ISSN 2616-5783 Vol.8, Issue 11: 195-200, DOI: 10.25236/AJHSS.2025.081128

The Formation Mechanism and Pluralistic Governance of Algorithmic Gender Discrimination in the AI Era

Yating Wang[1,*]

[1]Faculty of Humanities and Arts, Macau University of Science & Technology, Macau, 999078, China
**Corresponding author: 775081569@qq.com*

Abstract: *With the widespread application of algorithms and big data, the issue of algorithmic gender discrimination has become increasingly prominent, exacerbating social inequality and hindering the achievement of gender equality goals. This study adopts a dual perspective of technology and society to systematically analyze the endogenous risks and social roots of algorithmic gender discrimination. It proposes a diversified governance framework encompassing data collection, algorithm design, technological innovation, ethical norms, legal refinement, and social supervision to mitigate the risks of algorithmic discrimination. The research aims to promote the healthy development of AI technology, foster a fair (non-discriminatory), just (ethical), and transparent (explainable) algorithmic ecosystem, and provide theoretical foundations and practical pathways for gender equality in digital spaces.*

Keywords: *Algorithmic gender discrimination; Algorithmic fairness; AI; Technological innovation*

1. Introduction

In today's era, artificial intelligence (AI) is developing rapidly as a revolutionary technology, profoundly changing the world's operating rules. Algorithm and big data have become the core of information production and dissemination, clearly showcasing data's four core characteristics—often referred to as the "Four Vs": Volume, Variety, Velocity, and Value [1]. This technological progress constantly reconstructs the human living environment and challenges the definition of the essence of "human. "From the industrial revolution to the era of algorithm, technology has always presented dialectical characteristics: on the one hand, it promotes the development of civilization; on the other hand, its alienated use aggravates social risks. As W. Brian Arthur noted, "Contemporary technology has inevitably become a source of dependence and anxiety for humanity, and this paradox is profoundly reshaping the power structures and cognitive paradigms of human society." [2] However, everyone is the purpose itself, and everyone should be respected equally. [3] As a new form of social power, the algorithm exhibits characteristics of simplification in its development, processing, and application, often overlooking the diversity of gender expression. Consequently, this phenomenon gives rise to an "unconsciousness" that remains largely unexplained, which is subsequently transformed into a manifest "consciousness" upon the generation of the algorithm.

A review of extant research in academic circles reveals that the relationship between technological progress and gender equality manifests dual characteristics: historical inheritance and contemporary complexity. According to the analytical framework of "society, technology, gender," this connection is deeply rooted in a specific cultural context. In the AI era, algorithmic discrimination systematically reproduces the deep inequality in the social power structure more subtly, posing a new threat to gender equality. [4] Bruce Bimber found that in the process of using technology, the embodiment of gender differences and the male values inherent in technology often favor men's development in the digital space. [5] We aim to create an AI that prioritizes human well-being, ensuring that advancements in algorithms and technology promote the greater good of humanity. Based on this concept, this study aims to analyze the formation mechanism of algorithmic gender discrimination and propose collaborative governance strategies for these issues. It aims to improve the "coded gaze" in the context of the rapidly evolving algorithm propagation in the AI era, which is also a key issue that needs to be addressed urgently.

ISSN 2616-5783 Vol.8, Issue 11: 195-200, DOI: 10.25236/AJHSS.2025.081128

2. Conceptual Framework of Algorithmic Gender Discrimination: From Algorithmic Decision Making to Discrimination Characteristics

The algorithm is a concentrated expression of the computing system's capabilities and potential. It is a concise and clear operating instruction intended to solve a specific problem. [6] It is often used in computing, data processing, and automatic reasoning. At present, the algorithm has become an iconic cultural symbol, which benefits from the wide application of the concept of "algorithm" and the role of the algorithm itself. In other words, the influence of the algorithm is not only reflected in its functions, but also in its universality. [7] Personalized recommendation of an algorithm is widely used in judicial, military, financial, and medical fields. From existing research at home and abroad, it is evident that research on algorithm risks exhibits characteristics of interdisciplinary complexity, significant artificiality, and uncertainty. [8] With the development of AI technology, algorithmic decision-making has evolved from assisting human decision-making to fully replacing it in some fields. For example, the application of the algorithm is particularly significant in talent recruitment, job screening, credit approval, and crime risk assessment. Compared with human autonomous decision-making, algorithmic decision-making is widely used because of its advantages, such as its fast, accurate, objective, universal, and low-consumption nature. However, these algorithm-based decisions are susceptible to inaccurate or discriminatory outcomes, as well as potential discrimination.

