

Algorithm Interpretation Right—The First Step to Algorithmic Governance

Caixia Zou* , Fanyu Zhang

Shanghai University of Political Science and Law, Shanghai, China

Email: *caier7@163.com

How to cite this paper: Zou, C. X., & Zhang, F. Y. (2022). Algorithm Interpretation Right—The First Step to Algorithmic Governance. *Beijing Law Review*, 13, 227-246.
<https://doi.org/10.4236/blr.2022.132015>

Received: March 23, 2022

Accepted: April 24, 2022

Published: April 27, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

As artificial intelligence technologies are increasingly deployed in this digital era entitled to generate automated and semi-automated decisions, and the internal logic of machine learning algorithms is typically opaque, the absence of a right to explanation will put the individuals in a weak position, so the right to an explanation for such decisions has become a critical legal issue. The right to an explanation is first written in the “General Data Protection Regulations”, and this year, China lays out the “Personal Information Protection Law”, which also includes the right to explanation. However, there are still shortcomings in the relevant regulations. Therefore, this paper comprehensively sorts out the right to explanation by studying the relevant literature and using the method of comparison, points that to construct the applicable logic of the right to explanation in China to ensure its effective implementation, it is necessary to clarify the connotation of the right to explanation, establish its application path and build a collaborative governance system.

Keywords

Automated Decision, Right to Explanation, Personal Information Protection Law, Artificial Intelligence

1. Introduction

The advent of the era of big data has completely changed people’s way of life and work. Although we enjoy the great convenience brought by this era, at the same time, people’s personal information is abused, which makes the protection of personal information face severe challenges. A new business method in the era of big data is to analyze massive data containing personal information through various algorithm models to establish complete and specific user portraits, and

to push accurate information to customers so as to influence personal decisions or make some automated decisions that have legal effect on people. The so-called automated decision-making is a concept opposite to the decision-making of natural persons, which refers to the use of computer technology, algorithm programs, deep learning or neural networks to replace natural persons to process key data, and to automatically generate decisions with legal effects on the data subject (Tang, 2020). Today, this kind of automated decision-making based on algorithmic technology has spread all over people's lives, but in the final analysis, this kind of automated decision-making is developed based on algorithmic models.

The IDC report pointed out that the market size of China's artificial intelligence infrastructure will reach US\$3.93 billion in 2020, a year-on-year increase of 26.8%. And AI computing power has become a key factor for future breakthroughs in artificial intelligence, and complex algorithm models have greatly promoted the development of artificial intelligence in the industry (IDC & Inspur, 2021). In today's highly deployed algorithms, people are in a world controlled by algorithms every moment of every day. Digital technologies collect vast amounts of data and evaluate every aspect of people's lives: what they buy, what they do, what they think, how they work, and how they conduct their personal and intimate lives. Any information can be collected and encoded, and the information collected can be used in many situations, such as job applications, social benefits, or loans. This rating system is run by algorithms, not humans. A person's personal life may change as predictive algorithms have an impact on individuals making important decisions. More exaggeratedly, this means that many social activities will change: finances, marketing, insurance, employment, housing, education, political elections, judicial decisions, and more. Algorithms are increasingly being used in these areas to make decisions that matter to individuals. In 2020, the main customers of China's artificial intelligence market come from government urban governance and operations (public security, traffic police, justice, urban operations, government affairs, transportation management, land resources, prisons, environmental protection, etc.). The financial industry is close behind, accounting for 18% and 12% respectively (PIRI, 2021). It can be seen that algorithms are becoming more and more crucial in both the public and private spheres. The advantages of algorithmic decision-making to improve the efficiency of social governance and save decision-making costs determine that it will continue to develop at a geometric growth rate in the future, and the expansion trend is a foregone conclusion. In this case, the right to explanation was proposed and incorporated into the legislation, and it also sparked a wide-ranging discussion. These discussions also pointed to concerns about the development and implementation of this right in various countries. Of course, China is no exception. At present, China has transplanted the right to explanation, so doing more research on it and making this right better implemented in China has become the research topic of this paper.

2. The Proposition of the Right to Explanation

2.1. Algorithmic Nuisance under Algorithmic Decision-Making System

The concept of “algorithmic nuisance” was proposed by Professor Balkin, who likens the harm caused by algorithms to nuisance, and believes that the concept helps us understand how the harm of algorithmic society is generated by the cumulative decision-making and judgment of a wide range of public and private actors. And it elaborates on the harm that algorithms can cause, including: damage to reputation, discrimination, standardization or systematization, manipulation, lack of due process, transparency and explainability (Balkin, 2017). Today we have entered an algorithmic society, and artificial intelligence is advancing so rapidly, but the public may not realize until recently that their fate may be governed by systems they do not understand and cannot control. The covert, opaque nature of these algorithmic systems results in a lack of obvious means for dealing with or circumventing such systems when they produce unexpected, disruptive, unfair or discriminatory outcomes. As a result, various problems have arisen based on algorithms, such as algorithm discrimination, big data killing, and data gap. Relevant cases and events have attracted people’s attention and discussion. For example, Ctrip’s big data killing case (PCD, 2021), and a netizen named Father Drift published a post on the Internet “I was cut off by a member of Meituan” (Sina, 2020), which caused heated discussions on the Internet, as well as the crime risk commonly used in the United States. The evaluation algorithm COMPAS has been reported to have obvious bias (Freeman, 2016), and these incidents show that people’s rights are under threat of algorithm violation.

