed discriminatory patterns in AI outputs. Finally, we note that the true value of big data lies in its responsible use, since societal benefits from generative AI depend not only on accuracy but also on fairness, transparency, accountability, and equitable outcomes. Variability in big data over time affects the quality of training data for AI models, influencing whether the outputs are fair or biased.

## Research Methodology

A systematic literature review is a widely used research methodology across multidisciplinary domains (Varsha, 2023). This qualitative approach includes a thematic evaluation, a search for peer-reviewed papers in databases, and interpretive analysis. Online digital libraries are also considered in this search process to conduct the systematic literature review, consistent with Varsha (2023). In addition, the screening of the selected peer-reviewed papers includes the inclusion and exclusion conditions and criteria. The inclusion/exclusion strategies and criteria address filtering and extraction. Mainly, the filtering and extraction method includes one run of duplication and redundancy removal and three runs of inclusion and exclusion. Figure 1 shows the flowchart outlining the research methodology, including the sources (list of databases searched and consulted), research protocol features (search keywords and themes), the automated search, and the inclusion and exclusion steps or phases.

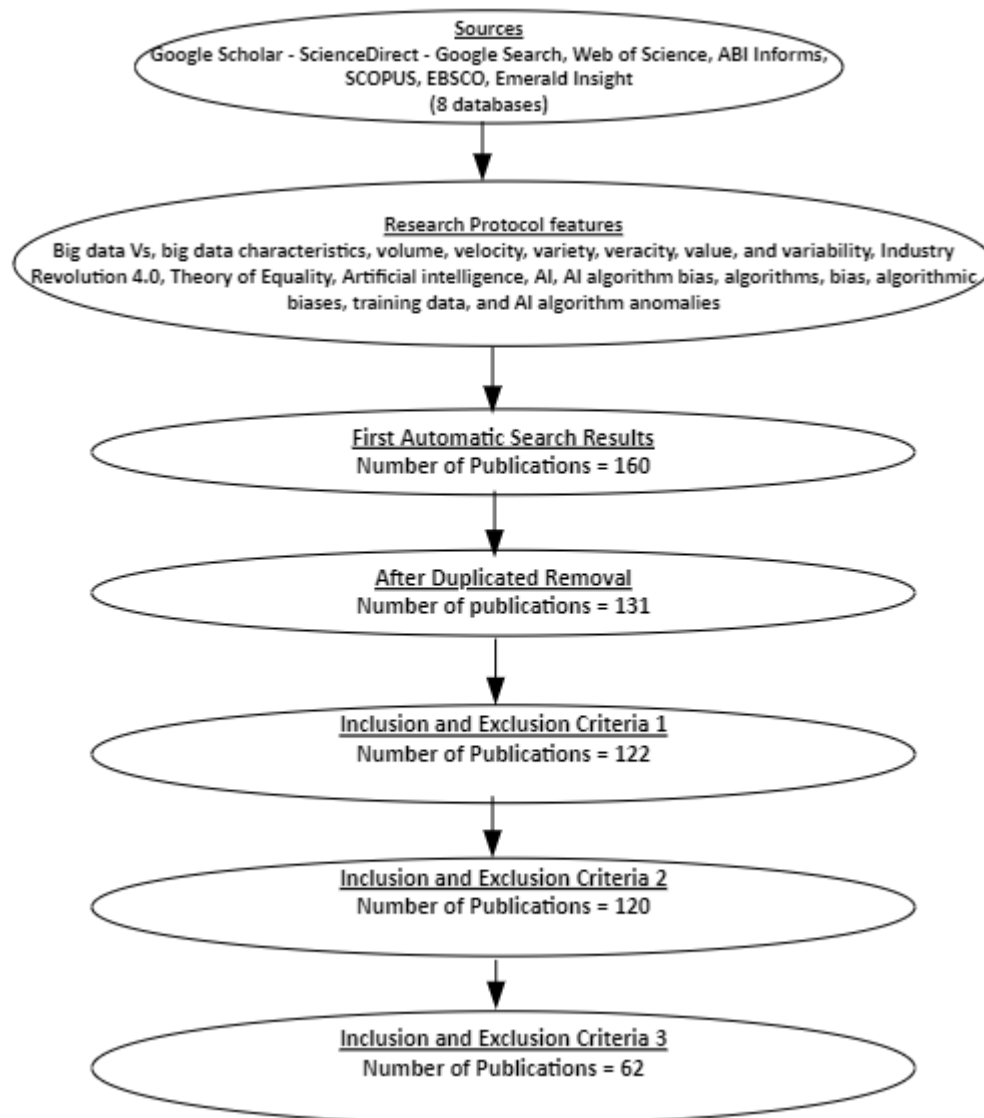
The thematic analysis covered peer-reviewed, SCOPUS-indexed, ABDC (Australian Business Dean Council) journals. More importantly, the thematic keyword search was done in social media using ScienceDirect, Google Scholar, Google Search, SCOPUS, SCImago, and EBSCO. We searched the databases for articles that mentioned the terms and concepts related to big data Vs or big data characteristics, artificial intelligence, training data, theory of equality, and AI algorithm bias in the titles or abstracts. Additionally, the search focused on publications associated with AI, data science, data analytics, and AI algorithms. Moreover, a keyword-based search was used, consistent with Samad et al. (2022). Search keywords include big data volume, velocity, variety, veracity, value, variability, bias, algorithmic bias detection, training data bias, and AI algorithm anomaly prevention. Furthermore, the interpretive qualitative method was used to complete the first portion of the systematic literature review approach. The authors of this paper undertook this part. Papers published between 2016 and 2025 were selected. The selection of the 2016 to 2025 range was used because AI is relatively new.

