

Legal Regulation of Algorithmic Discrimination in the Age of Artificial Intelligence

Xinhui Liu

Department of Economics and Law, Shanghai University of Political Science and Law, Shanghai, 201701, China

ABSTRACT

As mankind advances into the era of artificial intelligence, algorithmic discrimination, characterized by "mass micro-aggression", poses an undeniable hazard to the stable social order and urgently requires effective legal governance. Problems such as unclear identification criteria for algorithmic discrimination, the inability of existing laws to effectively regulate algorithmic discrimination, and the imbalance of the regulatory structure have adversely affected the effectiveness of algorithmic governance, and need to be overcome one by one. Under the premise of clear identification criteria of algorithmic discrimination, it is imperative to propose feasible legal regulation measures of algorithmic discrimination based on the causes of algorithmic discrimination and the regulatory effect of existing laws, and in combination with the needs of the people. The regulatory strategies proposed by domestic and foreign researchers are summarized through literature analysis and comparative study, on the basis of which a three-tier algorithmic discrimination governance framework is built with a mechanism for preventing algorithmic discrimination beforehand, supervising algorithmic discrimination during the process, and pursuing responsibility for algorithmic discrimination after the process, and proposals are made to set up a specialized algorithmic supervisory body, establish an algorithmic impact assessment mechanism, refine the algorithmic record-keeping system, establish a multifaceted supervisory mechanism, and set up a statutory procedure to facilitate class action lawsuits. These recommendations will help enrich and improve the details of legal regulation of algorithmic discrimination in China and fill the gaps in the existing laws.

KEYWORDS

Artificial intelligence; Algorithmic Discrimination; Automated Decision-making; Legal regulation

1 Introduction

From Alan Turing's initial conception of AI machines to the rapid development of AI technology based on the Internet, big data, cloud computing and other information technologies, AI has become an important force in boosting the development of the times. Algorithm is the core of artificial intelligence, as a tool to assist human problem solving or decision making, it is more and more widely used because of high efficiency and convenience, automatic driving, medical image processing, typhoon path prediction, video personalized recommendation, etc. There are traces of algorithm operation. Automated decision-making as one of the functions of algorithms is becoming more and more common in many fields, which helps human beings to save costs and improve work efficiency, and at the same time brings the legal risk of algorithmic discrimination. Algorithmic

discrimination may violate a variety of legal interests, such as citizens' right to privacy, equality, human dignity, etc. It is necessary to establish and improve the law to regulate algorithmic discrimination as soon as possible, which will help to make up for the shortcomings of the existing legislation, and better safeguard the equality and fairness of the social order, and at the same time promote the benign and healthy development of artificial intelligence technology.

2 Existing algorithmic definitions of discrimination and their limitations

Algorithmic discrimination, as a socio-technical phenomenon, has not yet been scientifically and widely recognized as a common definition. Even the term “algorithmic discrimination” itself has been confused with the term “algorithmic bias”. In response to this confusion, some researchers have clarified that “algorithmic discrimination” is more appropriate than “algorithmic bias”. Based on Hans-Georg Gadamer's theory of “prejudice”, (Bias is constituted by the historicity of understanding, is inherent and ineradicable, and is weakened, or strengthened, or altered by human cognitive activity.)¹ Meng Lingyu (2022) suggests that “treating people differently based on certain characteristics” and “causing negative consequences for the discriminated” are the necessary elements of discrimination, prejudice arises before discrimination, the concept of prejudice has more connotations than discrimination, discrimination can only be negative, and prejudice can be negative, neutral, or positive, and most people describe the phenomenon of algorithmic discrimination by emphasizing its negative impact. Discrimination can only be negative, while prejudice can be negative, or neutral, or positive, and the phenomenon of algorithmic discrimination is most often described in a way that emphasizes its negative effects, so it would be more appropriate to use the term algorithmic discrimination.

Currently, most researchers usually describe algorithmic discrimination from two perspectives: technical characteristics and the nature of the phenomenon, i. e., algorithmic discrimination is a number of systematic and repeatable errors in the operation of algorithms that may lead to the generation of unreasonable or unfair results . In layman's terms, any well-defined computational step (an input/output relationship designed to solve a problem) can be called an algorithm.²

The above definition of algorithmic discrimination is not particularly complete, and the extent to which differentiation can constitute algorithmic discrimination has yet to be discussed and refined. Obviously, as long as there is algorithmic differentiation, it will undoubtedly lead to over-intervention of the law. For example, excluding the scenario of “big data ripening” that really violates consumers' rights and interests, “big data ripening” in the eyes of consumers may just be that once the newcomer's discount period has expired, the merchant restores the due market price of the goods. However, even if algorithmic differentiation is qualified by the word “unreasonable”, its scope is still very broad. While it is difficult to formulate a quantifiable measure of “unreasonable differentiation” to help solve the problem of overly broad scope of legal regulation, interpreting “unreasonableness” from the perspective of the purpose of the law may serve as a breakthrough. The purpose of regulating algorithmic discrimination by law should be to seek a balance between safeguarding the interests of the individual and the public interest of society, and to seek to protect the legitimate rights and interests of the individual while maintaining the social order. On this basis, algorithmic discrimination that effectively infringes on the legitimate rights and interests of

1 Hans-Georg Gadamer: *Hermeneutics I - Truth and Method*, translated by Hong Hunding, Commercial Press, 2010 edition.

2 Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, *Introduction to Algorithms*, Fourth Edition, The MIT Press, 2022.

individuals, or that may have a greater adverse impact on the public interests of society or the social order if left unchecked, is algorithmic discrimination that should be regulated by law.

