THE ALGORITHM OF "BLACK BOX" AND ITS RESPONSIBILITY MECHANISM

Zhangmin-Liu[1] *and Eksiri Niyomsilp*[2]
[1]*Ph.D. Candidate, School of Management, Shinawatra University*
[2]*Ph.D. School of Management, Shinawatra University*
Corresponding Author. e-mail: 172538004@qq.com

วันที่รับบทความ (Received) 13 กุมภาพันธ์ 2565
วันที่ได้รับบทความฉบับแก้ไข (Revised) 10 มีนาคม 2565
วันที่ตอบรับบทความ (Accepted) 5 เมษายน 2565

ABSTRACT

Since 2010, a new round of artificial intelligence upsurge has emerged, with big data, powerful computing hardware and artificial intelligence algorithm as the main elements. The powerful ability of the algorithm to influence production, consumption and social life through automatic decision-making and big data, is in sharp contrast to its

incomprehensibility and lack of transparency. The algorithm "black box" is a perfect analogy of the algorithm that has been increasingly mysterious: more and more decisions are being transferred from people to algorithms, nevertheless the rules behind are invisible to most people. Therefore, to open the "black box" in the process of algorithm programming and application, to promote more responsible decision-making and conduct of algorithm development and operation enterprises, is a research fields with important practical significance under the background of current algorithm prevalence.

Keywords: Algorithm“black box”; Stakeholders engagement; Algorithmic responsibility

Introduction

A new round of artificial intelligence upsurge has sprung up around 2010, with internet big data, powerful computing hardware facilities, and artificial intelligence algorithms as key development elements. Among them, big data and hardware rely more on the overall development of relevant industries and technologies, and the algorithm has a single identifiable technical feature. Due to the strong professionalism of the algorithm, it is not easy to be clearly analyzed and interpreted, and it is more vulnerable to the subjective will of the algorithm writers. Therefore, the research of this paper will focus on the development and application field of the algorithm, integrate a variety of theories and research methods, analyze the “black box” of the algorithm and the problem of algorithm responsibility, and provide a reference for the construction of systematic algorithm responsibility.

At present, the relevant research on algorithms is mainly focused on the technical field, with only a few solutions. We try to solve the “black box” problem of algorithms from the technical perspective and pay attention to the role of algorithm application in promoting sustainable development. Most solutions are limited to the feasibility study of a certain type of technology, there is no practical technical scheme to solve the algorithm “black box” problem and improve the algorithm. In addition to the technical field, the attention to the algorithm “black box” and algorithm responsibility is mostly concentrated in the fields of law, ethics, communication, and so on. There is little research on the positioning and role of enterprises directly concerned with the algorithm development and application in the algorithm “black box” problem and the related algorithm responsibility problem. Due to the professionalism and concealment of the algorithm “black box” problem, the research on algorithm-related enterprises is not commensurate with the actual impact caused by these enterprises through the algorithm, which makes the research in this field an important research direction worthy of breakthrough.

Black box and concealment of algorithm power

"Black box" is a concept in cybernetics. As a metaphor, it refers to those unknown systems that can neither be opened nor directly observe their internal state from the outside. The "technology black box" refers to artifacts, which have been known by some people as knowledge, but others do not necessarily know (Yingchun Tao,2008). Ashby (1956: 53) systematically expounded the black box and the black box method in his introduction to cybernetics, "the black box problem occurs in electrical engineering. Give the electrician a sealed box with some input connectors on it, which can be connected with how much voltage, electric shock or any other interference at will; in addition, some output connectors can be used to observe what he can do.

In the algorithmic society, as algorithms are more and more deployed and applied to social fields such as justice, education, medical treatment, transportation, urban governance, community management and life services, power increasingly appears and exists in algorithms (Lash, 2007), and algorithms are even expressed as "algorithm authority" (beer, 2009) and "algorithm as power" (Gillespie, 2014). Because algorithm plays an increasingly important role in the exercise of power, it has become a new power broker in society (Kitchin, 2017) and a mechanism for building and realizing power (Kushner, 2013). Algorithm power is rising rapidly as a new power form.

Algorithmic power is a "double-edged sword". The proper use of algorithmic power will have a positive impact on economic and social development and create positive multi value increment for stakeholders. On the contrary, the improper use of algorithmic power will have a negative impact on stakeholders and social development. In serious cases, it will bring great damage and form a strong negative externality beyond the traditional way.This study will focus on the relationship between responsible research and innovation, stakeholder participation, meaningful transparency and algorithmic responsibility.