Data controllers use algorithm analysis to find out who is more likely to be manipulated, and how to effectively guide and control the behavior of these people, and then guide individuals to make predictable choices through algorithms. This practice greatly saves the decision-making cost for the data controller, but the process causes cumulative harm to the data subject. These cumulative harms are a side effect of algorithmic decision-making and a social cost of algorithmic activity, but this cost is borne by the information subject (Han, 2020). China is currently highly networked, and there are applications for automated decision-making in almost conceivable businesses such as food ordering, hotel booking, and flight booking. At the beginning, people felt very convenient for this, but the problems that were constantly exposed made people realize that such convenience and efficiency come at the expense of personal privacy, and the opacity of many algorithm models is violating the public’s right to know. Unfair algorithmic models put data subjects at risk of being treated unequally, and unfortunately people do not understand how these algorithmic models work against them, making it more difficult to enforce rights. Some scholars concluded that algorithm decision-making mainly faces privacy risks and discrimination risks (Zhang, 2019a). China’s huge economic size and huge network user

group determine that we will encounter more problems in the process of technological development. Compared with Europe, which has always attached great importance to human rights, it is more necessary and urgent for us to explore and establish an effective algorithm governance framework.

2.2. The Right to Explanation Was Born under Algorithmic Nuisance

Algorithms increasingly inform our lives to make corresponding decisions, but with few mechanisms to explain how they work, it is easy to cause bias, error, and discrimination. This kind of opacity and lack of comprehensibility caused by the algorithm black box allows companies to evade responsibility when making mistakes in decision-making, which makes more and more people feel uneasy and worried about the consequences of algorithmic decision-making. Today, automated decision-making systems seem to have a higher level of social and economic risk than ever before (Lu, 2020). In this context, organizations need to be subject to algorithmic supervision, and society needs more fairness, accountability, and transparency to challenge outcome bias. Therefore, the question of accountability for algorithms is naturally raised. Not surprisingly, advocates, policymakers and legal scholars are calling for machines that explain themselves to regulate algorithms. In its report “Algorithmic Decision-Making: Opportunities and Challenges”, the European Parliament stated that to achieve the comprehensibility of algorithmic decision-making, two issues are crucial, one of which is explainability (EPRS, 2019).

Data controllers make various automated decisions by collecting digital information from data subjects and then analyzing them. The internal logic of these automated decision-making systems is often opaque and unavailable to the public. The right to explanation envisioned in the General Data Protection Regulation (GDPR) constitutes an important development in this area. The emergence of the “right to explanation” is a compelling and powerful remedy because it intuitively makes it possible to open the algorithmic “black box” for greater transparency and to facilitate accountability goals. While the legal status of this right to explanation is widely debated, the truth is that today we need to trust automated decision making for many of the key decisions, so the real question should be whether the right regulation exists, not whether it should exist (Desai & Kroll, 2017). GDPR is the first law that stipulates this clause. It is necessary for us to study it and draw useful experience from it to provide ideas and reference for the formulation and implementation of relevant laws in China.

3. The Transplantation and Shortcomings of the Right to Explanation

3.1. The Origin of the Right to Explanation

3.1.1. The Reason Why the EU Established the Right to Explanation

Due to historical and cultural reasons, the European Union attaches great im-

portance to the protection of personal data. As early as 1950, the “European Convention on Human Rights” has included personal data in European human rights. It is precisely for this reason that the EU places personal data protection in a high position even at the expense of certain economic interests. Since its promulgation, GDPR has been called the strictest data protection act in history. In the regulations, the data subject is given a series of rights through the personal empowerment model. Perhaps it is precisely based on this personal empowerment model that many scholars believe that there exists right to explanation (Liang, 2020).

The GDPR passed by the EU replaced the 1995 Data Protection Directive (DPD), both of which provided for automated decision-making, but there are few cases developed around specific aspects of the law during its 23 years of existence. GDPR has renewed interest in automated decision-making provisions. The provision on automated decision-making in the original DPD was Article 15, which was originally designed to protect users from unsupervised automated decision-making. But at the time the provision did not consider dealing with the particular opacity found in complex machine learning (ML) systems, so this was changed to manage that opacity in Article 22 of the GDPR (Docherty et al., 2017), which states: (including profiles) decisions that would have legal effects relating to him or her. However, this improvement has not completely made up for the shortcomings of the previous directive, which has triggered a wide-ranging debate on the existence of the right to explanation.

3.1.2. Academic Disputes over the Right to Explanation

Goodman and Flaxman’s conference paper—EU Algorithmic Decision Regulation and “Right to explanation”—first popularized the core issue of GDPR “Right to explanation”, which sparked the debate on the right. In response to the widespread debate in the Goodman and Flaxman conference paper, Wachter et al. wrote a paper titled “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, which immediately responded to Goodman and Flaxman’s paper. The legality and technical feasibility of what Goodman and Flaxman called the GDPR’s “right to explanation” has been questioned. They insist that the GDPR does not provide for a right of interpretation, but rather a “right to know” (Wachter, 2017). In November 2017, just six months before the GDPR was enacted, the debate on the “right to explanation” quickly unfolded, with Selbst and Powles following Wachter et al. published “Meaningful Information and the Right to Explanation” joins this debate, whereby they provide a positive notion that the right to explanation lies within the text and purpose of the GDPR. They convincingly point out that right to explanation should be articulated in a functional, flexible way and should at least enable data subjects to exercise their rights (Selbst & Powles, 2017). Edwards and Weale accept the possibility of the right to explanation, but point the difficult of exercising this right from the aspect of machine learning algorithm (Edwards & Veale, 2017). Mendoza and Bygrave also support the idea of the existence of a