Some researchers have attempted to address the problem of overly broad regulation of algorithmic discrimination from the perspective of categorization. Li Zhiying (2024) argues that not all discrimination in the moral dimension is subject to legal regulation, and that discrimination subject to legal regulation is categorized into "algorithmic bias" regulated by anti-discrimination laws and "algorithmic inequality" regulated by laws of other domains, and proposes corresponding regulatory means based on this categorization. On the basis of this categorization, corresponding regulatory means are proposed.³ The problem with this proposition is that there is currently no specific anti-discrimination law in China, and furthermore, the ultimate point of departure for both types of discrimination is still the result of inequality, and such a categorization does not seem to be of substantial help in dealing with algorithmic discrimination.

3 Algorithmic discrimination retroactivity and attribution

3.1 Causes of algorithmic discrimination

Patricia Zivkovic and Rossana Ducato (2023) argue that algorithmic discrimination is usually a phenomenon caused by a lack of time, background, skills, and knowledge to evaluate algorithms and make appropriate decisions⁴, but this can only mean that inadequate evaluation of algorithms has made the potential risk of algorithmic discrimination a reality, i. e., that factors capable of causing algorithmic discrimination were already in place prior to algorithmic evaluation. This means that there were already factors that could cause algorithmic discrimination before the algorithmic evaluation was carried out.⁵

3.1.1 Perception projection: pre-existing biases inadvertently embedded in algorithmic programs

There are two main paths by which inherent biases in society may be reflected in an algorithmic program. One, inherent discrimination may be the reflection of human subjective cognition in the algorithmic program. Society is developing, and people in a particular historical context have limited insights, and the information they obtain may be the result of survivor bias. Based on limited insights, people tend to form one-sided perceptions, and algorithmic program designers are biased in certain aspects, and whether algorithmic program designers intentionally or unintentionally apply bias to the algorithmic program, the probability of discriminatory results may be increased. Second, inherent discrimination may also be reflected in the algorithm through the data. Specific historical and cultural contexts that create differences between people or groups of people are collected and fed into the algorithmic program as data, and the algorithmic program recognizes these differences, which may be further reinforced to create discrimination. Research has shown that due to the high rate of past criminal behavior among black people, ads suggesting a criminal history are more likely to appear in the search results when someone searches for a black person

3 Li Zhiying, "The Two Faces of Algorithmic Discrimination and Its Legal Regulation," in Jiao Tong University Law Journal, Vol. 1, No. 1, 2024.

4 Zivkovic Patricia; Ducato Rossana, Algorithmic Discrimination: A Blueprint for a Legal Analysis, EU and Comparative Law Issues and Challenges Series, Vol.7, 2023.

5 Shi Ying, "The Occurrence Logic and Legal Regulation of Algorithmic Discrimination," in Theory Exploration, No. 3, 2022

than when they search for a white person in the search bar of Google Chrome.⁶ Moreover, data is delayed and data collected from past records may not match current and future developments.

3.1.2 Technical shortcomings: technical limitations of the algorithm's predictive capabilities

Prediction, as one of the functions of an algorithm, is essentially the analysis of existing data to determine whether a target is more likely to do something, i.e., a “fact-based prediction” model. On the one hand, algorithms usually select a portion of the data to make predictions, which can allow the algorithm to draw one-sided conclusions based on one-sided data, increasing the likelihood that algorithmic discrimination will occur; on the other hand, life is inherently contingent, so accurately predicting what will happen in the future is extremely difficult. Sandra G. Mayson (2019) argues that all predictions are based on the past to speculate on future events. Any forecasting method may project past inequalities into the future.⁷ While this argument addresses racial inequality in risk assessment, it should be applicable in any algorithmic decision making of a predictive nature, as all algorithms of a predictive nature are judging future outcomes based on past data, which will inevitably and unavoidably deviate from the actual situation. For example, a team at Amazon once created an algorithmic hiring system to browse through the past ten years of candidate resumes, using specific keywords in the resumes to capture candidates on the web that the algorithm thought would be qualified for the position in the future. The team eventually disbanded the system after a year when it realized that because the majority of past resumes were owned by men, the algorithm assumed that men had a dominant position in the industry and therefore favored men when recommending candidates.⁸

3.1.3 Thought Bias: The Difference Between Algorithmic Logic and Human Logic

Deviations between the logic of algorithmic thinking and the logic of human thinking may lead to different ways of approaching problems, which in turn may lead to algorithmic discrimination. For example, induction, one of the algorithmic methods of thinking, is characterized by summarizing laws or experiences and drawing conclusions based on available facts, a process in which algorithms are more likely to oversimplify certain experiences than humans. For example, Amazon's algorithmic hiring system assumes that men perform better in the industry based on data. However, the fact that candidates' resumes have been more heavily weighted toward men in the past decade cannot rule out the fact that men have been more heavily weighted toward relevant education in the past decade, and the data does not negate the fact that women applying for a job are also capable of performing in that industry, much less the possibility that women will outperform men in certain areas of the job once on board. The use of oversimplified experience to screen candidates not only leads to discrimination, but also makes companies miss out on talent.