Methodology and Methods

By combing the relevant theories and methods, this paper puts forward the theoretical model of algorithm responsibility based on algorithm risk, designs and counts the questionnaire， Through further into quantitative models, data collection and analysis through quantitative models is further evidence that responsibility that the causal relationship between the phenomena of the conditions and algorithms and change relation, the change law of development and the responsibility to understand the algorithm, The development direction of future algorithm responsibility problem and the purpose of reasonable prediction of key problems are discussed.The research design of this paper involves four research variables: responsible research and innovation, meaningful transparency, stakeholder engagement, and algorithmic accountability.

Research Hypothesis

Three Hypothesis regarding the posited relationships between the independent and dependent variables are as follows:

Firstly, responsible research and innovation need to be introduced into algorithmic responsibility research. Make the following assumptions:

Hypothesis A: responsible research and innovation have a positive impact on algorithm responsibility.

Hypothesis B: meaningful transparency can positively regulate the relationship between responsible research and innovation and algorithm responsibility

Hypothesis C: stakeholder participation mediates the relationship between responsible research and innovation and algorithm responsibility.

The questionnaire variable scale

The questionnaire is divided into three parts. The first part is the enterprise and the basic information of individuals and working enterprises, including the number of employees and business income, as well as the positions and working years of the interviewees. The second part is the value judgment part, which requires the fillers to choose the identification degree of the statements related to each variable according to their own cognition. The third part is the enterprise practice part, which requires the fillers to choose the degree related to each variable according to the actual situation of the enterprise. The final questionnaire variable scale is as follows:

Serial number	Category	Project	Code	Measuring item
1	Responsible research And innovation RRI	Foresight	RRI1	In view of the uncertainty in the process of algorithm development and application, enterprises.The industry should implement the intervention in the early stage and establish a mechanism to prevent algorithm risks
		Reflection	RRI2	The algorithm development and application of enterprises should follow the maintenance of society.Principles of public safety and compliance with moral and ethical norms
		Response	RRI3	The adverse consequences caused by the development and application of its algorithm by enterprises, It should be responded to in time
		Inclusiveness	RRI4	Enterprises should adjust and improve algorithms, and constantly improve society and the public.Give feedback on your concerns and expectations
2	Meaningful transparency	Interpretability	MT1	An algorithm should be able to use mathematical logic symbols and natural language Some combination of words to explain

			MT2	When applying the algorithm, it is necessary to use the operation principle and mechanism of the algorithm Explain and disclose to stakeholders in an easy-to-understand way
		Verifiability	MT3	The common purpose of developing or applying the algorithm, data management, and access rights, algorithm automatic decision-making process, etc. should be
			MT4	For the final result of an algorithm to run, it is necessary to be able to Enough to deduce the complete process chain and reasoning logic in reverse
3	Stakeholder participation	Degree of trust	SHE1	In the process of algorithm development and application, the enterprise actively pays attention to, considers, and protects the rights and interests of core
		Cooperation efficiency	SHE2	In the process of algorithm development and application, enterprises work with developers,The communication between users, regulators, and other core stakeholders is harmonious
		Extensive participation	SHE3	In the process of algorithm development and application, enterprises have a wide range with the whole society
4	Algorithm responsibility	Bottom line compliance	AR1	Enterprises have never received administration for algorithms and algorithm-related products
		Meet expectations	AR2	Enterprises have established evaluation and control mechanisms for algorithms and algorithm-related products developed or applied to avoid cheating and abuse

			AR3	In the process of developing and applying algorithms and algorithm products, enterprises will adopt the knowledge and opinions of experts in the field of social sciences to ensure compliance and not
		Value co-creation	AR4	In the process of algorithm development and application, the enterprise can communicate frequently with core stakeholders such as developers, users, and regulators to protect their right to
			AR5	This enterprise has not refused to disclose it on the grounds of protecting trade secrets Risk factors in the process of algorithm development or application

Table 1 Questionnaire variable scale

Population and Sample Size

Considering the quality of sample data recovery and the breadth and diversity of sample data collection, the final questionnaires were distributed through the star line sample service and Shangden, a Shenzhen research and consulting institution. Among them, the Wenjuanxing platform distributed questionnaires to IT, industry practitioners planned to collect 200 questionnaires, 211 effective answers were actually collected, and the platform automatically and manually removed 265 invalid answers. A total of 93 questionnaires were collected from the upper study, of which 89 were valid. A total of 300 valid questionnaires were collected from the two platforms.