Initially, 160 publications were retrieved. After removing 29 duplications and redundancies, 131 articles remained from the databases and other sources. Applying three exclusion/inclusion criteria resulted in a final count of 62 articles, meaning 69 articles were excluded.

## Findings

According to the process flowchart for relevant papers, which consists of 7 steps or phases (see Figure 1), 8 databases were included in the search. About 18 features of the research protocol were used. The first automatic search generated 160 publications. After duplicate removal, we found 131 publications. After applying the first inclusion and exclusion criteria, dealing with having at least one protocol feature in the title or abstract, we found 122 publications. Next, we scanned the papers, focusing on the subtitles,

introductions, and conclusions. This step generated 120 publications. The third and final inclusion and exclusion criteria concerned reading each paper in depth and analyzing its relevance to the objective of this study. Table 1 lists the databases and social media websites visited, the search keywords, and the number of articles found in each outlet.



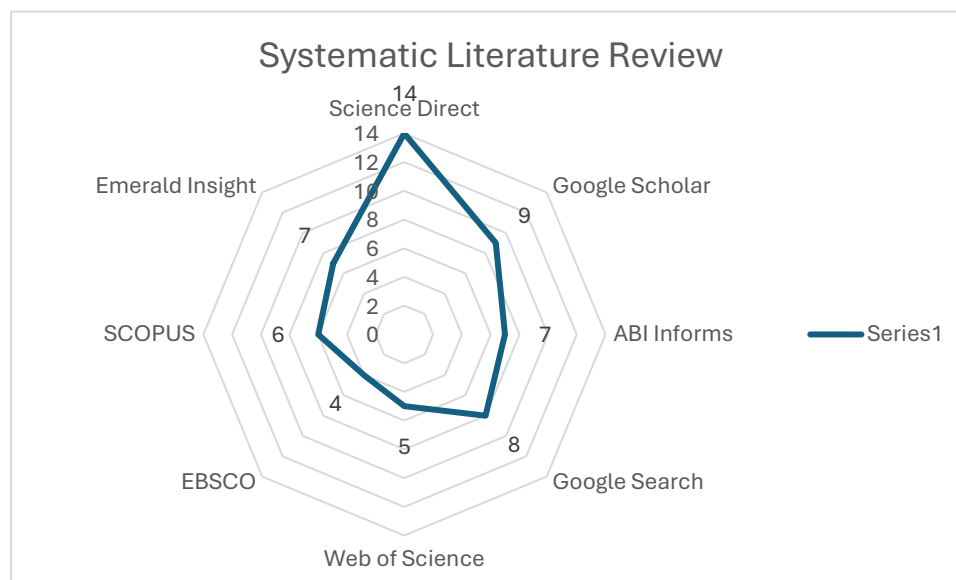
**Figure 1. Flowchart of relevant papers (inclusion and exclusion)**