right to explanation, arguing that such a right can be derived from GDPR Article 22(3), especially the provisions of the GDPR do not necessarily exclude the possibility interpretation after events (Mendoza & Bygrave, 2017). Bryan Casey, Ashkon Farhangi, Roland Vogl reiterate the debate on the right to explanation of the core scholarly work and claim that the GDPR introduces an explicit “right to explanation” (Casey, Farhangi, & Vogl, 2019). Maja Brkan provides another new idea for the existence of the right to explanation, that is, several GDPR clauses can be interpreted together. Compared with other interpretation methods, this method has two advantages. First, it considers not only the wording of the clauses, but also the broader purpose of the aforementioned clauses. In particular, Wachter et al. limit themselves to the literal interpretation of the relevant provisions only in a narrow sense. European courts often depend on the purposeful or a systematic approach to interpretation, the review of the text is only the first step in interpretive work, and it is precisely the latter two interpretations that have the greatest weight in the court’s jurisprudence. Second, the joint interpretation of certain provisions follows the use of the European Court of Justice Approach. In court case law, it is not uncommon jointing different data protection clauses in order to interpret certain rights of data subjects. In light of this, they argue that the right to explanation is provided through Article 13(2) GDPR(f), 14(2)(g), 15(1)(h), 22 and recital 71. The interpretation methods should help they have the right to inform the data subjects the reasons why automated decisions have legal or significant influence on them. And Maya Bracken also demonstrated the feasibility of the right to explanation from the legal and technical aspects in a later article.

At the beginning, the debate on the right of algorithm interpretation mainly focused on the existence of the right to explanation. As a result, there were two views on the interpretation of the GDPR legal text. One party believed that the right to explanation existed and the other believed that the right did not exist. But focusing too much on the legal status of the right to explanation could drive the debate in the direction of useless and unnecessary confrontation. Later, the debate discussed in depth what the right should explain, how to interpret it, and elaborated on the content of the right to explanation. The debate on the right to explanation is deepening and advancing step by step. In the extensive debate on the right to explanation, people from all walks of life pushed the content of the right to explanation become clear and definite by expressing powerful opinions (Brkan, 2019). Not only that, there was deep thought and reflection during the discussion, which ultimately pushed the discussion in the right direction (Brkan & Bonnet, 2020).

3.2. Considerations on Transplanting the “Right to Explanation”

3.2.1. Right to Explanation Effectively Inhibits the Arbitrary Use of Automated Decision-Making

Due to the opaqueness of the internal algorithmic logic of automated decision-making, it is actually difficult for people to prove that their rights and in-

terests have been infringed due to automated decision-making. If they cannot prove the infringement, they cannot claim their rights. Therefore, the significance of the right to explanation is that it can help the data subject obtain meaningful information on the logic of making automated decisions, so that the data subject can take further measures. The rapid development of China's network technology is accompanied by the widespread application of automated decision-making. While bringing convenience, some contradictions and problems have become prominent. Moreover, with the awakening of data subjects' awareness, the conflict between the application of automated decision-making and data subjects require independent decision-making will increase day by day. The EU established the right to explanation is to protect individuals from infringement caused by automated decision-making. Obviously, China is currently facing the same problem and situation. In addition, the EU GDPR, as a model for right to explanation has its advanced legislation and superb legislative skills, which is of great reference value for us. It is precisely based on the overlap of the problems faced and the advanced nature of EU legislation that it is necessary for us to transplant the right to explanation.

3.2.2. Technical Barriers Are Not a Reason Not to Explain

Those who believe that there is no right to explanation have two main reasons for objecting. One is that the right to interpret algorithms is only explicitly mentioned in the preamble of the GDPR, but the preamble has no legal effect. The second is that the right to interpret the algorithm faces technical obstacles, which may lead to unexplainable and meaningless explanations. But in fact, the 29 working group guidelines on this issue make it clear that while the "complexity of machine learning algorithms" used in such systems can make it challenging to understand how automated decision-making processes or profiles work, this complexity It cannot be used as an excuse for not providing information to the data subject (A29WP, 2018). And many scholars have also proposed corresponding solutions, such as counterfactual explanations for automated decision-making (Wachter, Mittelstadt, & Russell, 2018), and the development of easy-to-interpret algorithm models. Edwards distinguishes between model-centered explanations and subject-centered explanations. The subject-centered explanations are more conducive to data subjects seeking relief (Edwards & Veale, 2017). The "Report on Artificial Intelligence Development 2020" believes that one of the key development directions of artificial intelligence in the next decade includes Explainable AI, which also shows that technical obstacles are expected to be solved (AITR, 2021). Although we may not be able to fully overcome this technical obstacle at present, giving up the right to seek explanation because of technical obstacle may allow the data controller to evade the obligation of such explanation by developing more complex algorithm models, which is also against with the principle of transparency established by the GDPR.

Furthermore, a technical interpretation of a machine learning algorithm model to a data subject is not necessarily the best for the data subject, and the fact

that a computer system is interpretable does not mean that it is understandable to an individual. Therefore, formal explanations have explanatory value only for computer scientists (Nisevic, 2021). The vast majority of data subjects may not fully understand technical explanations, so it is more necessary to seek effective explanations so that they can understand the reasons behind automated decision-making, in other words, to seek explanations for the entire decision-making logic.

3.2.3. The Right to Explanation—Aiming at the Return of People-Oriented

The progress of artificial intelligence brings convenience to people, but also brings about issues of ethics. As algorithmic decision-making is more and more applied to all aspects of social life, this algorithmic decision-making gradually replaces and overrides human autonomy, people are effectively controlled by an opaque decision-making system. In order to give apply to the full potential of artificial intelligence and avoid the unexpected risks that come with artificial intelligence, there is a lot of discussion about how to design, implement and manage ethical artificial intelligence. What is morality? Morality is generally defined as the moral rules and values that govern human behavior or the conduct of activities, and the principles by which these rules are evaluated. These principles can form part of essential concepts to human nature, such as human dignity. The EU attaches great importance to this kind of ethics. According to the European data protection director, EU ethics is not a substitute for obeying the law, but the basis for true compliance with the law, so as to avoid undermining trust in digital services (Tsakiridi, 2020). The establishment of the right to explanation is precisely in line with this moral value. Interpretation is considered valuable because it forces the basis for making decisions to be known for public, thereby providing data subjects with a way to question the validity and legitimacy of decisions based on these reasons. More importantly, this interpretation has some inherently important implications for individuals to understand the system to which they obey, namely respect for individual autonomy. This return to autonomy makes the right to explanation instrumental in challenging the bias and discrimination of automated decisions.