Algorithmic black boxes can lead to algorithmic discrimination. Algorithm black box refers to the complexity, intelligence and professionalism of algorithms, which makes it difficult for the public, especially non-professionals, to distinguish the logic of the decision-making process, the basis for decision-making and the reasonableness of the decision-making results. In a certain sense, the algorithm black box is a manifestation of the deviation between the logic of algorithmic thinking

6 Latanya Sweeney, Discrimination in online ad delivery, at <https://queue.acm.org/detail.cfm?id=2460278>, Apr.2, 2013.

7 Mayson, Sandra G., Bias In, Bias Out, Yale Law Journal, Vol.128, 2019.

8 Jeffrey Dastin, Amazon scraps secret AI recruiting tool that showed bias against women, at <https://www.irishtimes.com/business/technology/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-1.3658651>, Oct.10, 2018.

and the logic of human thinking. In addition, algorithmic products are usually provided to the public by enterprises, and in order to protect commercial secrets, enterprises usually do not disclose the code used, so the general public has no way to monitor whether the results of the algorithm are reasonable and fair. Speaking of algorithmic products, it is necessary to mention machine learning, as a number of products or services based on machine learning have emerged. As a branch of artificial intelligence, machine learning is based on algorithms, which refers to allowing computers to automatically increase or improve their performance by learning from data and experience as a way to better continue to accomplish certain tasks. Machine learning is optimized to such an extent that the involvement of the algorithm designer is no longer required, as in the case of deep learning, so that even professionals cannot accurately determine whether the results of a series of algorithms contain discriminatory factors or are reasonable.⁹

3.2 Attribution analysis of algorithmic discrimination

3.2.1 Determination of algorithmic discrimination

There are a variety of views in the academic community regarding the identification of algorithmic discrimination. For example, the traditional anti-discrimination theory that categorizes algorithmic discrimination into direct and indirect discrimination based on the presence or absence of intent to discriminate, the associative discrimination theory that addresses discrimination based on factors other than the protected characteristics in the traditional anti-discrimination framework, the algorithmic game locking theory.¹⁰

The three elements of commercial necessity, disparate impact and substitutable algorithmic decision-making are adopted by U.S. courts in practice as the criteria for determining or reviewing algorithmic discrimination. Commercial necessity examines whether the design of the algorithm is directly related to the goal pursued by the commercial subject; disparate impact means that regardless of whether the subject of the algorithmic decision-making subject has subjective intent in implementing the behavior, it only looks at whether the behavior or decision of the algorithm has caused a differential adverse impact on a specific group of people; substitutable algorithmic decision-making means that unless the subject of the algorithmic decision-making subject proves that there are no other algorithms that can substitute the algorithm that has caused negative effects to the relative negative impact on the relative, then the possibility of algorithmic discrimination cannot be ruled out even if the original algorithm is indeed relevant to business needs .¹¹The meaning of differential review in the United States is basically the same as that of the "differential review model" proposed by domestic researchers.

This paper endorses the United States' three-factor standard of review because it takes into account the need to preserve the vulnerability of the discriminated against and emphasizes the focus on the disparate and adverse impacts on the discriminated against, while also allowing room for the business entity to defend itself.

3.2.2 Attribution of algorithmic discrimination

With regard to the types of rules for damages caused by algorithmic discrimination, theoretically,

9 Ignacio N. Cofone, *Algorithmic Discrimination Is an Information Problem*, *Hastings Law Journal*, Vol.0, 2019.

10 Niu Binbin, "The Recognition and Regulation of Algorithmic Decision-Making Discrimination on Digital Platforms," in *Journal of Tianjin University (Social Science Edition)*, Vol. 1, No. 1, 2024.

11 Zheng Yushuang and Bao Mengru, "Re-exploration of the Legal Regulatory Model of Algorithmic Discrimination," in *Digital Law*, Vol. 1, No. 1, 2023.

the general tort liability provisions of the Civil Code, the Internet provisions and the product liability provisions can be applied. It seems feasible to apply the general tort liability provisions, but the problem is that under the general tort liability, the injured party needs to bear the burden of proof for the subjective fault of the infringer, both in terms of professionalism and in terms of the difficulty of obtaining evidence, it is too onerous to let the injured party bear the burden of proof, so it is inconvenient to directly apply the general tort liability provisions; while the Internet provisions regulate the infringement of rights using the network. Algorithmic discrimination is mainly generated under the joint action of data and algorithm, and the network may not be utilized in the process, so this article is not suitable for directly regulating algorithmic discrimination; algorithmic products or algorithmic services can be regarded as products, and therefore the product liability provisions can be applied.¹² The European Parliament has also called for AI product liability in its 2020 proposal for a civil liability regime for AI.¹³

Article 2, paragraph 2 of China's "Product Quality Law" for the first time on the "product" generalized definition - "product refers to the processing, production, for sale of products", the current China's laws and regulations and judicial interpretations of the "processing, production" and "for sale" lack of a clear definition, but if the "processing, sales" is only understood as manual or industrial. However, if "processing and selling" is understood to mean only manual or industrial processing and production, which is not conducive to the protection of consumers' rights and interests, the act of influencing or producing the quality of any product should be regarded as processing and production. In this way, the algorithmic decision-making system to be influenced or controlled would naturally fall within the scope of "processing, production". And "for sale" if only understood as directly from the consumer to obtain monetary gain is obviously not conducive to consumer rights, because many algorithmic decision-making systems do not directly from the user to obtain benefits, it may be the use of the user to bring huge traffic or business prospects to obtain investors to invest or to third parties to collect advertising fees to Profit. Some algorithmic decision-making systems may provide some services to users for free, but if users want to enjoy better services, they have to purchase them by recharging their money, which does not mean that users who only use free services are not subject to algorithmic discrimination. If only the top-up users, i.e. consumers, are allowed to defend their rights against the results of algorithmic discrimination, it seems to be fair, but it is unreasonable.