**Table 1. List of Database and Academic Publication Search Engines**

| Databases and Academic Publication Search Engines | Number of articles | Search Keywords for Big Data                 | Search Keywords for AI algorithm bias     | Search Keywords for the theoretical background |
|---------------------------------------------------|--------------------|----------------------------------------------|-------------------------------------------|------------------------------------------------|
| Science Direct                                    | 14                 | Big data Vs. prominent data characteristics, | Artificial intelligence, AI, AI algorithm | Theory of equality, theory of equity, theory   |
| Google Scholar                                    | 9                  |                                              |                                           |                                                |
| ABI Informs                                       | 7                  |                                              |                                           |                                                |

|                 |    |                                                                                       |                                                                                       |                                  |
|-----------------|----|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|----------------------------------|
| Google Search   | 8  | volume, velocity, variety, veracity, value, variability, and industry revolution 4.0. | bias, algorithms, bias, algorithmic biases, training data, and AI algorithm anomalies | of equality for AI and big data. |
| Web of Science  | 5  |                                                                                       |                                                                                       |                                  |
| EBSCO           | 4  |                                                                                       |                                                                                       |                                  |
| SCOPUS          | 6  |                                                                                       |                                                                                       |                                  |
| Emerald Insight | 7  |                                                                                       |                                                                                       |                                  |
| Total           | 62 |                                                                                       |                                                                                       |                                  |

Figure 2 shows the categories of outlets and the influence and weight of each database or Academic Publication Search Engines. Science Direct generated more publications on AI algorithm biases (14) than any other publication website. Google Scholar showed nine (9) relevant publications in the field of big data Vs or characteristics and AI algorithm biases. Google Search showed eight (8) articles associated with AI algorithm biases and big data. The remaining publication outlets visited in this study showed less than 7.



**Figure 2. Articles distribution per outlet**

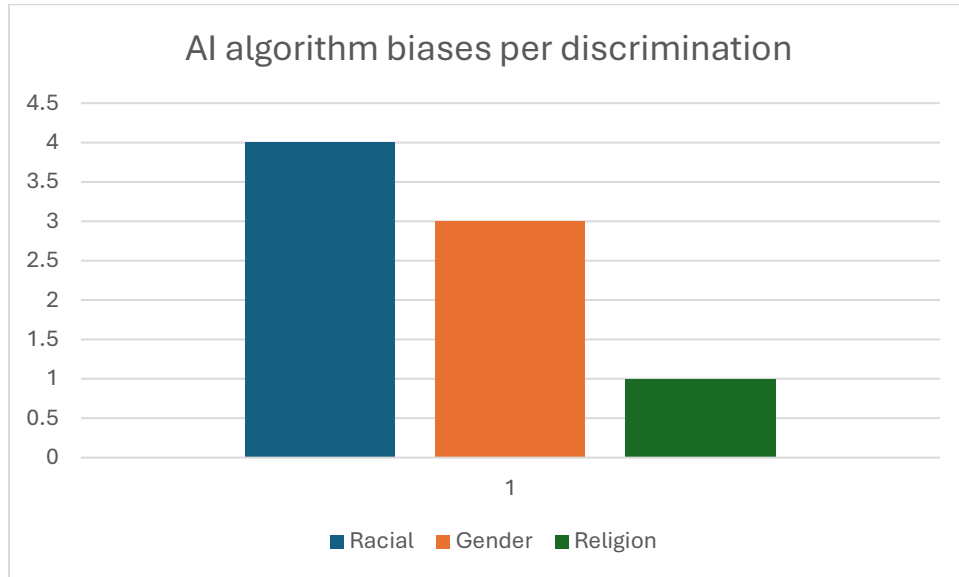
## Findings of AI algorithm biases for discrimination

The content in Table 2 demonstrates that AI algorithm bias is a reality. The results of this thorough literature review indicate the research community's awareness of these societal issues. Also, the research community found many examples of AI algorithm biases. This study presents a sample of AI algorithm biases, including their fields, definitions, and the types of discrimination involved. Most importantly, we found a type of discrimination based on race, gender, religion, age, political affiliation, education, healthcare, and workplace. In terms of fields, we found that AI algorithm biases exist in criminal justice, the health sector, business, politics, and employment. Figure 3 shows that gender and race are the most popular types of AI algorithm discrimination. Also, it shows that religion, age, and political affiliation are less popular types of AI algorithm discrimination.