3.3. Deficiencies after Transplanting the Right to Explanation in China

Although there is no high-tech blessing, the EU has exercised unprecedented global power through its laws, and has formed a phenomenon of “Brussels effect”. Through this effect, the EU has successfully exported GDPR regulations to various countries and regions in the world. This export process is achieved through the process of unilateral regulation of globalization. The EU externalizes its laws and regulations through market mechanisms, leading to the globalization of standards. This demonstrates the EU’s ability to set high industry standards within the region and reshape global rules in the process. Without exception, the “Personal Information Protection Law” that has been passed in China

also introduces content on automated decision-making and right to explanation. Article 24 stipulates the rules for personal information processors to use personal information to make automated decision-making and Individuals have the right to require personal information processors to explain automated decisions that have a significant impact on personal rights. However, this provision is too general and abstract, and there are still many deficiencies. In contrast to the EU's regulations, its clauses on right to explanation are rich and can be understood in relation to other laws and regulations, and there are corresponding collaborative governance systems, which make its implementation more feasible. Therefore, if we want to use the right to explanation as a tool for algorithm governance, more work needs to be done to improve and enrich its connotation. After examining the discussions of foreign scholars on the right to explanation, the legal status of the right has been substantially determined. Since there exists the right to explanation, and the corresponding clauses can also be found in China's legislation, how to implement this right in practice becomes the key. The next section will focus on this issue.

4. Construction of RTE'S Implementation Logic and Path in China

4.1. Localized Interpretation of the Right to Explanation

4.1.1. The Nature of the Right to Explanation

First of all, some people think that the algorithm interpretation right is an emerging right that needs to be justified. However, some scholars believe that due to the "practical importance" of rights, rights themselves must have the ability to deal with new problems; otherwise, there is no reason to take rights seriously (Chen, 2021). Moreover, a specific right is usually a right bundle composed of several right elements, so we can actually understand the right to explanation as a right element of the personal information right. In recent years, China has gradually paid more attention to the protection of personal information. In the section "Rights of Personality" of the Civil Code, personal information rights and privacy rights are written. In addition, the newly promulgated "Personal Information Protection Law" stipulates a series of personal rights to personal information. Generally speaking, the current legislative trend in China shows that we are more inclined to classify the right to personal information as a personality right, strengthens the protection of personal information and people have the right over their own information as the subject of information. The right to explanation can actually be an indispensable part of the data subject's exercise of the right to personal information to protect his rights and interests. Only after the interpretation can the data subject know, this is the premise to ensure that the data subject takes other measures to protect his rights and interests. Therefore, from this perspective in terms of the right to explanation as a right element of the right to personal information, it is also a proper way.

Secondly, it is difficult to understand who has the right to interpret the algorithm only from the method of literal interpretation. If the data subject has the right, is the data subject interpreting it? Therefore, the name can easily lead to confusion and misunderstanding among people who do not understand the right. With an in-depth understanding of the connotation of this right, we can know that this right is the right of the data subject to seek explanations from the data controller. It is a right of request, that is, when an algorithm decision has a legal or other similar significant impact on the counterparty, the counterparty has the right to object to the algorithm controller, and can ask him to provide an explanation of the algorithm decision and then request correction of the inappropriate algorithmic decisions, etc. (Xie, 2020). Therefore, calling this right an “right to request explanation” can better express the connotation of this right.

4.1.2. Look at the Right to Explanation from the Perspective of Rights and Obligations

The right to explanation should include two aspects. From the perspective of the data subject, it has the right to request explanation, and from the perspective of the data controller, it undertakes the obligation to explain automated decision-making. Some scholars believe that the GDPR constructs a limited and weakened version of the right to explanation at the legislative level, but it is reinforced at the legal implementation level through the data subject rights and data protection impact assessment system (Zhang, 2019b). The so-called restrictions and reinforcements can actually be viewed from the perspective of rights and obligations. By carefully examining the provisions of the GDPR, we can find that the GDPR not only stipulates that data subjects have the right to interpret algorithms, but also impose positive obligations on those data controllers through articles 13(2)(f), 14(2)(g), 15(1)(h) and 22. These obligations strengthen rights of the data subject, so that the right to explanation will not become a right of display. In contrast, Article 24 of China’s “Personal Information Protection Law” only makes simple provisions on automated decision-making and the individual’s right to require information processors to explain: first, Article 24 stipulates the personal information processing rules in Chapter 2 of the “Personal Information Protection Law”. It is not stipulated in Chapter 4, Individuals’ Rights in Personal Information Processing Activities. Secondly, this article does not stipulate how the data controller should explain, which may lead to those information subjects unable to truly exercise their own right to request for explanation. It should be known that if the data subject wants to exercise the right to request algorithm interpretation, he needs to request the cooperation of the enterprise, which can only be realized by the enterprise actively fulfilling its obligations. And the provisions of this obligation are also very important in the next step to hold the company accountable. If the company fails to fulfill this obligation, then we can hold it accountable for violating the law, so that the company can truly be responsible for the conduct of personal information.

