Research on big data enabling governance decision based on consistency model

Lei Qiong*

Human Resources Department, Guangdong University of Finance & Economics, Guangzhou, China

Lei Qiong, female, (1981 -), PhD, associate researcher, main research directions: big data governance decision, higher education management.

Abstract. Big data has become a new factor of production in the era of digital governance. Big data has changed the thinking and mode of decision-making, enabled governance decision-making, and created a new paradigm of decision-making. However, in the actual application process, due to the limitations of data availability, time, cost, cognitive and psychological factors, it is often inconsistent with panoramic data. How to judge and solve the inconsistency between the data subset and the data set is the premise of making scientific decision. From the perspective of "consistency", this paper constructs the basic assumptions of governance decision-making, combs and explains the solution path of consistency between the available data subset and the expected data set, expands the research perspective of measuring data similarity, and deepens the feasibility of data-driven enabling governance decision-making.

1 Introduction

With the rapid development of computer technology, big data has become a new development mode after experimental science, theoretical research and computational science. The characteristics of big data are that it doesn't care about the clutter of data, but emphasizes the amount of data; it doesn't require accurate data, but attaches importance to its representativeness; it doesn't deliberately pursue causality, but attaches importance to law summary. Big data has become a useful tool for all walks of life to see the essence from complex phenomena. [1] With the development of digital media, network information, touch media and other technologies, the application of big data in social governance is increasingly widespread, breaking through the limitations of social governance thinking, means and scope, breaking the traditional social governance model, and promoting the innovation of governance thinking. In 2015, the State Council issued the "action outline for promoting the development of big data"(GF[2015] No.50),emphasizing the deepening of research and development applications of big data and artificial intelligence. [2] In 2020, the State Council issued the opinions of the CPC Central Committee and the State Council on building a more perfect market-oriented allocation system and mechanism of factors, which juxtaposed data with traditional factors such as land, labor, capital and technology as production factors. As an important factor of production, data is of great value for enabling public administrative decision-making and improving public administrative decision-making ability. Big data

enabling governance decision-making is the era requirement of public governance development. However, due to the limited collection ability and availability of big data, comprehensive data is usually not available in the formulation of public administrative decision-making. The information asymmetry caused by the differences between big data and small data, complete set and subset affects the support ability of data for public administrative decision-making. How to reduce the difference between subset and complete set and improve the representativeness of subset to complete set has become an important issue to improve the decision-making ability and governance performance of public administration under data-driven.

2 Significance

2.1 Promote the transformation of governance thinking

In human thinking, causal thinking is one of the most habitual ways of thinking. People are used to judging the causal relationship between affairs through intuition and experience. For example, the high housing price increases the demand for land and leads to the rise of land price. Whether the land price is high or not is the result of the high or low housing price. Under the premise of limited data, traditional research pursues to accurately determine the causal relationship between transactions, so as to improve the real observation and response of the relationship between transactions. However, there are many causal relationships, such as

*The author: Lei Qiong, E-mail:36525380@qq.com.

whether a dog bite will definitely lead to rabies. This is a probability event, which requires a lot of data to verify the causal relationship between a dog bite and rabies. Big data has changed the traditional causal thinking. Through a large number of data mining and auxiliary algorithm system, big data can fully predict the colors and styles of clothes popular in the next season, and know which kind of food is most popular before the arrival of typhoon through statistical probability. In the governance work, the correlation of big data is of great value to analyze the complex dynamic relationship between transactions. For example, the increase of wage standard can improve people's sense of gain to a certain extent, but after the income reaches a certain level, happiness will not increase with the increase of wage. On the one hand, it has provided incentives for the government to raise the standard of people's income.

2.2 Innovation for effective governance

2.2.1 Get useful information quickly

Through data correlation analysis, governance decision-makers can grasp the information in a short time, track the process of public service, understand the feedback and psychology of the masses, and carry out safety warning. For example, the Internet provides an important platform for public network communication, but it also provides a convenient way for the government to understand the psychological reaction of the masses in the implementation of governance policies. Through the investigation of the public sentiment in the forum, we can understand the public's attitude and cognitive level towards public services and public events, so as to take measures or guide public opinion or make policy adjustments.

2.2.2 Improve the efficiency of public services

Big data technology strengthens the communication, interaction and cooperation between the vertical levels and horizontal parallel subjects of public administration, and breaks the information barriers between the government, the market and social organizations. Through the massive analysis data of big data management and control platform, the government integrates the administrative functions, simplifies the government service process, and realizes the integration and one-stop of government service. For example, in the process of license processing, we can build an integrated big data service platform to gather personal information: ID card, real estate information, business license, etc. on one platform, and use electronic license instead of physical license. Big data optimizes the approval process, handles daily government services through the data platform, and realizes automatic real-time verification and automatic approval. Because a large amount of information is automatically read by the system, the handler can automatically fill in the form through the platform, so as to reduce the error rate of information collection and improve the reliability of data sources.

Through the overall management of data element market, standardizing the data activities of market subjects, it provides institutional support for the cultivation of data element market.