The operator did not stipulate that users could not enjoy any services of the algorithmic decision-making system without recharging, but instead provided part of the services free of charge, which seemed to be a unilateral gift from which the operator did not make any profit and thus did not need to bear any responsibility, but in fact, the operator provided users with part of the free services in order to expand the popularity of the service, to attract more potential consumers and even to attract investment, which is a kind of business tactics. The partially free service provided by the operator for the users cannot negate the continuity and profitability of its business behavior. As the operator, the operator should fulfill the obligations stipulated in the law, and as long as the service is provided, the operator should be responsible for the damages caused by the service. Therefore, this kind of "trial service" or "trial sale" should also be categorized as sales behavior. Under this interpretation, the service provided by the algorithmic decision-making system should be categorized as a product.

12 Shang Jianguang, "Analysis of the Elements of the Composition of Algorithmic Decision-Making Damage Liability," in *Oriental Law Journal*, No. 6, 2022.

13 European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014 (INL)), Civil liability regime for artificial intelligence, at https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_EN.html, Oct. 20, 2020.

The liability of the producer of an algorithmic product or algorithmic service should derive from the fact that he enjoys a certain degree of control over both the operation and the functioning of the algorithmic product or algorithmic service, and that even if the producer becomes aware of the defects in the algorithmic product or algorithmic service after the algorithmic discrimination has arisen, he can intervene in modifying or eliminating the part of the algorithmic operation that contains risks of discrimination. This right of control makes the producer's rights exceed that of the user of the product, and according to the principle of consistency of rights and responsibilities, when the scope of the producer's rights is enlarged, the obligations that the producer should bear naturally become more. Shang Jianguang points out that the controller of algorithmic decision-making should bear the obligations of legality of algorithmic decision-making (i.e., it must bear the obligations explicitly stipulated by the law), the obligations of data appropriateness (i.e., it must ensure that the data processing is lawful and legitimate, and ensure that the results of algorithmic decision-making are just and correct), and the obligations of appropriateness of algorithmic system and application (i.e., algorithms must be designed with an emphasis on social fairness and safeguarding the interests of disadvantaged groups). Li Zhiying pointed out that the justification for algorithm designers and algorithm applicators to assume more obligations is based on the obligation of responsibility transfer (applying reversal of the burden of proof by drawing on the principle of tort liability and taking into account the difficulty of proof for victims of algorithmic discrimination) and the obligation of social law (providing protection for the disadvantaged groups and correcting the deviation of freedom of contract between the strong and the weak).

With regard to the attribution of responsibility for algorithmic discrimination, it is important to recognize the question of who the responsible subject should be. It can be confirmed that although algorithmic decision-making is one of the functions of artificial intelligence, at present artificial intelligence cannot be the responsible subject of algorithmic discrimination. Even if artificial intelligence enters into deep learning no longer requires the intervention of the algorithm designer, the generation of artificial intelligence requires professionals to design algorithms and build models in the early stages. Moreover, most algorithms require the designer to determine the context of use and how to use them. Too much overestimation of the power of algorithmic decision-making is actually an underestimation of intervention. Artificial intelligence systems not only do not have a legal personality but also do not have a human conscience, and the subject of responsibility for algorithmic discrimination should still be returned to human beings.¹⁴ Article 1202 of the Civil Code stipulates that "If a product is defective and causes damage to another person, the producer shall be liable in tort." This paper argues that reference can be made to the European Parliament's proposal to the Commission for a civil liability regime for AI, which states that producers in algorithmic discrimination can refer to manufacturers, developers, programmers, service providers, and back-end operators. It is important to note that the developer of an algorithm should be the person who is involved in most of the algorithm's design and who has a decisive role in determining "what the algorithmic product should do". A person who provides only part of an algorithmic program, such as a person who makes some algorithmic programs available for use in a public manner through Internet forums and book publications, but who does not foresee that the algorithmic program he provides may be used for discriminatory purposes should not be held liable, even if a large number of algorithmic programs provided by the person are found in the product that led to the algorithmic discrimination.

According to Article 1202 of the Civil Code, product liability should be composed of three parts, namely, that the product is defective, that the defective product has caused damage to others, and

14 Ragnar Fjelland, *Why general artificial intelligence will not be realized*, Humanities and Social Sciences Communications, Vol.7, 2020.

that there is a causal relationship between the defective product and the fact of the damage caused. Due to the algorithmic decision-making system once established has opacity, connectivity and even a certain degree of autonomy, want to algorithmic decision-making system based on the harmful behavior traced back to a particular person or a particular algorithmic design of the consumer is extremely difficult, is not conducive to the consumers to protect their rights, based on the principle of fairness and justice, when the algorithmic decision-making system so that the producer is in an advantageous position and the consumer is in a disadvantageous position, should be taken to burden of proof Inversion of the method, the producer is given no-fault liability or presumption of fault liability, that is, the producer proves that the algorithmic product does not have defects, and there is no causal relationship between the damage suffered by the consumer and the algorithmic product, or the producer needs to prove that the damage suffered by the consumer is not caused by the defects of the algorithmic product, or else it will be required to bear the corresponding responsibility; and the consumer needs to prove that the damage suffered by him damage and the defect of the producer's algorithmic product have a causal relationship.