4.1.3. The Content of the Right to Explanation

A large part of the current discussion focuses on the technical interpretation of the algorithm model for the data subject, and the current task of the right to explanation is to open the algorithm black box and explain the content. But this view has two drawbacks: first, because machines learn by themselves and human intervention is mainly focused on the definition of task-specific algorithms and data used, and machines reasoning cannot compare with natural intelligence, and machines do not think like humans. Therefore, humans cannot pursue the way of thinking adopted by machines, and the results cannot be transparent. Human understanding is sacrificed in favor of engineering views, and this is a “black box,” meaning we don’t understand the results and decisions that algorithms make (Castets-Renard, 2019). From this perspective, opening the algorithmic black box and explaining it does not seem to work well. Second, too much focus on explaining the logic of machine learning models to individuals completely obscures important and priority questions: How are the rules governing automated decision-making operations formed? To understand why the rules are so, one must seek an explanation of the process behind model development, not just the model itself (Selbst & Barocas, 2018). So we have to think outside the black box to get back to this important question, when people try to explain the rationality of decisions made by algorithmic models, they are actually asking about the institutions and subjective processes behind the decisions. The European Council has also stated that data subjects should have the right to know the logic that underpins the processing of their data to make a yes or no decision, not just information about the decision itself. If the data subject is not aware of these factors, other fundamental safeguards such as the right to object and the right to appeal to the competent authority cannot be exercised effectively (ETS, 2016).

As to what the “right to explanation” is to explain, it can be obtained from the foregoing that it should not only explain the information related to automated decision-making, but also explain the logic behind the decision and the rationality of the decision-making mechanism. Be aware that with simple and clear disclosures, people may vacillate unsure how to make a choice, let alone make too many technical explanations, so people must understand the relevance of these disclosures to the final decision, so that they may have some knowledge of automated decisions about themselves to make the right decisions. In addition, giving the right to explanation is essentially to achieve accountability, and its ultimate purpose is to hold data controllers accountable for the results of their algorithmic decisions, and to provide data subjects with a basis for exercising their rights. After understanding the ultimate purpose of the right to explanation, the content of the right becomes easy to grasp.

4.2. Application of the Right to Explanation

4.2.1. The Specific Implementation Logic of the EU’s Right to Explanation

As we all know, the GDPR consists of text and a detailed explanatory preamble.

The provisions of the preamble have no direct legal effect in the EU. They are merely clarifications and interpretations of legal rules, and themselves cannot constitute such rules. While the preamble is not a legally binding provision, it is often cited as the authoritative interpretation when the GDPR is ambiguous. And the preamble contains implications that go far beyond the GDPR itself, and in many cases it reflects the compromises of the parties during negotiations. Discussions on the GDPR also frequently cite interpretive guidance issued by an organization formerly known as the Article 29 Working Group, now known as the European Data Protection Commission. The European Data Protection Commission is made up of data protection authorities across the EU (the supervisory authority responsible for enforcing the GDPR) who reach consensus on the interpretation of data protection clauses. The data protection authorities of EU member states refer to the guidance issued by the working group (Data Protection Commission) when actually implementing the GDPR, although this interpretive guidance also has no direct legal effect. However, they strongly suggest how enforcers and ultimately courts will interpret the text. Now that the GDPR is in effect, these guidelines have more function, albeit indirect effect. So, while only the text of the GDPR is legally binding, in practice, the preamble and working group guidelines also play an important role in guiding the conduct of companies. Therefore, companies and institutions that are engaged in the business of processing personal data will largely abide by the guidance of the Preamble and the Working Group Guidelines. To this, Margot E. Kaminski shows that this is exactly how GDPR is intended to work, which is largely a system of cooperative governance. The GDPR text is full of broad standards, and over time, through constant dialogue between regulators and companies, will eventually come to concrete substance, upheld by the courts. Preambles and working group guidelines, as well as numerous mechanisms ranging from formal procedures for establishing codes of conduct to informal impact assessment requirements, are part of this collaborative approach. So, when scholars argue that what is in the preamble is not law, they not only insist on the distinction between the official text and the preamble in terms of legal effect, they also ignore the fundamentally collaborative, evolving nature of the GDPR (Kaminski, 2019).

Brian Casey et al. point out that in this debate, many have largely overlooked the potentially profound change of all the changes the GDPR heralds: the new regulation giving EU data protection authorities new sweeping enforcement powers (Casey, Farhangi, & Vogl, 2019). When people focus all their attention on how to interpret what the “right to explanation” entails, it means they ignore this important fact. The biggest different between updated rights set out in the GDPR and the provisions of the original Data Protection Directive in that the GDPR gives EU data authorities enormous investigative powers and the ability to levy fines thousands of times higher than the previous maximum limit. Given these truly threatening executive powers, EU data authorities will no longer be toothless regulators, but will instead play an important role in enforcement.

While the academic discussion on the right to explanation is in full swing, these data authorities have already begun work by describing their interpretations of the right to explanation in an attempt to provide those companies with a detailed framework for compliance with the GDPR, thereby providing a powerful opinion for the GDPR's vaguely worded authorization.

Under the general guidance of the GDPR, regions across the EU can develop their own data protection regulations, such as the latest guidance issued by the UK Information Commissioner's Office (ICO): "Explaining decisions made with artificial intelligence", which aims to help organizations understand the accountability for explaining how automated decisions are made. The guide covers everything from basic explanations to how organizations can implement such a process, with key legal frameworks including the EU General Data Protection Regulation and the UK Data Protection Act 2018. The interpretation work that has been carried out in practice and the formulation and promulgation of corresponding guidelines also illustrate the current acquiescence of the right to explanation. From this perspective, the debate over the legal status of the right to explanation may have the answer, and in this debate, those scholars who believe that there exists right to explanation seem to have the upper hand.