**Table 2. Sample of Algorithm Biases**

| Fields             | Definitions of Bias                                                                                                                       | Type of discrimination | Sources                   |
|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------|------------------------|---------------------------|
| Criminal Justice   | AI algorithms predict future criminality among African Americans twice that of white people.                                              | Racial                 | Pessach and Shmueli, 2021 |
| Health sector      | AI algorithms enrolled more white than black patients in healthcare programs.                                                             | Racial                 | Pessach and Shmueli, 2021 |
| Business           | AI algorithms charge higher interest rates for loans for black people than white.                                                         | Racial                 | Pessach and Shmueli, 2021 |
| Business           | In the insurance industry, when the system determines premiums based on religion instead of gender                                        | Religion               | (Varsha, 2023).           |
| Business           | In the insurance industry, when the system determines premiums based on gender                                                            | Gender                 | (Varsha, 2023).           |
| Business           | Google's ads target specific groups based on gender profiles                                                                              | Gender                 | (Akter et al., 2023)      |
| Business           | Apple's biased credit card offering to businesses                                                                                         |                        | (Akter et al., 2023).     |
| Work opportunities | On Facebook, customers of African American descent could not view targeted housing, credit, and employment ads.                           | Racial                 | (Akter et al., 2023).     |
| Aesthetic/Mode     | Face recognition AI algorithm systems have higher error rates for darker-skinned people and female subjects than for lighter-skinned men. | Gender                 | (Ferrara, 2024).          |
| Work opportunities | AI algorithm recruiting tools showed gender bias, favoring male candidates over female candidates for technical roles.                    | Gender                 | (Ferrara, 2024).          |
| Politics           | AI algorithm amplifies extremely left and incredibly right political groups more than moderate ones                                       | Political Affiliation  | Cecere et al., (2024).    |
| Work opportunities | AI algorithm fails to prevent discriminatory job posts targeting religion.                                                                | Religion               | Cecere et al., (2024)     |
| Business           | AI algorithm highlights gender disparity in advertising targeting and disadvantaging women.                                               | Gender                 | Cecere et al., (2024)     |

|                    |                                                                                                    |     |                       |
|--------------------|----------------------------------------------------------------------------------------------------|-----|-----------------------|
| Work opportunities | AI algorithm highlights advertising targeting and disadvantaging younger women's job opportunities | Age | Cecere et al., (2024) |
|--------------------|----------------------------------------------------------------------------------------------------|-----|-----------------------|



**Figure 3: The most popular AI algorithm biases**

## Analysis and Discussion

The theory of equality served as an interpretive framework throughout our data analysis. Initially, open-ended coding selected relevant articles from the literature review, but the subsequent analysis stages were guided by the theory. The identified codes were examined through the core principles of the Theory of Equality, helping us understand how the features of big data influence fairness, equitable treatment, access to AI resources, and the distribution of benefits and challenges related to generative AI. This process refined our themes and moved the analysis beyond mere description. Additionally, each major theme is now explicitly linked to aspects of the Theory of Equality. The revised findings highlight how big data characteristics either support or challenge theoretical expectations about equality, offering a clearer conceptual understanding. We also expanded the discussion to demonstrate how the findings extend and contextualize the Theory of Equality within the emerging field of generative AI, rather than just citing it as background.

## Addressing AI algorithm biases using big data volume

AI algorithms work with substantially larger data sizes, types, and volumes. Authors refer to these data as big data volume, encompassing quantitative and qualitative data types such as text and images (Marler, 2024). Because some algorithms update data fields, they are considered transactional algorithms. Some other algorithms perform more complex tasks, like learning from data to identify patterns and make predictions. We suggest sample data for the data-processing mechanism for AI algorithm training. Training data for deep learning is less subject to AI algorithm bias than training data for machine learning. Machine learning can be limited by the amount of data, whereas deep learning can handle unlimited data. Once the

machine learning algorithm is trained on limited data, adding new data for further training is impossible due to its characteristics. If data for a specific community is not available when supervised and unsupervised machine learning algorithms are trained, it would be difficult to add more data later due to data limitations. Data used to train deep learning algorithms can be augmented to correct biases. For instance, if the first category of trained data is biased, the additional data can be corrected by including data representing the underrepresented population in the second data group. That flexibility allows deep learning to minimize the risk of bias compared to machine learning algorithms.

Some researchers' attitudes toward the size of big data vary widely. In this sense, some scholars argue that big data's large sample size enables them to answer any questions regarding real-world phenomena (Lyu et al., 2022). Other researchers suggest that a large sample size sometimes provides valid information (Lyu et al., 2022). Big data's large sample size can reduce AI algorithm bias more than a small sample size can. A large sample size can represent more underrepresented minorities or populations than a small sample size.