2.3 Promote all sectors of society to implement policies together

According to modern governance theory, government, society, market and citizens are the legal subjects of social governance and have equal rights to participate in public affairs. The Third Plenary Session of the 19th CPC Central Committee deliberated and passed the decision of the CPC Central Committee on deepening the reform of Party and state organs (hereinafter referred to as the decision). The "decision" proposed that we should further promote the streamlining of administration and decentralization, improve the efficiency and fairness of resource allocation, and greatly reduce the institutional transaction costs. Big data has broken through the data barriers among multiple governance subjects such as government, society and market, created a unified data camp, simplified government functions, reduced government administrative burden, and promoted the decentralization of governance power. The opening of data can provide convenience for other governance subjects besides the government: market and social organizations. At the same time, the information of the market and social organizations can also be timely fed back to the government to realize the flow and sharing of information among the multiple governance subjects, improve the communication and coordination ability between the multiple governance subjects, change the traditional management mode of "taking all responsibility" and "doing things by oneself" of the government, return what should have been decided by the market to the market, and realize the separation of government and society and the separation of government and enterprise Separation promotes and cultivates the ability of multiple governance subjects.

2.4 Improve the scientificity of decision making

The development of big data has given birth to the digital economy, which makes the use of super scale, rich media, low-density, streaming information characteristics of big data become a reality, and the use of these data has become a new driving force for scientific decision-making. [3] The traditional decision-making paradigm has changed and spawned new decision-making methods and processes in the context of data field, formed decision-making subjects and theoretical assumptions under big data, and transformed into the decision-making paradigm of big data. [4] In the process of governance, the government improves the efficiency of data collection and decision-making and the accuracy of policy implementation by developing and applying new big data application technologies, such as smart city construction and information cloud. Especially in large-scale and important public projects,

the government can explore and promote the scientific way of linking policy implementation with process flow with the help of "Internet plus" and AI technology, thus changing the traditional decision-making mode, breaking through the way of decision-making based on human will, reducing the subjectivity of decision-making and improving the scientific nature of decision-making. Decision makers can see the planning and progress of a project through mobile terminals such as mobile app, obtain the phased progress and analyze the existing problems. Big data improves the work efficiency of decision-makers. With the support of big data, administrative decision-makers will achieve a more accurate portrait of the political ecology and turn data advantages into efficiency. The wide application of big data in government decision-making will reshape the behavior of staff, build the basic criteria for data processing, coordinate the organizational structure of public data management, and establish an evaluation mechanism based on the public data standard system.

3 CONSISTENCY OF BIG DATA ENABLING GOVERNANCE DECISION MAKING CAPABILITY

Under the big data driven, the significance of governance decision-making is to establish decision-making situation in the view of global analysis, and whether the decision makers can obtain global data has a significant impact on the accuracy of decision-making. However, most of the data sets that decision makers obtain in reality are not equal to the whole data set. This is mainly due to the objective conditions of data collection capacity. Under the influence of availability, cost, time, cognitive ability and psychology, people can get information only a part of the data set, not theoretical big data. For example, only part of the results can be queried through keyword retrieval; government managers only have limited time to read part of the Forum; decision makers only have limited practice and energy to understand a part of all public opinions. In reality, this phenomenon often occurs, even on the basis of the computer can collect all the data, but because any algorithm is based on a certain premise of analysis, the data analysis results inevitably have some tendency to summarize the data set and describe the characteristics. Especially when decision makers need to understand more dynamic psychology or personality, or need to judge, evaluate and interpret high-risk public events, or need to interpret the public service model and related relations, the subset and examples that the decision makers can master are more concrete, rich, situational and more practical description methods. Therefore, the subset selected by decision makers will be of great practical significance to reflect the whole set. However, the significance of subset is to be based on the consistency with the whole set, or whether the real case set is consistent with the whole set will become an important factor affecting the quality of decision-making. Therefore, it is very important to discuss the consistency between the subset and the whole set.

In order to further analyze the consistency between subset and complete set, we use {x} to represent complete set data and {x} to represent instance subset. Subset is a part of complete set, that is $\{x\} \in \{x\}$. When the eigenvalues of the subset {x} and the whole set {X '} are consistent, the decision made by the subset will be consistent with the governance decision based on the whole set. For example, for the decision-making of increasing the salary standard of teachers in public colleges and universities, the concerned indicators are usually the consumption index of residents in a certain area, the minimum living expenses of local employees and their supporting population, the consumer price index of urban residents, the social insurance premium and housing accumulation fund paid by individual employees, the average wage of employees, the level of economic development, employment status and other factors. However, investigators only have limited time to obtain a limited data subset from the massive data. If the characteristics of the data subset are consistent with the characteristics of the data set, then the formulation of teachers' salary standard will be scientific. With the development of computer technology, the level of data collection and feature expression is improved, such as text summary, statistical mean, feature vector, label and so on. At the same time, sound, image, video and other non structural data are also collected with the innovation of the algorithm. These methods of complete set deconstruction are of great significance to obtain the information of complete set. However, it also confirms once again that scientific decision-making will depend on the degree of information symmetry between big data, that is, the complete set of data, and small data, that is, the subset of data.

