Methodological Innovation of Public Management Research based on Big Data and Social Computing

Xingzhou Gao, Tesheng Sun, Mei Zhao and Mengyuan Li
Northwest Normal University, School of Social Development and Public Administration, Lanzhou, China

Keywords: Big Data, Public Management, Data Security, Social Calculation.

Abstract: As the scale and scope of human socio-economic activities have expanded, modern society has become a complex system with uncertainty and rapid change. The complexity of public problems and the prevalence of "governance failure" suggest that there is an urgent need to change the thinking of public management research from that of the industrial society to that of the information society. Through the characteristics of big data, such as large scale, variety, fast generation and low density, which have great value, big data analysis and social computing methods are used in combination to build big data analysis models for public management, to mine effective information and reshape public management research. While big data research serves public management, it is also important to pay attention to the possible problems of "finiteness".

1 INTRODUCTION

As social and economic prosperity moves forward, public administration is faced with many new challenges and opportunities, and the existing public management methods are often inadequate in dealing with these complex problems, and the phenomenon of public management "failure" occurs repeatedly. The advent of the information age has led to The Internet of Things, intelligence and big data are gradually coming into view and becoming an important support for society to maintain its operation.

The use of big data thinking has gradually become an integral part of public management research, and the practical results it has shown have proven to be a natural choice in public management practice. Big data provides a large amount of high-quality information for public management. By using data mining and data analysis, public management can deeply and systematically observe and grasp the complex behavior of the operation of social systems, avoiding subjective assumptions and factual distortions based on statistical data in the traditional public management process, thus effectively improving the objectivity, systematization and relevance of public management. It also helps to identify new problems in public management, so as to accurately portray complex public issues, reduce uncertainty and complexity in the process of public management, provide auxiliary decisions for public management decision makers, optimize the level of public services, and ultimately achieve scientific, accurate and intelligent public management (Zhang 2013). This is not only an expansion of the public management system, but also an important innovation of its way of thinking and behavior.

In today's society of big data and digitalization, public management should be based on major theoretical achievements and technological advances, introduce advanced public management methods, achieve innovation in research methods, research objects and research environments, strengthen multidisciplinary interaction and exchange, and promote the continuous development of public management theory and practice.

2 THE DATA SOURCE DILEMMA OF PUBLIC MANAGEMENT RESEARCH

Data is a fundamental component in carrying out public management. In the information age, the full use of big data thinking and advanced technological methods to optimize government public decision-making and improve service levels is an important part of public management research. In industrialized

Gao, X., Sun, T., Zhao, M. and Li, M.
Methodological Innovation of Public Management Research based on Big Data and Social Computing.
DOI: 10.5220/0011342500003437
In *Proceedings of the 1st International Conference on Public Management and Big Data Analysis (PMBDA 2021)*, pages 263-268
ISBN: 978-989-758-589-0

societies, as public issues often exhibit lower levels of complexity and uncertainty, then methods such as empirical analysis, statistics and metrics, and logical analysis can be used to explore and implement public management laws with universality from individual structured data and conclusions of limited size and discontinuity, relying more on personal experience and statistics for decision making due to limited data and high acquisition costs.

However, statistics do not accurately reflect the facts because they often suffer from their endogenous flaws, such as limited sample size and authenticity (Li 2015). In order to reduce uncertainty and complexity in public management research, the most effective way is to expand the sample size and type, obtain big data with complete information including public management issues, and control research errors, so big data naturally becomes an important choice for public management analysis. In the era of big data, research in the field of public administration no longer simply emphasize the accuracy of statistical data, but rather the search for cause-and-effect relationships, and quantitative analysis of complex issues is no longer just sample data, but full-scene data. For this reason, some scholars consider big data research as the "fourth paradigm" of research (Tab 1).

Table 1: The four paradigms of scientific discovery.