Big data volume represents data size (Mishra & Singh, 2020). Data should be voluminous enough to fix, prevent, or minimize AI algorithm bias at the level of the training data. Given this, AI algorithm-trained data must consider a sufficient number of countries, regions, ethnicities, genders, and religions to justify the volume requirements. Mishra and Singh (2020) added that big data volume must include sufficient products and periods to meet the volume condition and criteria for justifying bias-free training data for AI algorithms.

### **Addressing AI algorithm biases using big data variety**

Variety encompasses multiple data types, formats, and sources. When considering these three facets of big data variety, the AI-trained data must represent data from multiple countries, groups of people, organizations, genders, races, religions, and cultures. Mishra and Singh (2020) added that big data variety must cover multiple periods with varying demand, multiple products, and multiple countries. AI algorithm training data must be drawn from many sources. These sources can include electronic records, sensor-generated data, mobile data, reports, wearable fitness data, social media data, administrative data, etc (Hassan et al., 2021). AI training data must include structured, semi-structured, and unstructured data to reduce AI biases. Combining different big data sources is a fruitful approach for identifying the data's strengths and weaknesses and understanding possible biases in AI algorithms (Schmucker & Reif, 2022).

Uber, 23andMe, and Netflix use big data to manage their business operations, so all communities use their services (Johnson et al., 2017). Diversifying the training data by collecting large-scale data from these three companies expands coverage, enabling communities to build training sets used to train and test AI algorithms.

### **Addressing AI algorithm biases using big data velocity**

In the training data step of AI algorithms, developers must assume that data does not exist all at once but rather arrives sequentially in large chunks from a data stream (Lee et al., 2020). The algorithm must be trained on data, given that new data will arrive quickly. With the advent of big data velocity, static training data alone is no longer sufficient. Continuous training data is the only way to keep the AI algorithm up to date; otherwise, it will produce biases. As data keep coming in, new data must be continuously used to train the algorithms. The training data will always provide insights into new needs, development, behavior, and societal attitudes under these conditions. Data will go nowhere soon. The more society grows, the more data will be produced and consumed. There is no point in training AI algorithms to wait or stop the training process of existing or current algorithms. Continuous training of data is recommended in this study for the early detection and prevention of AI algorithm biases.

In today's business environment, technology, competitors, regulatory events, and significant changes in economic and political conditions (Lin & Kunnathur, 2019) generate new behaviors and attitudes in society. Training data must continuously account for these changes to avoid AI algorithm biases.

The sensing power of big data enables it to perceive turbulent environments, helping the community develop new capacities in response.

### **Addressing AI algorithm biases using big data veracity**

Veracity is the need to process big data accurately (Johnson et al., 2017). When AI algorithm training data does not represent all communities or is not sufficiently representative of the world population, it is not accurate. The AI algorithm must improve the accuracy and precision of the training data. Big data may be correct. Still, when it is not well representative of the population for which the AI algorithm is designed, that accuracy is invalid. Therefore, more than big data accuracy is needed to guarantee free algorithm bias. In general, data can be accurate but not helpful. Sometimes, data must be in the proper context to make it useful. For instance, AI training data may be correct, but when the AI algorithm produces bias, there is a mismatch between accuracy and context. Training data for AI algorithms that generate bias is unsuitable. The AI algorithm is supposed to help make decisions. When the decision is wrong, the training data is not appropriate. The truthfulness characteristics of big data determine trust in organizations and communities.

The accuracy of the big data provided to the AI algorithm builder for testing needs improvement. We recommend another run of accuracy testing before using the data for training. Only accurate training data may lead to the right direction and decisions, and contribute to AI algorithm biases.

### **Addressing AI algorithm biases using big data variability**

The changing nature and inconsistency of big data must be considered when training AI algorithms. Data used to train algorithms change over time because people's needs are dynamic. For instance, after five years, most YouTube videos must be updated. Automatically, the periodic peek at those videos also drops tremendously. The AI algorithms trained on data collected from these videos must be updated; otherwise, they will produce results that run counter to society's expectations. When the data used to train algorithms changes over time, the AI algorithm does not track these trends and produces anomalies and biases. Historical data must be used with caution to train data. We posit that historical data has changed over time, so training an AI algorithm on historical data violates the characteristics of big data variability. Also, non-volatile data is no longer appropriate for the community because people's needs are changing quickly. Old data can become obsolete and no longer be valid; therefore, using it as training data may introduce AI algorithm biases.