4 Current Status and Dilemma of Legal Regulation of Algorithmic Discrimination

4.1 Current status of legal regulation of algorithms in China

The Cybersecurity Law (2017) clarifies the security protection obligations of network product and service providers and network operators, and increases penalties for violating the law, but does not clarify the procedures for network operators to maintain the security of personal information; the Data Security Law (2021) requires the establishment of a data security system, and clarifies China's data sovereignty, but there is still a need to further improve and standardize the data classification and classification management system and data transaction management system; the Personal Information Protection Law (2021) establishes rules for the identification and protection of sensitive personal information and regulates automated decision-making, but the implementation of the norms needs to be refined.

The Regulations on the Administration of Algorithmic Recommendation of Internet Information Services (2020) is the first special regulation on algorithmic recommendation services in China, which clarifies that users have the right to be informed of algorithms and the right to choose algorithms, regulates the providers of algorithmic recommendation services for special subjects (e. g., minors, the elderly, laborers, and consumers), and clarifies the filing obligations of algorithmic recommendation service providers, the obligation to conduct security assessments, and the obligation to cooperate with supervision and inspection. obligations, safety assessment obligations and obligations to cooperate with supervision and inspection, etc. However, there are problems such as the content and procedures for assessing algorithms are not clearly defined, the responsibilities of supervisory departments are not clear and the punishments are relatively light.

Overall, the legal regulation system for algorithms in China is relatively weak, and needs to be further improved and perfected according to the actual use effect, and there is a gap in the legal regulation of algorithmic discrimination.

4.2 The Dilemma of Legal Regulation of Algorithmic Discrimination

4.2.1 Institutional Challenges: The Damaging Characteristics of Algorithmic Discrimination Are Different from Conventional Traditional Torts

Unlike traditional "clear and substantial harm", harm caused by algorithmic discrimination is an example of a "mass micro-tort". These "mass micro-torts" may cause less legal harm to individuals,

but the number of people involved may be twice as large as in traditional torts, or even exponentially so, and the character of the harm varies according to the specific scenario and the outcome of the harm is difficult to predict. For society as a whole, the negative impact of algorithmic discrimination is no weaker than that of traditional copyright infringement, and even its hidden damage may exist for a longer period of time, and the scope of the damage may be wider with the advancement of time. The original legal provisions may not be fully compatible with the problems caused by algorithms, and it may be necessary to constantly re-interpret or even modify the original law to effectively regulate algorithmic discrimination. The lagging nature of the law is in stark contrast to the wild growth rate of algorithmic technology.

4.2.2 Technical Barriers: Conflicting Legal and Regulatory Requirements and Algorithmic Technical Characteristics

Some of the requirements in laws and regulations may seem straightforward, but the specialized nature of algorithms makes it difficult to guarantee some of the legal requirements in practice. For example, the legislation stipulates that automated decision-making should ensure the transparency of decision-making, but how to ensure transparency in practice is a problem, the law has not yet clarified what the algorithm designer or algorithm decision-maker should disclose, and what penalties there are for non-disclosure. Some researchers believe that algorithm designers or decision makers should disclose the source code, input data and output results. However, for the purpose of protecting intellectual property rights and trade secrets, algorithm designers or decision makers should prefer not to disclose such information, and even if the disclosed information is false, it will be difficult to be detected. Moreover, in the context of deep learning algorithms, algorithm designers and decision makers are not able to grasp the logic of the algorithm's operation, and even if they can do so, it does not mean that the algorithmic problems can definitely be discovered.

4.2.3 Regulatory imbalance: a multi-party governance system led by government departments has yet to be established

At present, China's algorithmic governance is still dominated by the participation of public power. There are relevant regulations expressly written to "encourage relevant industry organizations to strengthen industry self-regulation", strengthen industry self-regulation usually need to be driven by the leading enterprises, the development of the Internet so far has slowly tended to natural monopolies, than to take the lead to strengthen industry self-regulation, leading enterprises in the barbaric development of the profits gained under the more, without external interference, it is very difficult for enterprises to have the power to Develop industry rules widely recognized by the industry. While the government and other relevant departments have been given the responsibility to regulate, but without clarifying the specific responsibilities of the premise, it is easy to shirk their responsibilities for the thankless work of each other. Although the public has the right to monitor and report, but limited to the ability, even if the algorithm is found to be problematic, but also do not know through what channels upward feedback and difficult to come up with strong evidence.

4.2.4 Lack of Talent: Lack of Complex Talent with Proficiency in Algorithmic Technology and Legal Expertise

Composite talents are always in short supply, people who understand legal expertise may not understand the results of the cutting edge of algorithmic technology, and people who stand at the

forefront of science and technology may not be willing to cooperate with the law to restrain themselves. Although the vision of the relevant laws to maintain fairness and justice is good, but people want to do in accordance with the relevant laws usually know the direction but don't know the path, don't know how to do it is difficult to reach the realization of the beautiful pursuit of fairness and justice. Some of the norms need to be more detailed, but how to be detailed is the need for professionals to work together. Perhaps it is not only necessary to regulate the problem of algorithmic discrimination from the legal point of view, but also to utilize technology to achieve the practical effect of regulation.