This paper systematically summarizes the existing literature and finds that the concepts of "algorithmic bias" and "algorithmic discrimination" are mixed up in current academic circles. The fuzziness not only affects the differential analysis between them but also weakens the accuracy and validity of related research. Therefore, this paper systematically defines and distinguishes the core concepts to establish a clear research framework. "Algorithmic discrimination" is a standardized concept that mainly involves direct or indirect unfair prediction and decision-making in data processing. It also involves rules, standards, and behaviors that seem neutral on the surface, but actually have a different effect on protected individuals or groups. [9] There are various forms of algorithmic discrimination, including but not limited to algorithmic gender discrimination, algorithmic racial discrimination, algorithmic religious or belief discrimination, etc. Algorithm gender discrimination is a typical embodiment of algorithmic discrimination, which mainly involves the differential treatment and consequences based on "gender characteristics" in the process of algorithm development, training, data collection, processing, and automatic decision-making by applying data. [10] It can be concluded that algorithm gender discrimination is a social phenomenon with new characteristics. Its uniqueness lies in the following aspects: First, its manifestation is not limited to a single gender. Even if gender is not directly involved, discrimination may still occur through data association. Second, it presents a multi-intersection, which is often associated with factors such as race and age. In comparison with discrimination experienced in daily life, it is characterized by its systematic nature, covert nature, and extensive scope. Moreover, the academic community's cognition of the issue has evolved. Initially, the focus was on female discrimination. Subsequently, attention shifted to male reverse discrimination. Finally, the scholarly discourse expanded to encompass transgender groups. [11] In conclusion, algorithmic gender discrimination, which is both hidden and widespread, has significantly undermined the concept of gender equality and has negatively affected the social atmosphere.

3. Analysis of the Causes of Algorithmic Gender Discrimination: Technology and Society

3.1 Technical Factors: From Data Defects to Algorithm Enhancement

The algorithm is based on the digital presentation of massive amounts of information, utilizing big data, which is accurate, and emphasizes the relevance of content rather than causality. [12] Simultaneously, the operation process of the algorithm is presented as a "system that cannot be accessed or directly observed from the outside." [13]. Its core mechanism is that the algorithm designer transforms data input into information output through opaque programs, thus forming an "algorithm black box". In this process, the intricate and ever-evolving world is often "simplified," resulting in an oversimplification that disregards the intricacies and complexity of the world itself. It also means that there is a discriminatory tendency behind this indiscriminate algorithmic generalization, or "selective presentation". In the process of algorithm statistical analysis, only a subset of specially screened samples is calculated, while the whole process of sample screening is ignored. This approach may lead to the omission of key factors in statistical analysis, thereby affecting the accuracy of the research results and potentially leading to deviations in the research findings. [14] Ultimately, what we see is only the "survivor bias".

ISSN 2616-5783 Vol.8, Issue 11: 195-200, DOI: 10.25236/AJHSS.2025.081128

The subversive nature of the data age is that the attributes and laws of all things can be transmitted to another homogeneous thing through "coding" to realize "intact" holographic expression. [15] Moreover, the algorithm is the projection of human thinking, which requires value judgment, not just logical judgment. Its risks lie in the inheritance of social prejudices, as well as the cultural or social confinement that these prejudices may cause. [16] This is the product of the co-shaping of human cognitive structure and social form. Algorithm, as a concrete expression of human thinking, is showing its potential defect-discrimination in an out-of-control way. [17] Because AI integrates the inherent prejudice and discrimination of human nature into the data in the process of algorithm generation and processing, the "unconscious" behavior of algorithm design developers becomes a "conscious" algorithm in the process of data processing. In addition to these explicit and easily recognizable discriminations, there are still some unconscious discrimination phenomena, which are subtly inherited and amplified by the algorithm. Additionally, the algorithm reinforces the fixed and simplified ideas held by individuals or groups. When it processes data that contains stereotypes, it will continue to strengthen these stereotypes throughout its learning and training process. They can worsen discrimination, which undermines the diversity and inclusiveness of society.