Scientific Paradigm	Duration	Method
Empirical	More than 1,000 years ago	Description of natural phenomena
Theoretical	Hundreds of years in the past	Use models, generalizations
Computing	The last few decades	Simulating complex phenomena
eScience	Today	Collecting data using tools, generating data using simulators, processing data using software, storing data using computers, and analyzing data

At present, the academic community has not yet formed a complete consensus on the definition of Big Data, but the different definitions given based on different perspectives all have their common features, namely the generalization of the characteristics of Big Data. Through the elaboration and summary of the key characteristic attributes of Big Data, a generally agreed definition of the concept has been formed, among which the more representative one is "5V" (Volume, Variety, Velocity, Veracity, Value). In contrast to the definition of the concept, the key to big data lies in the technical means to fully acquire and mine big data, and then explore the value of it through data analysis, so that it can bring into play social and economic benefits, and achieve the transformation from "digital" to "data-based" (Fang 2014). Big data is the life source of information, and a series of data collection, transmission, processing and application A series of technologies related to data collection, transmission, processing and application constitute Big Data processing technologies, which are collectively referred to as "Big Data technologies" for processing a large amount of structured, semi-structured and unstructured data generated in the real world with the help of non-traditional tools in order to obtain analysis and prediction results. (Fig.1).

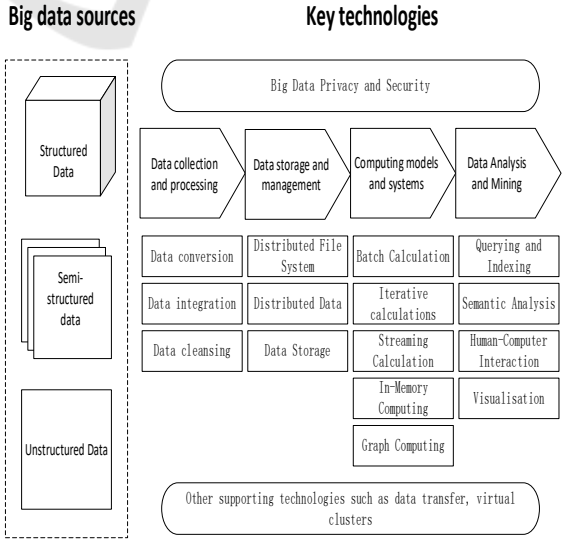

Figure 1: Big Data Technology Road Map.

3 PUBLIC MANAGEMENT IN THE PERSPECTIVE OF BIG DATA

Big Data, as a product of the information society, is a huge and diverse collection of information that can be processed to provide strong decision-making aids, scientific foresight and the ability to optimize redundant processes, majorly showing the characteristics of "big", "miscellaneous" and " correlation" (Cheng 2014). "It is proved that big data is an effective technical method for value mining and behavior pre-decision analysis of complex social systems. Complexity theory shows that even simple rules in a deterministic system can produce complex and unpredictable behaviors, and only breakthrough technologies and methods can be used to face increasingly complex public management problems. Data analysis methods, which are commonly used today, can be very useful in public management research (Fig. 2).

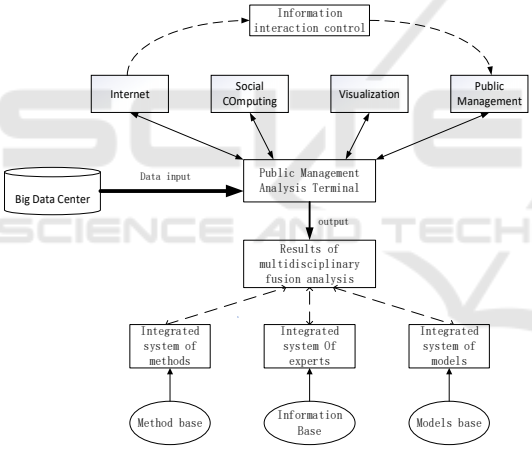

Figure 2: Big data analysis model for public management issues.