Variability is the data's tendency to change inaccurately or in near real time (Mishra & Singh, 2020). It represents data fluctuation with time. AI algorithm training data must account for fluctuations and changes over time.

Old data cannot be used to train AI algorithms intended for the young generation. The generation gap can become a massive issue in the training data. Most recent training data should represent the current generation's needs and, therefore, be less subject to bias.

### **Addressing AI algorithm biases using big data value**

The value extracted from the data may be used for training. In other words, data extracted from data becomes essential for early detection and prevention of AI algorithm biases. The value of the data extracted must reveal hidden patterns, truths, and knowledge. Otherwise, the valuable data used to train the AI algorithm will produce biases. Data that cannot deliver value is not appropriate for training data. There is no point in using data that cannot create value for the users to train AI algorithms.

## Conclusion

Artificial intelligence (AI) is the driving force behind the technologies of the fourth industrial revolution. This study aimed to use big data Vs or characteristics intelligently to detect and prevent AI algorithm biases. This would allow AI developers to minimize and eliminate discrimination, unfairness, and unethical behavior arising from the use of AI in society.

Following the qualitative procedures of Samad et al. (2022) and Varsha et al. (2023), a systematic literature review based on the AI algorithm's technical and thematic evaluation was conducted. Interpretive analysis was accomplished using articles from 8 databases and social media research outlets.

The main findings include sixty-two (62) publications representing distinct factors: big data volume, veracity, velocity, variety, variability, and value on one hand, AI, AI algorithms, AI algorithm biases, and the theory of equality on the other.

Additionally, this paper provides a holistic conclusion, transforming big data Vs or characteristics into techniques for predicting AI algorithm biases.

Moreover, it presents valuable tools that public leaders and governments can use as a starting point for proposing regulations to address biases in AI algorithms.

The study also extends the theory of equality and explores the relationships between big data Vs or characteristics and AI algorithms. This study uses big data Vs or characteristics in data training to minimize AI algorithm bias. In other words, biases in training data may be caused by measurement and AI developers' interpretation of big data Vs or characteristics. Biases may also arise from missing data and misunderstanding, resulting in datasets that lack some characteristics of big data. Our study responds to calls for big data Vs or characteristics in designing training data for AI algorithms.

The results are as follows. First, large data volume and a large sample size can reduce AI algorithm bias more than a small sample size. A large sample size can represent more underrepresented minorities or populations than a small sample size. Second, the variety of big data allows for diversifying the training data by collecting data in different formats and types from other sources. This extends coverage so that communities can collect training data for training and testing AI algorithms. Third, big data velocity enables continuous use of training data for AI algorithms, as new data arrives quickly. Fourth, big data veracity allows for truthfulness and accurate training data for AI algorithms. Fifth, big data variability enables training data to change and be volatile over time for AI algorithms. Sixth, big data provides valuable training data that benefits AI algorithms.

### Theoretical Contributions

This study makes significant theoretical contributions to the literature. First, it expands the existing body of research on AI algorithm biases. This research adds a systematic literature review methodology to the list of AI algorithm bias detection and prevention studies. Second, in terms of theoretical implications, this research advances the foundations of the theory of equality in AI. This study is among the first to use the theory of equality as the theoretical framework for research on AI algorithm bias. Therefore, it extends the theory of equality. Third, this research theoretically compares AI algorithm training and big data.

### Practical Contributions

The findings of this study are valuable for the public and private sectors. The public sector will benefit from the results of this research because it is an area where large amounts of data are collected, generated, and consumed. Therefore, the training data of the AI algorithm can meet the conditions of big data volume.

From a private-sector perspective, this research has the merit of reducing AI algorithm biases, as the private sector has more resources than the public sector. Additionally, the private sector must survive to meet the competitive advantage objective in that sector. The results of this research can help private businesses gain strategic advantages by leveraging big data characteristics to prevent and eliminate AI algorithm biases. Moreover, both sectors will know from this study which of the big data Vs they must prioritize to combat AI-algorithm discrimination.