From the above, we can conclude that the implementation logic of the algorithm interpretation right in the EU is based on the promulgated GDPR text. Article 22 provides protection and control measures for individuals not to be bound by automated decision-making, and Articles 13 to 15 provided with additional protection for data subject, based on the guidance issued by the 29 working group and the preamble, combined with GDPR Chapters 6 and 8 to give data protection agencies powerful law enforcement powers, and continuous dialogue through judicial practice and regulations, finally obtain the essence of the right to explanation.

4.2.2. Construction of the Application of Right to Explanation in China

When it comes to the right to explanation, in addition to the right of individuals to request explanations from relevant agencies, this right must also contain another aspect: how should data processors interpret them? This is the key point for individuals to exercise their right to interpret algorithms. In this regard, the "Explaining decisions made with artificial intelligence" issued by the British ICO established a series of tasks for institutions to provide guidance for them to explain: 1) Select priority description by considering domain, use case and impact on individuals; 2) Collect and preprocess data in a way that can identify explanations; 3) Build systems to ensure that relevant information can be extracted for a range of explanation types; 4) Translate basic theory for system results into usable and understandable reasons; 5) Train implementers to deploy AI systems; 6) Consider how to build and present explanations. In addition, six specific types of interpretation methods are specified: rationale interpretation, responsibility interpretation, data interpretation, fairness interpretation, security and performance interpretation, and impact interpretation. Institutions can choose appro-

priate interpretation methods according to different scenarios to ensure that data subjects receive useful information (ICO, 2020). These guidelines are highly practical and provide great practical advice for institutions to interpret to individuals affected by automated decision-making.

Therefore, in order to construct the implementation of the right to explanation in China, two aspects should be done: first, the connotation of the right should be explained, the conditions for exercising its rights should be clearly defined, and channels for data subjects to exercise their rights should be established, such as the reporting system, the appeal system, and judicial channels. Second, and even more important, is to set a series of obligations for companies, because the realization of accountability presupposes that companies violate these obligations. The data controller should implement the specific employee responsibility mechanism from the design and establishment of the algorithm model to the specific application, determine the personnel responsible for interpreting the artificial intelligence decision, and ensure that the data subject has a corresponding contact person who can promptly contact the enterprise when inquiring or questioning the decision. Provide data subjects with logically meaningful information about the processing of personal information; How data controllers explain includes: what is the function of explanation, help people understand what, what data controllers need to show, and what information will go into data interpretation scope, to provide specific guidelines for those companies. In addition, it should clarify the concept of some proper terms (such as what is automated decision-making, information subject), establish basic principles that should be followed in interpretation (such as the principle of transparency), and establish basic knowledge such as the legal framework for this right.

At present, China has transplanted the right to explanation, and in 2021, the Hangzhou Internet Court announced on its official account an online service dispute that it concluded for the first time, establishing the judgment rule that users have the right to ask the platform to make a reasonable interpretation of the algorithm logic. It shows that the existence of algorithm interpretation rights has also been recognized in practice. All these actually provide a good demonstration for the implementation of the right to explanation in China. With the refinement of relevant laws and regulations and the continuous development of judicial practice, the standards and content of the right to explanation will be more enriched and improved.

4.3. Collaborative Governance for the Application of Right to Explanation

In addition to Articles 22, 13 to 15, GDPR also stipulates many collaborative governance systems for algorithm interpretation rights, such as the data protection impact assessment system and the data protection officer system. Therefore, the GDPR adopts an extensive collaborative governance system to implement the right to explanation, and in this way, the right gradually obtains substantive

content. In fact, China also reflects this idea in the legislative process. Article 62 of China's "Personal Information Protection Law" stipulates that the national network information department will coordinate relevant departments to promote personal information protection work in accordance with this law, including formulating specific rules and standards for personal information protection. Based on this, it can be seen that there will be a large number of supporting systems and standards to be formulated in China in the future. Obviously, to realize the right to explanation, we need to formulate more standards and rules than those in the Personal Information Protection Law. Then, when formulating these specific rules, certain ideas and principles should be followed:

4.3.1. Adhere to the Principle of Contextualization

The contextualization principle was first proposed by American scholar Helen Nissenbaum, who creatively used contextual integrity as a privacy benchmark to capture the essence of the challenges brought by information technology. Scenario integrity ties adequate protection of privacy to scenario-specific norms, requiring information collection and dissemination to be tailored to a particular scenario and to comply with the governance norms in that scenario. If the integrity of the scene is compromised, then personal privacy will be violated (Nissenbaum, 2004). One of the reasons why it is difficult for people to understand their privacy rights or for companies to make wise decisions is that privacy is not easy to define, which is why Helen Nissenbaum's contribution to privacy theory is so important, and the theory has also become one of the most critical theoretical foundations for managing data usage. Many provisions in the GDPR reflect the principle of context theory and the EU Article 29 working group has issued many guidelines based on specific scenarios, such as "Opinion No. 2/2010 on Online Behavioral Advertising", "Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation", "Opinions No. 2/2017 on Data Processing at Work", etc. In the UK ICO's "interpreting decisions made by artificial intelligence" guidelines, there is a section dedicated to the principles of scenario theory, and in this section, five factors should be considered in scenarios: the domain (the settings used to make decisions), impact (how decisions effect individuals), data (what type of data is used to make decisions), urgency, audience (provide explain to sb) (ICO, 2020).