5 Suggestions for Improving the Legal Regulation of Algorithmic Discrimination in China

5.1 Establishment of the “Recognize, Evaluate, Record” ex ante preventive mechanism

To regulate the criteria for legal determination of algorithmic discrimination, it can be judged whether it constitutes algorithmic discrimination based on factors such as the existence of “differentiation of treatment based on certain characteristics”, “adverse consequences for the discriminated person”, “greater risk of disadvantage to the public interests of the society or social order”, etc. and “a greater risk of disadvantage to the public interest or social order”. Regarding the object of regulation of algorithmic discrimination, reference can be made to the European Parliament's proposal to the Commission on the civil liability regime for AI, which identifies the subject of regulation of algorithmic discrimination as the operator, including front-end operators and back-end operators. Operators are usually able to control or influence the operational effects and related functions of the algorithmic decision-making system, and they can be held liable according to the source of the discriminatory results. With regard to the measurement of the reasonableness of algorithmic decision-making, the law should provide for a statutory bar to algorithmic discrimination, such as the necessity or irreplaceability of the factors taken into account by the algorithmic decision-making system or the choice of decision-making solution. Without a clearer legal standard for algorithmic discrimination, even if a producer's algorithmic product does lead to algorithmic discrimination, the producer can find plenty of room to argue that he or she has escaped punishment or has been mitigated.

Establish an algorithmic impact assessment mechanism. Relevant government departments should clarify the content, standards and procedures of the algorithmic impact assessment mechanism, and standardize the process of algorithmic assessment for enterprises on the premise of giving them the obligation to conduct algorithmic impact assessment. On the issue of “who should conduct the assessment”, the mode of “government-enterprise cooperation and government-monitoring-enterprise assessment” can solve the problems of insufficient manpower in government departments and low professionalism of algorithms, on the one hand, algorithm operating enterprises should conduct self-assessment of algorithms internally, and on the other hand, the government should authorize another qualified algorithm assessment organization to assess the impact of algorithms. On the one hand, algorithm operators should conduct self-assessment of their algorithms internally, and on the other hand, the government should authorize another algorithm assessment enterprise with relevant qualifications to issue an assessment report on enterprises participating in the impact assessment of algorithms. The content of the assessment may include, but is not limited to, the legitimacy of the data source, the reasonableness and necessity of the decision-making model and data selection, the privacy risk and privacy leakage prevention measures involved in the decision, the possible adverse impact of the decision on the user and remedial measures, as well as the ways to realize the rights and interests of the user, and so

on. The evaluation criteria should be as objective and clear as possible, for example, the use of data must be authorized by the user, and the operator must clearly inform the user of what data will be collected and included in the relevant decision-making considerations. The Algorithmic Impact Assessment (AIA) phase is intended to identify algorithmic discrimination in the algorithmic decision-making system as early as possible before the algorithmic product is formally introduced into the market, and should be tested with a large amount of data input to see if there are any algorithmic discriminatory results. Given that the process of inputting large amounts of data may also be an invasion of personal privacy, the input data may not be real data, or the data may be anonymized. Considering the current high cost of algorithm evaluation, algorithms that have a greater impact on public decision-making or high-risk areas can be required to be evaluated first, and then widely promoted after a more complete algorithmic impact evaluation mechanism has been established and the cost of the evaluation has been reduced to a range that is acceptable to the majority of enterprises.

Refine the algorithm filing system. The algorithm filing system can strengthen the transparency of algorithms, and the relevant departments can regularly check whether there are any problems with the algorithms through the filing information. However, if the filing content is too rough, even after going through the filing procedure, it is difficult to find problems with the algorithms. Algorithm filing system needs to be refined in a number of places, for example, in addition to the clear requirements to fill in the form of algorithms, types of services, application areas, etc., it can be required to fill in the algorithms in the actual operation of the source of the data obtained and the way of obtaining the data, the data in the algorithm system to save the period of time and based on the data, the data of the subsequent results of the processing, and so on. In addition, the required contents of the algorithm self-assessment report and how to recognize the qualified algorithm self-assessment report can be written.

5.2 Building a multi-dimensional monitoring mechanism for “government, industry and users”

Specialized regulation of relevant government departments. Government departments hold a wealth of data on people, and if the government generates algorithmic discrimination through automated decision-making systems, it may be difficult for ordinary people to detect discrimination based on their trust in the government, let alone adduce evidence of their damaged rights and interests. It is important to clarify the areas in which the government can utilize algorithmic decision-making systems, the procedures by which the government will utilize algorithmic decision-making systems, what agencies will monitor the government's use of algorithmic decision-making systems, how citizens can monitor the government, and what avenues of redress are available to citizens, among other things. It is also desirable to centralize the management of government-acquired data, rather than using disconnected data stores for different departments. Perhaps multiple work sub-systems and a work of the total system can be connected, different departments only have different sub-systems of the authority, and then according to the different requirements of the release of the corresponding authority to give the right to use the data, the specialized monitoring body has the right to regularly review the corresponding departments to use algorithms to automate decision-making is compliant.