3.2 Social Structural Constraints: Lagging System and Cognitive Dilemma

Given the widespread utilization of algorithmic technology, traditional normative systems intended to prevent discrimination are facing unparalleled challenges. Traditional discrimination is strictly restricted by law, and its restraint mechanism is based on the principle of prohibiting discrimination and follows the logical chain of "violating obligations-bearing consequences". In addition, some scholars summarize it as two core elements: "information constraint" and "responsibility mechanism". Algorithms are produced in the social environment and coexist with it, which constitutes an indispensable part of society and has long been integrated into practice and results. [18] The rapid development of algorithms makes the current anti-discrimination legal system face a double dilemma: On the one hand, the rapid pace of technological iteration often outpaces the frequency of legal updates, resulting in a lag in regulation. On the other hand, because of the complexity and opacity of algorithmic decision-making, it is difficult to identify and prove discriminatory behavior. The lagging system and technical characteristics together constitute the structural conditions for the breeding of algorithmic gender discrimination. The algorithm system essentially embodies social power relationships, and its design and application often serve as a technical means to consolidate the existing ruling order. The invention and use of technology are essential to meet the needs of human society and political ideas. Technology is also "political" and may become an inevitable means to maintain power. [19] Therefore, other factors will interfere with the algorithm and affect its data generation, processing, and application.

Furthermore, while big data offers valuable insights into online behaviors and surface-level outcomes, it falls short of capturing the complexity of real-world people. The digital footprint leaves no trace of a person’s inner world—their private thoughts, emotions, beliefs, and imaginative spark remain beyond its grasp. [20] Especially in the aspect of gender presentation, the expression forms show significant diversity. Since its birth, AI has always taken "simulating human intelligence" as its ultimate goal. The hidden gender discrimination present in algorithms is an extension of societal biases, revealing a complex cognitive dilemma and deep-seated gender inequality. This discrimination is woven into every aspect of the algorithm's life cycle. To effectively address the hidden and recurring problem of algorithmic gender discrimination, we must implement comprehensive and systematic governance strategies that are diverse in nature. This approach will ensure that every individual retains their value and dignity as human beings while interacting with technology, and it will help to promote gender equality in the era of AI.

4. Pluralistic Governance of Algorithmic Gender Discrimination

The development of AI should prioritize the promotion of human welfare as its primary goal. Aiming at the key problem of algorithmic gender discrimination, this paper puts forward a comprehensive governance framework based on data level, design and development, technological innovation, laws and regulations, and social supervision (see Figure 1), aiming at preventing it from the source and promoting the development of AI in a more fair and inclusive direction.

ISSN 2616-5783 Vol.8, Issue 11: 195-200, DOI: 10.25236/AJHSS.2025.081128

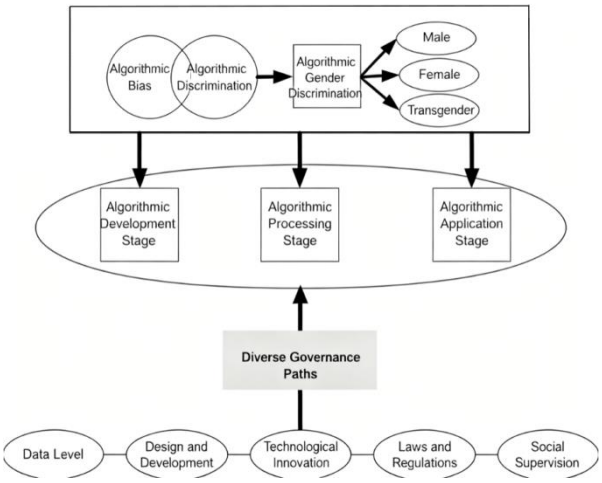

Fig.1 Pluralistic governance of algorithmic gender discrimination

4.1 For Data: Comprehensiveness of Data Collection and Fairness of Auditing and Cleaning Data

The algorithm's operation is predicated on the existence of sufficient data, underscoring the necessity for comprehensive data collection practices. In data collection, all relevant variables should be covered as comprehensively as possible to prevent the omission of information that may significantly impact the algorithm results. Moreover, the timeliness of data is also a key factor, so we should ensure that the data used truly reflects the social situation and gender characteristics. During the data auditing and cleaning process, fairness is also a principle that must be adhered to. First, it is necessary to pay attention to the sources and collection methods of data to avoid data being manipulated or tampered with by specific interest groups. Second, it is necessary to exclude those discriminatory or misleading data to ensure that the algorithm is not interfered with by Dirty Read in the training process. Third, strictly protect the data containing personal privacy and sensitive information to ensure its safety and compliance. It is necessary to establish a data quality monitoring mechanism, regularly check and update it, and ensure the accuracy and reliability of data. Additionally, the audited and cleaned training data can effectively eliminate potential discrimination factors and provide high-quality data for the algorithm, thus reducing the risk of gender discrimination.