Data Collection Procedures

Sample Data Overview

As can be seen from Table 2, the number of employees in enterprises mainly ranges from 11 to 100, accounting for 44.4% of the total sample. The operating income of enterprises is mainly distributed between 500,000 yuan and 100 million yuan, accounting for nearly 70% of the total sample. In terms of job distribution, middle managers and IT technology-related personnel accounted for the largest proportion, accounting for nearly 60% of the total sample. In terms of the distribution of the understanding degree of the algorithm responsibility, the proportion of understanding degree and understanding degree is the largest, reaching 71.3%, which further ensures the quality of the questionnaire.

Representation attribute	Dividing standard	Sample size	Percentage
Enterprise personnel	10 persons or less	12	3.9%
	11-100 persons	133	44.4%
	101-300 persons	68	22.6%
	More than 300 people	85	28.2%
	Uncertainty	2	0.6%
Business income of enterprises	Under 500,000 yuan	18	5.9%
	500-10 million yuan	132	44.0%
	10 million-100 million	76	25.3%
	More than 100 million	58	19.4%
	Uncertainty	16	5.2%
Post distribution	Founder	5	1.6%
	Top management	23	7.6%
	Middle managers	78	26.0%
	IT-related professional and technical personnel	101	33.6%
	Ordinary employees or others	93	31.0%
Understanding of algorithm responsibility	Know very well	41	13.6%
	Have a better understanding	117	39.0%
	Moderate understanding	97	32.3%
	Less understanding	43	14.3%
	Very ignorant	2	0.6%

Table 2 Descriptive Statistical Analysis of main features (N=300)

Data Analysis Procedures
Scale Test Analysis and Factor Analysis

SPSS21.0 was used to analyze the reliability of the scale.

1.CITC and Cronbach's α reliability analysis of the scale

Firstly, CITC and Cronbach's α reliability analysis of the Responsible Research and Innovation Scale. CITC method and Cronbach'sα reliability coefficient method purify the measurement terms of the scale, and the results are shown in Tables 5-8. As can be seen from the table below, the RANGE of CITC values of items in the RESPONSIBLE Research and Innovation scale is 0.558 ~ 0.454, which is higher than the threshold of 0.30, the minimum acceptable standard. In addition, the Cronbach's α value of the scale did not increase significantly after deleting any item. The overall Cronbach'sα value of the RESPONSIBLE Research and Innovation Scale was 0.713. According to the credibility comparison table of Wu Minglong (2008), the range of 0.7~0.8 had high reliability for the level constructs. Therefore, the reliability of the scale meets the requirements of the study.

Measuring item	CITC value	The value of A after deleting the item	Cron bach's a value
RRI1	0.558	0.616	Cronbach's a=0.713
RRI2	0.457	0.676	
RRI3	0.536	0.630	
RRI4	0.454	0.680	

Table 3 CITC and reliability analysis of the RESPONSIBLE Research and Innovation Scale (N=300)

Secondly, the CITC and Cronbach's α reliability analyses of meaningful transparency scales. As can be seen from Table 5-9 below, the range of CITC value of items in the meaningful transparency scale is 0.538 ~ 0.473, which is higher than the threshold of 0.30, the minimum acceptable standard. In addition, the Cronbach's α value of the scale did not increase significantly after deleting any item. The overall

Cronbach's α value of the meaningful transparency scale was 0.717, which met the analysis requirements.

Measuring items	CITC value	The value of A after deleting the item	Cron bach's a value
MT1	0.473	0.674	Cronbach's a=0.717
MT2	0.523	0.643	
MT3	0.538	0.636	
MT4	0.485	0.666	

Table 4 CITC and Reliability analysis of the RESPONSIBLE Research and Innovation Scale (N=300)

Thirdly, CITC and Cronbach's α reliability analysis of stakeholder participation scale. As can be seen from Table 5-10 below, the range of CITC value of items in the stakeholder participation scale is 0.546 ~ 0.433, which is higher than the threshold of 0.30, the minimum acceptable standard. In addition, the Cronbach's α value of the scale did not increase significantly after deleting any item. The overall Cronbach's α value of the meaningful transparency scale was 0.676. According to the credibility comparison table of Wu Minglong (2008), the range of 0.6~0.7 was acceptable for the level construct.