3.1 Data Visualization

Data visualization is closely related to information mapping and information visualization. The main objective of data visualization is to transform data information in order to present it in a clear and graphical way. Generally speaking, charts and maps can help people to understand information quickly and accurately based on human cognition, but when the amount of data to be represented increases to the level of big data, traditional techniques such as converting spreadsheets or pictures are no longer able to handle the huge amount of complex data, and big data visualization that can assist in algorithm design and software development has now become an active research area (Ma 2017).

3.2 Statistical Analysis

Statistical analysis techniques can be briefly divided into descriptive and inferential statistics. Descriptive statistical techniques provide a description and generalization of the overall condition of a data set, while inferential statistics enable inferences to be drawn about processes. In addition, more multivariate statistical analyses include regression, factor analysis, clustering and discriminant analysis.

3.3 Data Mining

Data mining is an important part of big data processing and analysis, mainly through specific algorithms to obtain the key information hidden in a large amount of data, but also the computational process of discovering data patterns in big data sets. At present, many representative and innovative data mining algorithms have been applied in statistics, artificial intelligence, machine learning, pattern recognition, databases, visualisation techniques and so on. The 10 most influential data mining algorithms, including C4.5, k-means, SVM, Apriori, EM, PageRank, AdaBoost, kNN, Parsimonious Bayes, were summarized at the 2006 ICDM International Conference and CART, covering the directions of classification, clustering, regression and statistical learning (Schönberg 2013). At the same time, as science and technology continue to advance, new products such as neural networks and genetic algorithms are being used to mine data for different scenarios, and the development of multidisciplinary cross-fertilisation has led to the fading of boundaries between many research methods and approaches.

In 2012, KDNuggets conducted a survey of 798 professionals on "Big Data, data mining, and data analytics software used in real-world projects in the past year" (A research study 2013). With the help of big data processing and analysis methods, a variety of public management objectives can be achieved (Tab 2). The use of big data in the social aspect promotes the change of social forms, makes social classes flow, changes the basic norms of social life and the behavior norms of the social public, forms new social characters, and becomes a new driving force of social evolution; the use of big data in the political aspect changes the traditional political ecology, promotes the rapid development of network politics and The use of big data in politics has changed the traditional

political ecology, promoted the rapid development of network politics and network democracy, and led to the change and transformation of realpolitik; in economy, it can better carry out economic micro-macro regulation and control, which is conducive to the stable and prosperous development of national economy (Liang 2013).

Table 2: Frequency of use of big data analysis tools.

Big Data Analytics Tools	
Software Name	Usage frequency
R	30.70%
Excel	29.80%
Rapid-Rapidminer	26.70%
KNMINE	21.80%
Weka/Pentaho	14.80%

4 SOCIAL COMPUTING THINKING IN PUBLIC MANAGEMENT

Modern society is a multi-level, interrelated, dynamic and dissipative complex system, showing the networked and computable complexity characteristics of "networking everywhere and computing everything", and moving from a "computable society" to a "society of From a "computable society" to a "digital society" and then to a "digital society" (Wang 2015). In reality, the most direct and effective way to address the complexity of public management is to conduct social experiments. However, due to resource, ethical, moral, psychological, cost and risk constraints, social experimentation is not a viable research method for public management (Fig. 3).

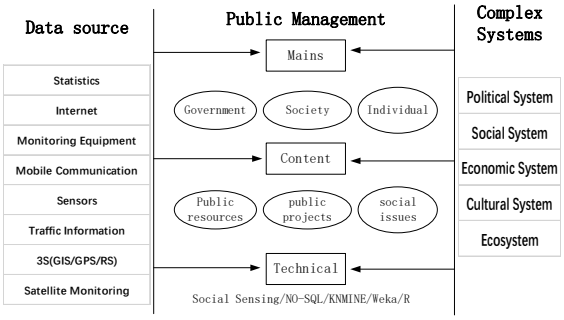

Figure 3: A social computing model for public management problems.