The algorithm is not a fixed object, and its properties are changed according to different scenarios and different requirements. For example, the risks of an Internet company and a coffee shop applying algorithms to process personal data are definitely different. Naturally, the obligations that data controllers need to perform in these two scenarios are also different. The construction of the legal framework for regulating algorithms should also be based on different scenarios. Therefore, the 29 Working Group has released the Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation aimed for automated decision-making. It is precisely because there is no one-size-fits-all method to explain AI-assisted decision-making, it is even more necessary to es-

establish a scenario-based principle to establish a dynamic standard so that the corresponding rules can cope with the increasingly complex and changeable situation in the future.

4.3.2. Implement the Principle of Risk Path

A very important principle in the legislative thinking of GDPR is to implement the risk path. Article 24 stipulates that, taking into account the nature, scope, context, and purpose of data processing, as well as risks to the rights and freedoms of natural persons of different degrees and sizes, the data controller shall take appropriate technical and organizational measures to ensure that data processing is in compliance with the GDPR, and these measures shall be regularly assessed and updated. This is a general provision on risk paths in the GDPR. In addition, the GDPR also stipulates a data protection impact assessment system, which is a typical representative of implementing risk paths. The data protection impact assessment system requires data controllers to conduct risk assessments before processing data, and to take certain measures based on the assessed risks (A29WP, 2016). Article 35(3) of GDPR also stipulates that data protection impact assessments are specifically required in some scenarios, including scenes based on automated data processing, including digital profiling, systematic and extensive assessment of the personal aspects of the data subject, and the assessment of legal effect or similar significant effect on people. In its report, the European Commission also sets out a series of risk-based obligations for companies (EC, 2018). Therefore, it is conceivable that both from the perspective of legislation and from the perspective of the company, they have paid certain attention to “risk”. A general static standard is established in GDPR through general regulations, and based on the application of scenario-based theory, each data controller can formulate their own dynamic standards based on the risks in different data processing scenarios, and finally achieve a transition from static to dynamic process protection effect. The design of this flexible system makes the risk path suitable for different occasions and continuously improved with the development of practice. China also pays attention to the risk path, and the formulation of the “Information Security Technology Personal Information Security Impact Assessment Guidelines” follows this principle, and stipulates the basic principles and implementation procedures of personal information security impact assessment. Therefore, the formulation of specific rules and standards for personal information protection in the future is essential to consider the principle of risk path.

4.3.3. Focus on the Establishment of System Accountability

A large proportion of scholars in the broad discussion of the right to explanation are experts in privacy, which makes the debate start from a privacy perspective, and the GDPR is also generally considered to be a law empowering personal data. This is determined by historical origins. However, the debate on the “right to explanation” has actually overshadowed the important algorithmic accountabil-

ity system established by the GDPR. The algorithmic accountability system not only includes provisions on the right to explanation, but also provisions on the principle of transparency and the basis for “consent”, and the data protection impact assessment system. The establishment of accountability means that we not only rely on the personal empowerment model established by the GDPR to protect the information rights of data subjects, but pay more attention to the supervision of enterprises. System accountability aims to make enterprises responsible for their own data processing behavior, which can prompt enterprises to establish their own internal accountability and disclosure systems, and establish a protective barrier from the source. Scholars have pointed out that as we develop comprehensive governance structures to address issues arising from the use of machine learning in decision-making, we should move beyond frameworks that rely on individuals to exercise their rights and move towards developing a systems approach to establishing and maintaining accountability, which means going beyond privacy as a lens to observe algorithmic decision governance (Gillis & Simons, 2019). This change of observation perspective is of great significance for the further discussion of right to explanation.

GDPR is currently the most comprehensive legislation on algorithmic decision-making governance. Although the debate on the right to explanation to some extent obscures the higher-order issue of system accountability, the focus on algorithmic governance originally stemmed from this debate. And the value of “explaining” in this process is that it is necessary to achieve system accountability over time. Because accountability must have reasons, and reasons must be explained, interpretation is the first step, the front end of the procedure, and the prerequisite for ensuring the realization of the accountability system.

5. Conclusion

If the right to know is a prerequisite for data subjects to exercise other data rights, then the right to explanation may be the key to unlocking the world of opaque algorithms. The right to explanation is to allow people to understand how automated decision-making affects themselves to the greatest extent in the digital age. This right reflects our call and return to personal dignity and personality when individuals have more and more digital attributes. It's more like a correction process. At present, the right to explanation faces many obstacles, so exploring how to better exercise and implement the right is a project we need to complete. Although the right to explanation is only a small part of algorithm governance, the extensive debate on the right to explanation has made people begin to deepen the research on algorithm governance. As this debate continues, people's understanding has become more and more profound and a lot of insightful cognitions have been drawn. Today, when algorithms are highly deployed, we should understand that the discussion of right to explanation is definitely not the only topic that needs to be clarified. In the future, algorithm governance requires continuous exploration and summary.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- A29WP (2016). *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is “Likely to Result in a High Risk” for the Purposes of Regulation 2016/679*.
- A29WP (2018). *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (wp251rev.01)*.
- AITR (2021). *Report on Artificial Intelligence Development 2020*.
https://www.sohu.com/a/461233515_120056153
- Balkin, J. M. (2017). Sidley Austin Distinguished Lecture on Big Data Law and Policy: The Three Laws of Robotics in the Age of Big Data. *OSLJ Ohio State Law Journal*, 78, 1217-1241.
- Brkan, M. (2019). Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and Beyond. *The International Journal of Law and Information Technology*, 27, 91-121. <https://doi.org/10.1093/ijlit/eay017>
- Brkan, M., & Bonnet, G. (2020). Legal and Technical Feasibility of the GDPR’s Quest for Explanation of Algorithmic Decisions: Of Black Boxes, White Boxes and Fata Morgana. *European Journal of Risk Regulation*, 11, 18-50.
<https://doi.org/10.1017/err.2020.10>
- Casey, B., Farhangi, A., & Vogl, R. (2019). Rethinking Explainable Machines: The GDPR’s “Right to Explanation” Debate and the Rise of Algorithmic Audits in Enterprise. *Berkeley Technology Law Journal*, 34, 143-188.
- Castets-Renard, C. (2019). Accountability of Algorithms in the GDPR and Beyond: A European Legal Framework on Automated Decision-Making. *The Fordham Intellectual Property, Media and Entertainment Law Journal*, 30, 91-137.
<https://doi.org/10.2139/ssrn.3391266>
- Chen, J. (2021). Could Rights Be New?—Two Propositions of Emerging Rights and Their Criticism. *Law and Social Development*, No. 3, 90-110.
- Desai, D. R., & Kroll, J. A. (2017). Trust but Verify: A Guide to Algorithms and the Law. *Harvard Journal of Law & Technology*, 31, 1-64.
- Docherty, C., McLean, F., & van der Merwe, D. (2017). GDPR Series: The New Data Subject Rights. *PDP Journals*, 17, 9-11.
- EC European Commission (2018). *The GDPR: New Opportunities, New Obligations*.
- Edwards, L., & Veale, M. (2017). Slave to the Algorithm? Why a “Right to an Explanation” Is Probably Not the Remedy You Are Looking for. *Duke Law & Technology Review*, 16, 18-84. <https://doi.org/10.31228/osf.io/97upg>
- EPRS (2019). *Understanding Algorithmic Decision-Making: Opportunities and Challenges*.
[http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU\(2019\)624261](http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_STU(2019)624261)
- ETS (2016). *Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data [ETS No. 108]*. Draft Explanatory Report, Paragraph 75.
- Freeman, K. (2016). Algorithmic Injustice: How the Wisconsin Supreme Court Failed to