Measuring items	CITC value	The value of A after deleting the item	Cron bach's a value
SHE1	0.433	0.651	Cronbach's a=0.676
SHE2	0.546	0.503	
SHE3	0.491	0.579	

Table 5 CITC and reliability analysis of stakeholder Participation Scale (N=300)

Fourthly, CITC and Cronbach's α reliability analysis of the algorithm responsibility scale. As can be seen from Table 5-11 below, the range of CITC value of items in the algorithm responsibility scale is 0.561 ~ 0.435, which is higher than the threshold of the minimum acceptable standard of 0.30. In addition, the Cronbach's α value of the scale did not increase significantly after deleting any item. The overall

Cronbach's α value of the meaningful transparency scale was 0.668, which met the requirements.

Measuring items	CITC value	The value of A after deleting the item	Cron bach's a value
AR1	0.561	0.464	Cronbach's a=0.668
AR2	0.452	0.609	
AR3	0.435	0.639	

Table 6 CITC and reliability analysis of algorithm responsibility scale (N=300)

2.Exploratory Factor Analysis (EFA) of Scales

Firstly, exploratory factor analysis of the RESPONSIBLE Research and Innovation Scale. Table 5-12 shows the KMO and Bartlett test results of exploratory factor analysis of the responsible Research and Innovation Scale. As can be seen from the table below, the KMO value is 0.729 > 0.7, the Chi-square test value of the Bartlett spherical test of sample distribution is 222.407 (when the degree of freedom is 6), and the significance level is 0.000. The statistical value of the Bartlett test is significantly different from 0. The original hypothesis of the unit correlation matrix was rejected, indicating that there were common factors among the items of the scale, which was suitable for factor analysis (Ma Qingguo, 2002).

Kaiser-Meyer-Olkin Measure of Sample Adequacy		0.729
Bartlett's Test of Sphericity		
	Approx. Chi-Square	222.407
	df	6
	Sig.	0.000

Table 7 KMO and Bartlett tests of exploratory factor analysis of responsible research and innovation (N=300)

Factor analysis was carried out on four items of the RESPONSIBLE Research and Innovation scale. Factors were extracted as the characteristic root was greater than 1. The principal component method was adopted for factor extraction, and the rotation method was the maximum variance method. The results of exploratory factor analysis were shown in Table 3-10. A factor with an eigenvalue greater than 1 was extracted, and common factor 1 explained 53.99% of the variation. According to the factor loads in Table 3-11, the factor loads of responsible research and innovation items are both greater than the threshold value of 0.5, indicating acceptable validity.

Regression Analysis

Descriptive statistics of main variables

From the correlation analysis between variables, it can be seen that there is no multicollinearity between variables. The test results of Variance Inflation Factors (VIF) show that the mean value of VIF is 1.2 and the VIF value of each variable is no more than 10, indicating that there is no multicollinearity problem of variables and subsequent regression analysis can be carried out.

Variable	Mean	S.D.	Min	Max	1	2	3	4
1. RRI	4.064	0.673	1.750	5.000	1			
2. MT	4.174	0.684	2.000	5.000	0.633	1		
3. SHE	3.806	0.752	1.667	5.000	0.409	0.429	1	
4. AR	3.790	0.789	2.000	5.000	0.353	0.438	0.518	1

Table 8 Descriptive statistics of variables and correlation analysis among variables

Common Method Deviation Test

In this study, cross-sectional survey data were used and each questionnaire was filled by the same person, which could easily lead to the problem of “common method bias” (CMV) (Podsakoff et al., 2003). To examine potential CMV problems, the study formally evaluated them using the following two steps. First of all, we followed the suggestions of previous studies in designing the questionnaire, such as conducting pretests to avoid ambiguity, minimizing the length of the questionnaire, providing clear instructions for filling the questionnaire, and ensuring the privacy of respondents. Secondly, Harman’s single factor test was used. Specifically, this study tested the factor structure of the study variables through exploratory factor analysis (EFA), as suggested by Podsakoff et al. All items of the constructs such as responsible research and innovation, meaningful transparency, stakeholder participation, and algorithm responsibility were taken as input variables. Principal component analysis was used for exploratory factor analysis. With the criterion of eigenvalue greater than 1, a total of 4 factors were extracted, which accounted for 60.28% of the total variance variation. The variance explanation degree of the first factor was 32.38%, and one factor did not explain most of the variation. In summary, the common method bias in this paper is not a serious problem.