4.2 Design and Development: Optimize the Algorithm Design Team and Establish a Feedback Mechanism for Discrimination Detection

In the process of generating algorithms, designers with specific professional competencies have developed technical standards for algorithms that encompass technical, social, and commensurability aspects. These standards facilitate and guide the systematic and sustained advancement of algorithmic fairness technology. [21] During the operation of the algorithm, it is normal for the technology to be redefined and modified once it is separated from the designer. Studies have shown that gender discrimination within algorithms is particularly pronounced in practical applications. Therefore, it is particularly important to implement ethical review at the design stage and establish an efficient algorithm feedback mechanism. Because the lack of feedback may lead to errors being ignored or omitted, we deal with the problem according to the established internal logic and determine whether the output result is accurate according to the standards. The continuous evolution and innovation of technology can effectively alleviate various problems, including algorithmic gender discrimination, to a certain extent. Many scholars, technical experts, and industry leaders are actively engaged in the research and development of algorithm design optimization and risk prevention. Their shared objective is to address gender discrimination stemming from reality through technological advancement.

4.3 Technological Innovation: Create a Gender-Equitable Algorithmic Model to Enhance the "Gender Sensitivity" of the Algorithm

Technological development frequently adheres to a cyclical logic characterized by self-reinforcement. The introduction of new technologies is intended to solve existing problems, but it often leads to new challenges, which in turn create a demand for more advanced technologies. Therefore, to truly achieve

ISSN 2616-5783 Vol.8, Issue 11: 195-200, DOI: 10.25236/AJHSS.2025.081128

the goal of eliminating discrimination, we must start from the source of technological innovation. For the specific practice, in addition to the use of synthetic data sets, such as generative adversarial networks (GAN), synthetic minority oversampling technology (SMOTE), or adaptive synthetic sampling (ADASSYN) (the extension of SMOTE). [22] More importantly, it is necessary to establish a "gender-equitable algorithmic model", which systematically eliminates gender bias in traditional algorithms and ensures that the whole process of data processing follows the principle of gender neutrality. A truly effective gender-equitable algorithmic model should have dual characteristics: it should not only ensure the fairness of the output results, but also ensure the gender neutrality of the algorithm operation mechanism itself. It requires us to: firstly, deconstruct and analyze the underlying logic of the algorithm to identify and correct the causes of potential gender discrimination. Secondly, it is necessary to introduce the evaluation mechanism for "gender sensitivity." In other words, in the process of algorithm design and optimization, the influence weight of gender factors on data interpretation and decision output is systematically considered. The double guarantee mechanism can greatly enhance the accuracy and fairness of algorithmic decision-making on gender issues.

4.4 Laws and Regulations: The Structural Expansion and Digital Innovation of the Anti-Sex Discrimination Normative System

To effectively address algorithmic gender discrimination, the development of laws and regulations is crucial. We need to constantly improve the existing laws and regulations, clearly define the behavioral boundaries of algorithmic gender discrimination, and formulate corresponding punishment measures, so that the "legal framework" can continuously adapt to the new challenges brought about by technological progress. "Digital innovation" should rely on modern information technology to improve the efficiency and accuracy of law enforcement. For example, the establishment of "online reporting" and "censorship mechanisms" to fight against algorithmic discrimination, and the use of big data analysis to identify potential discrimination patterns. We mainly consider the following three core principles: First, we must prioritize people and safeguard their fundamental rights against technological erosion. Second, we should uphold the principle of human dominance to ensure ultimate control over algorithms. Third, adhere to the principle of technology neutrality and avoid instrumental rationality covering up value rationality. Together, these principles form the basis of governing algorithmic discrimination. In the process of concrete implementation, we should make concerted efforts from all walks of life to ensure the effective implementation of laws and regulations and promote the healthy and sustainable development of technology in the application.

4.5 Social Supervision: Guided by Algorithmic Fairness, Promote the Spread of Gender Equality Awareness

Lately, AI has evolved from academic research to everyday utilities, influencing every sector of society. As the key to promoting social progress, AI has shown remarkable value in improving production efficiency and quality of life. This transformative potential constantly inspires public attention and expectation. In this context, it is urgent to build a multi-party collaborative social participation mechanism. On the one hand, it is necessary to improve the AI literacy of the whole people through diversified channels, including popular science education and public forums. On the other hand, we must vigorously advocate the concept of "algorithm fairness" and promote the scientific popularization of gender equality awareness. Especially in the design and application of algorithms, it is urgent to build a long-term mechanism to ensure social justice and gender equality. In addition, it is necessary to improve the public participation mechanism and encourage social forces to systematically supervise algorithm decision-making. There is a need to develop a closed-loop management system that incorporates problem feedback, deviation detection, and rectification optimization, thereby comprehensively enhancing society's ability to identify algorithmic gender discrimination and increasing awareness of prevention. By adopting a comprehensive and multi-faceted governance approach, we can cultivate an AI development ecosystem that prioritizes algorithmic fairness.