- Protect Due Process Rights in State v. Loomis. *North Carolina Journal of Law & Technology*, 18, 75.
- Gillis, T. B., & Simons, J. (2019). Explanation < Justification: GDPR and the Perils of Privacy. *Journal of Law & Innovation*, 2, 71-99. <https://doi.org/10.2139/ssrn.3374668>
- Han, S. (2020). How Algorithms Are Equal: Establishment of Algorithmic Discrimination Review Mechanisms. *The South China Sea Law Journal*, 4, 114-124.
- IDC & Inspur (2021). *2020-2021 China Artificial Intelligence Computing Development Evaluation Report*. <https://finance.sina.com.cn/tech/2021-01-11/doc-iiznezxt1769886.shtml>
- Kaminski, M. E. (2019). The Right to Explanation, Explained. *Berkeley Technology Law Journal*, 34, 189-218. <https://doi.org/10.31228/osf.io/rgeus>
- Liang, Z. (2020). On the Exclusive Right of Algorithm: A New Choice for Correcting Algorithmic Bias. *Political Science and Law*, No. 8, 94-106.
- Lu, S. (2020). Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence. *Vanderbilt Journal of Entertainment and Technology Law*, 23, 99.
- Mendoza, I., & Bygrave, L. A. (2017). The Right Not to Be Subject to Automated Decisions Based on Profiling. In T.-E. Synodinou, P. Jougleux, C. Markou, & T. Prastitou (Eds.), *EU Internet Law: Regulation and Enforcement* (pp. 77-98). Springer.
- Nisevic, M. (2021). The “Right to an Explanation” of Automated Decision-Making Systems—Highlights of the EU Legal Landscape Referring to Explainable AI: Part 1. *Computer and Telecommunications Law Review*, 27, 29-32.
- Nissenbaum, H. (2004). Privacy as Contextual Integrity. *Washington Law Review*, 79, 119-120.
- PCD (2021). *Say No to Big Data Killing with Judicial Judgment*. http://rmfyb.chinacourt.org/paper/html/2021-07/17/content_207484.htm
- PIRI (2021). *Panorama of China’s Artificial Intelligence Industry in 2021*. <https://www.qianzhan.com/analyst/detail/220/210803-dad34c8d.html>
- Selbst, A. D., & Barocas, S. (2018). The Intuitive Appeal of Explainable Machines. *Fordham Law Review*, 87, 1085-1139. <https://doi.org/10.2139/ssrn.3126971>
- Selbst, A. D., & Powles, J. (2017). Meaningful Information and the Right to Explanation. *International Data Privacy Law*, 7, 233-242. <https://doi.org/10.1093/idpl/ix022>
- Sina, Finance & Economy (2020). *I Was Cut Off by a Member of Meituan*. <https://finance.sina.com.cn/chanjing/cyxw/2020-12-17/doc-iiznezxs7345623.shtml>
- Tang, L. (2020). The Illusory Promise of “Separating from Algorithmic Automated Decision-Making Power”. *Oriental Law*, No. 6, 18-33.
- Tsakiridi, S. (2020). AI Ethics in the Post-GDPR World: Part 1. *PDP Journals*, 20, 13-15.
- UK Information Commissioners’ Office (ICO) (2020). *Explaining Decisions Made with AI*.
- Wachter, S. et al. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, 7, 76-99. <https://doi.org/10.1093/idpl/ix005>
- Wachter, S., Mittelstadt, B., & Russell, C. (2018). Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology*, 31, 841-887. <https://doi.org/10.2139/ssrn.3063289>
- Xie, Z. (2020). Regulating Algorithmic Decision—Focusing on the Right to Explanation of Algorithm. *Modern Law Science*, No. 1, 179-193.

- Zhang, E. (2019a). Background, Logic and Structure of the Right to Explanation of Algorithmic Decision-Making in the Age of Big Data. *Legal Forum*, 34, 152-160.
- Zhang, X. (2019b). Research on Right to Explanation and Algorithm Governance Path. *Peking University Law Journal*, No. 6, 1425-1445.