Promote and strengthen industry associations. Industry associations, as a kind of non-governmental organization, can unite individual strengths in the form of organization and grouping, and can satisfy the common needs of association members while at the same time constraining their misconduct through communication and coordination and punishment. The industry association can stipulate the ethical standards that should be followed in designing

algorithms, the technical standards adopted by algorithms, and the code architecture of algorithmic programs. As a kind of self-generated and spontaneous order, reasonable industry rules derived from the internal game and equilibrium of the industry are a powerful supplement to the national legal order, which can help the whole industry develop more healthily and sustainably. If the industry association cannot spontaneously establish reasonable and lawful industry rules, the government needs to supervise and provide assistance within a certain range.

Improve the algorithm filing information open system. Build an open platform for algorithmic filing information on the network, make public the filing information that enterprises can make public, and users can put forward problem descriptions, evidence, and suggestions or demands (such as the actual situation is inconsistent with the filing information, and enterprises are required to correct the filing information according to the actual situation) on the open filing information in the relevant function boards, and the platform will forward the feedback to the supervisory department, and then the supervisory department will make a decision whether or not to cooperate with the relevant person in charge of the enterprise according to the feedback content. The platform will forward the feedback to the regulator, who will then make a decision on whether or not to dock with the relevant person in charge of the enterprise based on the feedback, and feedback the relevant progress to the relevant user. Users must have a real name within the platform and can choose whether to make the feedback public or not, and the real name of the user who chooses to make the feedback public will be anonymized when it is made public. Other users can supplement the description of the problem, supplement the evidence, and supplement the suggestions or demands on the same issue in the designated function area of the page of the publicized feedback content. This channel serves as one of the means to ensure the right of users to monitor.

The use of algorithmic decision-making system is a long-term, dynamic process, whether the operator based on factors such as efficiency to replace the algorithm, or change the algorithmic decision-making system input data may produce new discrimination, the algorithmic platform algorithmic decision-making system for long-term, continuous review is conducive to timely detection of new discrimination risks, to reduce the adverse impact on the user.

The review method can be categorized into regular review and irregular spot check. Considering the huge workload of the review, it is possible to focus on material review and algorithmic program review separately. Regular review mainly checks whether the relevant materials of all algorithmic platforms filed and in operation are complete. Irregular spot checks, on the other hand, focus on technical audits, in which the regulator has to check the reasonableness of the design of the algorithmic decision-making system, whether there is any potential risk of discrimination in terms of data and program, whether the algorithm has already caused discriminatory results, and whether the problems that have been ordered to be corrected in the past have been resolved, and so on. The algorithmic regulator should be given the power to obtain detailed information about the algorithmic system and to review and assess it, as well as the obligation to protect business trade secrets. If, according to the assessment of the algorithmic regulator, the algorithmic decision-making system has produced unfair discriminatory results and caused adverse impacts, the regulator shall, in accordance with the law, take reasonable and specific administrative actions and require the operator of the algorithmic platform to make corrections within the prescribed period, and if the operator fails to make corrections within the prescribed period, the operator shall stop operating the algorithmic decision-making system. The regulatory authority's opinion on the handling of the operator or the algorithmic platform shall be publicized to the public.

The operator shall conduct regular self-inspection of whether the algorithmic decision-making system it operates poses a risk of discrimination, keep a good record of the self-inspection and submit it together with other materials to the reviewing authority on a regular basis, and make a

clear and concise announcement to the public regarding the self-inspection on a regular basis. The operator shall keep the results of the algorithmic decision-making system on a regular basis and regularly analyze what factors are associated with the results and whether these factors imply discrimination. The operator's self-inspection may focus on whether the algorithmic decision-making system poses certain discriminatory risks according to the specific characteristics of the application scenarios of the algorithmic decision-making system. For example, operators should focus on whether the algorithmic decision-making system applied in the field of job search and employment has the risk of gender discrimination, and whether the algorithmic decision-making system applied in the field of online consumption has the risk of price discrimination. Operator self-checks can raise operators' awareness of risk prevention and mitigate the potential damage caused by risks. Self-inspection by operators can also reduce the burden of algorithmic regulators to a certain extent. The self-inspection disclosure system is conducive to safeguarding users' right to know, and at the same time can force operators to take proactive measures to optimize their algorithmic procedures. In order to prevent the system from becoming a useless decoration, operators that do not publicize their self-inspection on schedule or falsely disclose their self-inspection should be held accountable by law.

5.3 Improvement of “reasonable, convenient and efficient” recourse and relief mechanisms

Setting up a specialized algorithm supervisory body. At present, algorithms are mainly supervised by the competent organizations of various industries in China, and there are problems of different supervisory standards and varying supervisory efforts. Setting up a specialized algorithmic supervisory body will, on the one hand, make the powers and responsibilities between state agencies more clearly defined, and, on the other hand, be conducive to the centralized and unified management of algorithm-related issues. In addition to algorithmic discrimination, the algorithmic supervisory body can formulate corresponding management programs for other algorithmic risk issues. At the same time, because algorithmic issues have a high technical threshold, requiring a certain degree of computer knowledge to deal with the relevant issues, the establishment of a specialized algorithmic supervisory body is conducive to the management of the relevant personnel.