5. Conclusion

This paper analyzes the mechanism of algorithmic gender discrimination, proposes collaborative governance strategies, and seeks to improve the "coded gaze" in the AI era, addressing an urgent issue to ensure that technology serves human well-being. While advancing civilization, technology also poses risks of alienation. Algorithmic simplification overlooks gender diversity, reproducing social power

ISSN 2616-5783 Vol.8, Issue 11: 195-200, DOI: 10.25236/AJHSS.2025.081128

inequalities subtly through systemic discrimination—posing new threats to gender equality. This paper examines the conceptual framework of algorithmic gender discrimination and analyzes its causes. The research findings show that many factors have an impact on gender discrimination.

Additionally, a comprehensive governance framework was proposed in this research, encompassing data level, design and development, technological innovation, laws and regulations, and social supervision. It is recommended that society implement an oversight mechanism for algorithmic decision-making processes. This oversight should encompass establishing a closed-loop management system that incorporates mechanisms for problem feedback and for optimizing deviation correction. It aims to foster the optimal development of AI technology while concurrently offering theoretical frameworks and pragmatic pathways to promote gender equality in digital environments.

References

[1] Zhang Y H, Qin Z G, Xiao L. The discriminatory nature of big data algorithms[J]. Studies in Dialectics of Nature, 2017, 33(5): 81-86.
[2] Brian Arthur. The Nature of Technology[M]. Hangzhou: Zhejiang People's Publishing House, 2018.
[3] Immanuel Kant. Principles of the Metaphysics of Morals[M]. Shanghai: Shanghai People's Publishing House, 2002.
[4] ZOU J, SCHIEBINGER L. AI can be sexist and racist—it's time to make it fair[J]. Nature, 2018, 559(7714): 324-326.
[5] BIMBER B. Measuring the gender gap on the Internet[J]. Social Science Quarterly, 2000, 81(3): 868-876.
[6] Cormen T H, Leiserson C E, Rivest R L, et al. Introduction to algorithms[M]. MIT press, 2022.
[7] David B, Qu H Y. The social power of algorithms[J]. China Communication Review, 2023, 18: 3-13.
[8] Hui Z B., Li J. Public Security Risk Governance in the AI Era[M]. Shanghai: Shanghai Academy of Social Sciences Press, 2021.
[9] Karen, Y., & Martin, L. Taming Algorithms[M]. Translated by Lin, S. W. & Tang, L. Y. Shanghai: Shanghai People's Publishing House, 2020.
[10] Gan T, Ma L. Algorithmic Gender Discrimination: Characteristics, Types and Governance[J]. Journal of Shandong Academy of Governance, 2024(2): 97-105.
[11] Liu X N, Gong X L. An Overview of the Seminar on Theoretical and Practical Issues of Anti-discrimination[J]. Journal of Chinese Women's Studies, 2013(6): 114-117.
[12] Zheng Z F. Beware of the potential discrimination risks of algorithms[N]. Guangming Daily, 2019-06-23(07).
[13] Tao Y C. Knowledge issues in technology - the black box of technology[J]. Science and Technology Association Forum, 2008(7): 54-55.
[14] Song L L. Reflections on Communication Studies under the Survivor Bias Theory[J]. Drama Home, 2015(24): 263.
[15] Li D W. Isomorphism: The mathematical and philosophical foundation of big data[N]. Guangming Daily, 2012.
[16] Peng L. Illusions, Algorithmic Prisoners, and the Transfer of Rights: New Risks in the Era of Data and Algorithms[J]. Journal of Northwest Normal University (Social Sciences), 2018, 55(5): 20-29.
[17] Li C. Legal Governance of AI Discrimination[J]. China Legal Science, 2021(2): 127-147.
[18] David B, Qu H Y. The social power of algorithms[J]. Chinese Communication Studies Review, 2023, 18: 3-13.
[19] Guo X P, Qin Y X. Deconstructing the data myth of intelligent communication: the causes of algorithmic bias and the path of risk governance[J]. Modern Communication, 2019, 41(9): 19-24.
[20] Guo X K. Big Data[M]. Beijing: Tsinghua University Press, 2013: 104-105, 84, 86.
[21] Xu X D, Wang Y X. Causes, impacts and countermeasures of algorithmic bias in intelligent communication[J]. Chinese Journal of Journalism & Communication, 2020, 42(10): 69-85.
[22] Liu Z H. Algorithmic bias and discrimination in AIGC: identification, evaluation and mitigation methods[J]. Electronic Components & Information Technology, 2024, 8(2): 109-111,115.