## Limitations and Future Research Avenues

This study suggests several perspectives for follow-up research. First, future research avenues can study legal regulations to prevent and detect early AI algorithm biases. Governments can appoint an AI authority to oversee the algorithms behind all AI applications and technologies. It will be a challenging task. Governments can hire a team of AI experts who will issue permits to IT companies to create and publish AI algorithms for society. That team will test the AI algorithms and ensure there are no biases before allowing them to be used by the population. AI language biases have become alarming in many communities because most NLP algorithms have been trained solely on English. Future research can explore using additional international languages to test NLP algorithms and reduce bias in the field. Culture is another concern because many people in developing countries use AI algorithms mainly designed for Western culture. Future research can examine opportunities to expand the AI algorithm-trained data to include cultures in developing countries.

## References

- Acheampong, A., Opoku, E., Amankwaa, A., Dzator, J. (2024). Energy poverty and gender inequality: Unpacking the transmission channels. In education, *Technological Forecasting and Social Change*, 202, Article e123274. <https://doi.org/10.1016/j.techfore.2024.123274>
- Ahmed R, Shaheen S, Philbin S. The role of big data analytics and decision-making in achieving project success. *Journal of Engineering and Technology Management*. 2022;65: Article e101697. <https://doi.org/10.1016/j.jengtecman.2022.101697>
- Akter, S., Sultana, S., Mariani, M., Wamba, S. Spanaki, K. & Dwevedi. Y. (2023). Advanced algorithm bias management capabilities in AI-driven marketing analytics research, *Industrial Marketing Management*, 114, pp. 243–261. <https://doi.org/10.1016/j.indmarman.2023.08.013>
- Alter, S. (2022). Understanding artificial intelligence in the context of usage: contributions and smartness algorithmic capabilities in work systems, *International Journal of Information Management*, 67. Article e102392. <https://doi.org/10.1016/j.ijinfomgt.2021.102392>
- Burton, J. (2023). Algorithm Extremism? The securitization of artificial intelligence (AI) and its impact on radicalism, polarization, and political violence, *Technology in Society*, 75, Article e102262. <https://doi.org/10.1016/j.techsoc.2023.102262>
- Cecere, G., Jean, C., Le Guel, F. & Manant, M. (2024). Artificial Intelligence and Algorithm Bias? Field tests on social networks with teens. *Technological Forecasting & Social Change*, 201, Article e123204. <https://doi.org/10.1016/j.techfore.2023.123204>
- Chen, J., Li, Z., Wang, Y., & Pang, S. (2026). Unpacking the impact of Big Data on business performance: Do open innovation and digital technology adoption matter? *Technology in Society*. Vol. 86. <https://doi.org/10.1016/j.techsoc.2026.103285>

- Chiarella, S., Torromino, G., Gagliardi, D., Rossi, D., Babiloni, F. & Cartocci, G. (2022). Investigating the negative bias towards artificial intelligence: effects of prior assignment of AI- authorship on the aesthetic appreciation of abstract paintings, *Computer in Human Behavior*, 137, Article e 107406. <https://doi.org/10.1016/j.chb.2022.107406>
- Chou, Y., Moreira, C., Bruza, P., Ouyang, C. & Jorge, J. (2022). Counterfactuals and causality in explainable artificial intelligence: theory, algorithms, and applications, *Information Fusion*, 81, pp. 59-83. <https://doi.org/10.1016/j.inffus.2021.11.003>
- Delecraz, S., Eltarr, L., Becuwe, M., Bouxin, H., Boutin, N. & Oullier, O. (2022). Responsible Artificial Intelligence in Human Resources Technology: An Innovative Inclusive, and Fair Design Matching Algorithm for job recruitment purposes. *Journal of Responsible Technology*, 11, Article e 100041. <https://doi.org/10.1016/j.jrt.2022.100041>
- Ferrara, E. (2024). The butterfly effect in artificial intelligence systems: Implications for AI bias and fairness, *Machine Learning with Applications*, 15, Article e100525. <https://doi.org/10.1016/j.mlwa.2024.100525>
- Fountain, J. (2022). The moon, the ghetto, and artificial intelligence: reducing systematic racism in computational algorithms, *Government Information Quarterly*, 39, Article e101645. <https://doi.org/10.1016/j.giq.2021.101645>
- Ghasemaghaei M. Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics. *International Journal of Information Management*. 2021;57: Article e102055. doi: <https://doi.org/10.1016/j.ijinfomgt.2019.102055>
- Ghasemaghaei M, Calic G. Assessing the Impact of Big Data on Firm Innovation Performance: Big Data is not always Better Data. *Journal of Business Research*. 2020;108:147-162. doi: <https://doi.org/10.1016/j.jbusres.2019.09.062>
- Ghasemaghaei M, Calic G. Does Big Data Enhance Firm Innovation Competency? The Mediating Role of Data-Driven Insights. *Journal of Business Research*. 2019b;104:69-84. doi: <https://doi.org/10.1016/j.jbusres.2019.07.006>
- Gonzales-Sendino, R., Serano, E. & Bajo, J. (2024). Mitigating bias in artificial intelligence: fair data generation via causal models for transparent and explainable decision-making. *Future Generation Computer Systems*, 155. pp. 384-401. <https://doi.org/10.1016/j.future.2024.02.023>
- Jia, S., Ma, C., Chi, O. H., & Fan, A. (2026). A socio-technical exploration of bias in generative AI travel planners. *Tourism Management*. Vol. 116. <https://doi.org/10.1016/j.tourman.2026.105442>
- Johnson J, Friend S, Lee H. Big Data Facilitation, Utilization, and Monetization: Exploring the 3Vs in a New Product Development Process. *Journal of Production Innovation Management*. 2017; 34(5):640-658. [doi:10.1111/jpim.12397](https://doi.org/10.1111/jpim.12397)
- Kim, W. & Ryoo, Y. (2026). The literacy paradox: How AI literacy amplifies biases in evaluating AI-generated news articles. *International Journal of Information Management*. Vol. 86.