Computational thinking is a way of thinking that uses computer science theory to understand human social behavior through different levels of abstraction, to design artificial systems, to solve complex problems, and to form reliable solutions to social problems using mathematical models and methods. It is a way to control complex tasks or design complex social systems using abstraction and automatic computation, and it is also a general understanding and a universal skill for people to observe, analyze and solve problems in complex social forms.

In the process of understanding the laws of social development, human beings have been trying to answer the philosophical question of whether "appearance" and "phenomenon" are consistent, that is, the real relationship between the appearance we see and the objective phenomenon. For public management issues, big data is the appearance of social operation and has distinct social attributes, but public management big data is not the public issue itself. Even if we can obtain real and complete big data of public issues, we cannot achieve real responses to public issues and reveal the essential characteristics of public issues without logical analysis, value analysis and cause-and-effect analysis. The data itself does not "speak", but the multi-state public big data needs to be crawled, mined and analyzed under the guidance of theories of computational science, statistics, sociology, political science, philosophy and psychology before it can "speak" and reveal the true public values and social attributes hidden in the surface of the data. Therefore, the analysis requires the presence of the "spiritual world". At the same time, big data analysis of public management should also address the conversion of complex public management from homogeneity to heterogeneity, and should rely on "spiritual world" analysis to clarify the social context and discursive basis of the public actions measured by the data, otherwise, even the perfect excavation and correct data inference will only be the product of the researcher's discursive construction of the public issues under study. Otherwise, even if the data are perfectly mined and the inferences are correct, they will only be the product of the researcher's discourse on the public issue under study. If the "mental world" analysis leaves the field, some seemingly "accurate" big data studies may be deviated from the public issue phenomenon itself and become worthless from the very beginning.

5 SECURITY MANAGEMENT OF BIG DATA APPLICATIONS

In the field of information technology, security and privacy have always been critical issues of concern to all sectors. In the era of big data, all aspects of social management generate all kinds of data, and with the dramatic increase and accumulation of large amounts of data, the data are exposed to serious security risks, and the traditional data protection methods are no longer applicable to the rapidly changing social environment, and the security and privacy of big data are facing serious challenges (Feng 2014) .At the same time, the continuous progress of science and technology, although bringing new security risks and challenges, also brings significant opportunities for the field of information security, the development of science and technology is a double-edged sword, based on big data generated by information security technology can also be counterproductive to big data, to achieve security and privacy protection of big data (Tab 3).

Table 3: Big Data Security and Key Technologies Related to Privacy Protection.

Key technologies for big data security	
Big Data Security and Privacy Protection Technology	Big data service and information security technology
Data release anonymity protection technology	Big Data-based threat discovery technology (IBM Big Data Security Intelligence)
Social network anonymity protection technology	Big Data-based authentication technology
Data watermarking technology	Big Data-based data authenticity analysis
Data Provenance Technology	Big Data and Security-as-a-Service
Role mining	
Risk-adaptive access control	

5.1 Privacy of Big Data

In the era of big data, the privacy issue of data includes two aspects: on the one hand, the protection of personal privacy, with the development of data collection technology, in the user is not aware of, personal interests, habits, physical characteristics and other private information can be more easily accessed; on the other hand, even with the permission of the user, personal privacy data in the process of storage, transmission and use, there is a risk of being leaked. The analytic power of big data leads to the possibility that seemingly simple information can be mined for privacy, so privacy protection in the face of big data era will become a new proposition.

5.2 Data Quality

Data quality affects the utilization of big data, and low quality data not only wastes transmission and storage resources, but also cannot be used. There are many factors that affect the quality of data, which can be affected in the process of generation, acquisition, transmission and storage. Data quality is expressed in terms of accuracy, completeness, redundancy, and consistency. Although there are many measures to improve data quality, the problem of data quality cannot be completely eradicated. Therefore, there is a need to investigate a method that can automatically detect data quality and can repair some of the data with quality problems by itself.