Table1: Content and problems of big data-small data

Data type	Significance	Content	Consistency problem
Big data	All eigenvectors	General semantics	Similarity, representative relationship, feature weight of subset
Small data	Part eigenvectors	Instance semantics	

4 MODEL CONSTRUCTION BASED ON CONSISTENCY

Consistency is data consistency. In a distributed system, it can be understood that the values of data in multiple nodes are consistent. At the same time, consistency means that the basic characteristics of transactions are the same, while other characteristics are similar. The consistency test of subset matrix and total matrix is compared in AHP. Building a consistency analysis model between data sets and subsets will help to solve the problem of consistency between big data and small data.

Suppose that the data set {x} represents the data set on which the decision depends and {X '} represents the subset of the data set concerned.

Hypothesis 1: when the complete set {x} and the subset {x} can infer the same parameter values, the subset {x} is similar to the complete set {x}.

Hypothesis 2: under the premise that the complete set {x} is similar to the subset {x}, the decision specified by the torus subset {x} is scientific.

Therefore, based on the hidden Markov model, we assume that the complete set X = {x1, X2 XM}; subset X '= {x'1, X'2 X'n};

In order to measure the similarity between X 'and complete set X, the kernel function is defined as:

$$k(x,x') = \sum_{\sigma}' p(x|z) p(x'|z) p(z) \qquad (1)$$

The similarity between the two is measured by judging the information of the model parameters in the obtained data. If more of the same parameters can be obtained, it is proved that there is similarity between the subset X 'and the whole set X. The data subset provides a large number of parameter values that are the same as the whole set, which indicates that they may be generated under the same parameter, and the consistency between the data subset and the whole set is high.

In this way, we can define kernel function

$$K(y,y') = q(g,y)^t f^1 q(g,y') \qquad (2)$$

and

$$F = Ex[q(g,y)q(q,y)t] \qquad (3)$$

The Fisher information matrix ensures that the value of kernel function remains unchanged when we transform the parameter vector nonlinearly. When P changes greatly relative to the parameter, it means that the parameter is the correct value, and the data set provides a lot of information about the parameter, which means that the collected data information has representative value, or that the data subset is generated under all the same parameters and has consistency.

5 CONCLUSION

Big data provides the possibility of management decision with global view, and gives birth to a new decision paradigm driven by big data. However, due to the limitation of data collection, time, cost, human cognitive limitations and other factors, governance decisions are often based on a limited subset of data. This asymmetry between data sets and subsets will affect the accuracy of governance decisions. In order to reduce the asymmetry of decision information, the parameters of the data subset should be consistent with the data set. In this paper, we discuss the meaning of the big data set from the perspective of the consistency of the form and the decision of the big data set. In the context of digital government, with the application and increasingly prominent of big data, the consistency of data subset for the whole set will be more prominent. This paper provides a mathematical model to solve the consistency of parameters between data sets and subsets, in order to solve the big data enabling management decision, insight and solve the application of big data in public administration, and enhance the value of big data driving administrative decision.

6 DISCUSSION

Digitalization has changed the old thinking and conventional mentality of the government, improved the scientific level of administrative decision-making, changed the process of public governance, optimized the supply mode of public services, and created a new paradigm of administrative decision-making. In the future, the protection of network security and privacy should become an important prerequisite for the government's digital transformation. The degree of data sharing and the protection of privacy should be defined. While strengthening the availability of data, the infrastructure of big data and artificial intelligence should be built, and the open and scientific data policy should be implemented. Of course, big data not only brings scientific decision-making, but also brings some problems, which should be paid attention to in the research. First, the protection of data security and privacy. [5] In the field of public administration, data is not only a tool for decision-making, but also has strong political significance. Some countries monitor other countries' information through we media and other devices, and disclose Smith's data, which poses a threat to political security. Second, there is a shortage of high-level information technology talents. In general, the main body of data used in administrative decision-making is usually separated. Due to the restriction of salary level within the system, most of the high-level information professionals flow into the high paid enterprises outside the system, and the government departments lack high-level information technology talents.

Acknowledgment

References

1 . Wu Hequan: big data thinking, science and society, 2014, issue 4.

2 . Circular of the State Council on printing and distributing the outline of action for the development of big data http://www.gov.cn/zhengce/content/2015-09/05/content_ 10137.htm.

3 . Hilbert, M. and Lopez, P., 2011, "The World's Technological Capacity To Store, Communicate,

and Compute Information", Sci‐ence, 332 （6025）, pp. 60~65.

4. Chen Guoqing, Zeng Dajun, et al. Decision paradigm transformation and enabling innovation in big data environment. Management world.2020 (2)

5. Zhang Chengfu, Xie Kan Kan. The government transformation and digital government in the digital age. Seeking truth.2020 (24)A. Mecke, I. Lee, J.R. Baker jr., M.M. Banaszak Holl, B.G. Orr, Eur. Phys. J. E **14**, 7 (2004).