Findings/Results

In this paper, OLS was used for basic regression analysis, and robust standard errors were used for all OLS regression results.

1. Responsible research and innovation and algorithm responsibility. The regression results are reported in column 1 of Table 4-1. Results of the model (1) showed that the regression coefficient of RRI is 0.414, which is significant at the 1% level, indicating that responsible research and innovation significantly improves the algorithm responsibility of enterprises, and hypothesis A has been verified.

	AR (1)	*SHE* (2)	*AR* (3)	AR (4)
RRI	0.414***	0.457***	0.199***	0.156*
	(0.063)	(0.059)	(0.062)	(0.080)
SHE			0.470***	
			(0.056)	
MT				0.430***
				(0.080)
*RRI*MT*				0.055
				(0.067)
Constants	2.107***	1.948***	1.191***	1.347***
	(0.260)	(0.242)	(0.258)	(0.328)
N	300	300	300	300
F-value	43.01	60.60	62.13	27.37
Adj-R2	0.122	0.164	0.288	0.195
Sobel Z			5.722***	
Sobel Z-p value			(0.000)	
Goodman-1 Z			5.701***	
Goodman-1 Z-p value			(0.000)	
Goodman-2 Z			5.744***	
Goodman-2 Z-p value			(0.000)	
The proportion of			0.519	

Table 9 Responsible research and innovation and algorithmic responsibility: effects, boundaries, and mechanisms

2. Meaningful transparent regulation was not significant. Before the formal analysis of the moderating effect, in order to reduce the possible multicollinearity, RRI and MT were decentralized according to the practice of general literature. The relevant results are listed in Column 4 of Table 4-1. Model (4) shows that the regression coefficients of RRI and MT are significantly positive, while the regression coefficients of the interaction term MTRRI * are positive, but not significant in the traditional sense, indicating that meaningful transparency does not affect the relationship between responsible research, innovation and algorithm responsibility. Let's say B is not verified.

3. According to the hypothesis above, we believe that responsible research and innovation influence the responsibility of the algorithm through the channel of stakeholder participation. Model (2) shows that the regression coefficient of responsible research and innovation is 0.457, which is significant at the 1% level, indicating that responsible research and innovation significantly improve stakeholder participation. Hypothesis C is verified.

Research Limitations

Limitations of the analytical framework. It should be noted that the object of algorithm responsibility analysis is the algorithm that is continuously improved with the growth of hardware and the improvement of technical level, so the problem of algorithm responsibility is always an open problem, that is, the framework proposed in this paper can only achieve logical self-consistency, and cannot cover all the contents of algorithm responsibility.

Conclusion

Through the theoretical analysis and construction of the previous chapter and the model quantification and empirical analysis of the previous chapter, this study initially established a research and analysis framework for “algorithmic responsibility”, an emerging interdisciplinary field of science and technology ethics and corporate social responsibility, and tested the validity of the framework through empirical analysis. Since the application of algorithms is still evolving with the improvement of computing power as a material basis and the improvement of data collection and storage means, the analysis of the responsibility of algorithms also needs to grasp the emerging technological development and ethical frontier at any time, and at the same time needs to closely combine the leading practice of algorithm development and application enterprises.

References

Ashby, W.R. (1965). *Introduction to cybernetics[M] .Translated by Zhang Lijing*, Beijing: Science Press.

Beer, D.Power (2009). through the Algorithm?Participatory Web Cultures and the Technological Unconscious[J]. *New Media &Society*, 6:pp.985-1002.

Gillespie, T. The Relevance of Algorithms[M]. in Gillespie, T., Boczkowski, P., J. and Foot, K., A (eds), (2014). *Media Technologies: Essays on Communication, Materiality, and Society(Inside Technology).* Cambridge:The MIT Press.

Lash, S. (2007). Power after Hegemony:Cultural Studies in Mutation?[J].Theory Culture & Society, (3), 55-78.

Kitchin, R. (2017). Thinking Critically About and Researching Algorithm[J]. *Information Communication & Society*, (1), 4-29.

Kushner, S. (2013). The Freelance Translation Machine:Algorithmic Culture and the Invisible Industry[J]. *New Media & Society*, (8):pp.1241-1258.

Yingchun Tao. (2008). Knowledge problem in technology -- technology black box[J]. *Forum of Association for science and technology*, (7), 54-55.