5.3 Big Data Security Mechanisms

Big data brings challenges to data encryption in terms of data size and data variety. Previous encryption methods for small and medium scale cannot meet the requirements of big data in terms of performance, and efficient big data cryptography needs to be studied. For structured, semi-structured and unstructured data with different structures, there is a need to study how to effectively perform security management, access control and secure communication. In addition, in the multi-tenant model, it is necessary to achieve the

isolation, confidentiality, integrity, availability, controllability and traceability of tenant data while ensuring efficiency.

5.4 Big Data Applications in the Field of Information Security

Big data not only brings challenges to information security, but also injects new momentum into the development of information security. For example, through big data analysis of log files of intrusion detection systems, potential security vulnerabilities and advanced sustainability threats (APT) can be identified. In addition, information such as virus characteristics, vulnerability characteristics and attack characteristics can be more easily grasped through big data analysis. In summary, the security issue of big data has gained much attention from domestic and foreign researchers, however, the current research on the representation, metrics and semantic understanding methods of multi-source heterogeneous big data, modeling theory and computation.

6 CONCLUSION

In the era of big data, the collection, acquisition and analysis of data are faster, and these massive data will have a profound impact on human society. The application of big data to public management process is to explore the potential value from big data through the method of data analysis, and according to the different ways of data generation and structural characteristics, it can act in different areas of public management. It is worth noting that the structural complexity and meaningful complexity of big data brings the problem of complexity in social computing. No matter how much data is used, predictions inevitably encounter subjective value judgments and cannot be truly accurate, making the effect of big data analysis limited. At the same time, the connotation, technology and methods of big data application to public management are still immature, and will face a variety of problems and technical challenges in the process of its development. The technologies of efficient data storage, effective data acquisition, data analysis, data presentation and data security in big data analysis are yet to be further developed.

REFERENCES

Cheng Xueqi, Jin Xiaolong, Wang Yuanzhuo, Guo Jiafeng, Zhang Tieyuan, Li Guojie. (2014). A review of big data systems and analytic techniques[J]. Journal of Software, 25 (09).

Fang Wei, Zheng Yu, Xu Jiang. Big Data: A Review of Concepts, Technologies and Applications Research[J]. Journal of Nanjing University of Information Engineering (Natural Science Edition), 2014, 6(05): 405-419.

Feng Dengguo, Zhang Min, Li Hao. Big data security and privacy protection [J]. Journal of Computer Science, 2014, 37(01): 246-258.

Li Xuelong, Gong Hegang. (2015). A review of big data systems[J]. China Science: Information Science, 45(01): 1-44.

Liang Na, Zeng Yan. (2013). Advancing data-intensive scientific discovery to enhance science and technology innovation: new paradigms, new methods, and new challenges: the publication of the translation of The Fourth Paradigm: Data-Intensive Scientific Discovery [J]. Journal of the Chinese Academy of Sciences, 28(01): 115-121.

Ma Liang. (2017). The increasing quantification of empirical public management research: response and adaptation [J]. Xuehai, (05).

Rawat, Seema. (2014). Challenges and opportunities with big data. Routledge.

Viktor Mayer-Schönberg, Zhou T. (2013). Big changes in life, work and thinking in the era of big data[J]. Human Resource Management, (03):174.

Wang Feiyue, Wang Xiao, Yuan Yong, Wang Tao, Lin Yilun. (2015). Social computing and computational society: the foundation and necessity of smart society[J]. Science Bulletin, 60(Z1): 460-469.

What Analytics, Data mining, Big Data software you used in the past 12 months for a real project? 2012 [2013-07-11]

Zhang Yin, Chen Min, Liao Xiaofei. (2013). Current status and outlook of big data applications[J]. Computer Research and Development, 50(S2): 216-233.