- Li, Y. (2022). Research on neural network algorithm in artificial intelligence recognition, *Sustainable Energy Technologies and Assessments*, 53, Article e102691. <https://doi.org/10.1016/j.seta.2022.102691>
- Liu, C., Yang, S., Hao, T. & Song, R. (2022). Service risk of energy industry international trade supply-chain based on artificial intelligence algorithm, *Energy Reports*, 8, pp. 13211-13219. <https://doi.org/10.1016/j.egy.2022.09.182>
- Marler, J. (2024). artificial intelligence, algorithms, compensation strategy: challenges and opportunities, *Organizational Dynamics*, 53, Article e100039. <https://doi.org/10.1016/j.orgdyn.2024.101039>
- Peng, X., Fu, Y. & Zu, X. (2024). Gender equality and green development: a qualitative survey. *Innovation and green development*, 3, Article e100089. <https://doi.org/10.1016/j.igd.2023.100089>
- Pessach, D. & Shmueli, E. (2021). Improving Fairness of Artificial Intelligence Algorithms in Privilege-group Selection Bias Data Settings. *Expert Systems with Applications*, 185, Article e115667. <https://doi.org/10.1016/j.eswa.2021.115667>
- Rodgers, W., Cardenas, J., Gemoets, L. & Sarfi, R. (2023). A smart grid knowledge transfer paradigm supported by experts' throughput modeling artificial intelligence algorithm processes, *Technological Forecasting & Social Change*, 190, Article e123373. <https://doi.org/10.1016/j.techfore.2023.122373>
- Rodgers, W., Murray, J., Stefanidis, A., Degbey, W. & Tarba, S. (2023). An artificial intelligence algorithm approach to ethical decision-making in human resource management processes. *Human Resource Management Review*, 33, Article e100925. <https://doi.org/10.1016/j.hrmr.2022.100925>
- Rosca, C., Stancu, A. & Ariciu, A. (2024). Algorithm for child adoption process using artificial intelligence and monitoring system for children, *Internet of Things*, 26, Article e 101170, <https://doi.org/10.1016/j.iot.2024.101170>
- Sheng, Z., Yilin, W., Jintao, W., & Keli, X. (2026). AI-driven big data acquisition and optimization in cold chain logistics using energy-aware IoT systems. *Sustainable Computing: Informatics and Systems*. Vol. 50. <https://doi.org/10.1016/j.suscom.2026.101362>
- Varsha, P. S. (2023). How can we manage biases in Artificial Intelligence Systems – A systematic literature review. *International Journal of Information Management Data Insights*, 3, Article e100165. <https://doi.org/10.1016/j.ijime.2023.100165>
- Wood, E. X. (2025). AI and big data in disaster response: Ethical and practical challenges. *Journal of Dynamic Disasters*. Vol. 1. <https://doi.org/10.1016/j.jdd.2025.100041>
- Zhang D, Pee LG, Pan S, Cui L. Big data analytics, resource orchestration, and digital sustainability: A case study of smart city development. *Government Information Quarterly*. 2022;39: Article e101626. doi: <https://doi.org/10.1016/j.giq.2021.101626>