Setting up statutory procedures to facilitate class action lawsuits. Algorithmic discrimination exists in a wide range of infringing objects, the infringement of individual legal interests is small, under normal circumstances, the infringer will be on the infringement of algorithmic discrimination will be carelessly ignored, even if there are infringers realize that algorithmic discrimination is potentially very harmful, but also because of the high cost of litigation, litigation is cumbersome, the pressure of proof is too large and give up the use of litigation to solve the problem. At this point, class action lawsuits become a breakthrough to solve the problem. For some people who want to protect their own rights and interests but are deterred by the cost of litigation, if someone can sue on their behalf based on the same interests, they can not only protect their own interests, but also save the cost of litigation. The difficulty is how to bring together these groups of people who share the same interests. If the relevant departments can develop a system in which parties suffering from algorithmic discrimination, after filling in basic information such as identity, can fill in information such as the object of the lawsuit, the legal interests chosen to be infringed upon, the lawsuit request, the facts and reasons, and the lawsuit fee they are willing to pay, etc., the system will categorize and integrate the cases according to the object of the lawsuit and the legal interests infringed upon, and at the same time, the system will notify the users of the number of people who have the same interest claims as the users. At the same time, the system notifies the user of the number of people who have the same interests as the user. When the number of people and the

amount of litigation meet certain requirements, the system sends the general situation of the case to the user again and sends the group chat portal, and the user identifies a representative in the group. The user negotiates whether or not to agree with the system to publish the blurred case information and the representative's contact information in the form of a commission in a specific section of the system, and once agreed, a lawyer certified by the system can get in touch with the representative according to the contact information left by the representative. The system displays the progress of the case. Any member who does not wish to participate in the class action must apply in person to withdraw, otherwise it counts as automatic participation.

Establishment of a reasonable review process. The court may judge whether the design, operation logic and results of the algorithm are reasonable and justified based on the algorithm impact assessment report originating from the regulatory authorities. At the same time, the principle of reversal of the burden of proof shall be adopted, requiring the subject of the algorithm to prove that the relevant damage was not caused by the algorithm; if the subject of the algorithm refuses to provide relevant evidence or the evidence provided is not effective in proving that the damage was not caused by the algorithm, the court may combine the ascertained facts with the presumption that the algorithm caused the damage. The procedural rights of the parties should be safeguarded by granting them appropriate explanations, adequate time limits for defense, adequate time limits for cross-examination, and setting confidentiality obligations for the parties.

6 Results

The urgency of managing algorithmic discrimination is becoming more and more prominent, and government departments should not just put pressure on the main body of algorithm design and application when managing algorithmic discrimination, but should intervene moderately, encourage enterprises to take the initiative to assume social responsibility while guiding the public to correctly understand the positive and negative impacts of algorithms, and set up a sound and long-term regulatory system. On the basis of analyzing the causes of algorithmic discrimination, this paper tries to establish an *ex ante* preventive mechanism of "identification, assessment, and filing", build a multifaceted monitoring mechanism of "government, industry, and users", and improve a "reasonable, convenient, and efficient" remedial mechanism of recourse and responsibility. "We will build a governance framework for algorithmic discrimination in three aspects, and discuss the implementation of this governance framework from the perspective of clarifying the rights and responsibilities of the government, enterprises, and the public, with the expectation that it will be of positive significance to accelerate the construction of a comprehensive algorithmic risk governance system in China. The development of institutional norms should follow the pace of technological development and closely follow the real needs of the people, so as to achieve a balance between the maintenance of social order and economic development.

About the Author

Xinhui Liu, Doctor of Law, Lecturer, Research Interest: Science and Technology Law.

References

- [1] Hans-Georg Gadamer: *Hermeneutics I - Truth and Method*, translated by Hong Hunding, Commercial Press, 2010 edition.
- [2] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, *Introduction to Algorithms*, Fourth Edition, The MIT Press, 2022.
- [3] Li Zhiying, "The Two Faces of Algorithmic Discrimination and Its Legal Regulation," in *Jiao Tong University Law*

Journal, Vol. 1, No. 1, 2024.

- [4] Zivkovic Patricia; Ducato Rossana, *Algorithmic Discrimination: A Blueprint for a Legal Analysis*, EU and Comparative Law Issues and Challenges Series, Vol.7, 2023.
- [5] Shi Ying, "The Occurrence Logic and Legal Regulation of Algorithmic Discrimination," in *Theory Exploration*, No. 3, 2022
- [6] Mayson, Sandra G., *Bias In, Bias Out*, Yale Law Journal, Vol.128, 2019.
- [7] Ignacio N. Cofone, *Algorithmic Discrimination Is an Information Problem*, Hastings Law Journal, Vol.0, 2019.
- [8] Niu Binbin, "The Recognition and Regulation of Algorithmic Decision-Making Discrimination on Digital Platforms," in *Journal of Tianjin University (Social Science Edition)*, Vol. 1, No. 1, 2024.
- [9] Zheng Yushuang and Bao Mengru, "Re-exploration of the Legal Regulatory Model of Algorithmic Discrimination," in *Digital Law*, Vol. 1, No. 1, 2023.
- [10] Shang Jianguang, "Analysis of the Elements of the Composition of Algorithmic Decision-Making Damage Liability," in *Oriental Law Journal*, No. 6, 2022.
- [11] Ragnar Fjelland, *Why general artificial intelligence will not be realized*, Humanities and Social Sciences Communications, Vol.7, 2020.