Future Generation of Communication and Internet of Things (FGCIOT)

Journal homepage: http://fgciot.semnaniau.ac.ir/

PAPER TYPE (Research paper)

The Role of Metadata in Enhancing the Quality of E-Government Services: A Case Study at Islamic Azad University

Yasser Elmi Sola, Samane Ebrahimi

Department of Computer Engineering, Sab.C., Islamic Azad University, Sabzevar, Iran.

Article Info

Article History:
Received 20 April 2025
Revised 22 June 2025
Accepted 25 June 2025

Keywords:
Big Data Volume, Big Data Velocity, Big Data Variety, E-Government Services.

*Corresponding Author's Email Address:
Yasser.elmi@iau.ac.ir

Abstract

The growth and development of e-government communities is shown as a sovereign asset in developing societies, as a way for governments to employ new technologies, make people more accessible to government information and services, improve the quality of these services and provide the possibility of people's participation in democratic processes and activities. The purpose of this study is to evaluate the feasibility and importance of using metadata in the implementation of e-government services. In this research, which was conducted over a period of 6 months, 146 senior, operational and middle managers of Islamic Azad University in Khorasan Razavi province were selected through stratified random sampling. The research tool was a questionnaire whose validity was confirmed by experts and its reliability was confirmed by Cronbach's alpha coefficient. Based on partial least squares analysis (PLS) which is a model fitting based on validity and reliability analysis, the results showed that macro data volume, macro data speed, and macro data variability are based on three criteria of e-government implementation ie e-government core. , The extended layer of e-services and the outer layer of e-services based on the t-statistic, which is higher than 1.96, has a positive and significant effect.

Introduction

With big data, a wide range of information can be stored, managed and processed. Big data is a term used for a very large, diverse, complex (and difficult to store) data set. The process of researching large amounts of data to uncover hidden patterns and uncover correlations is called "big data analysis". This useful information helps organizations and companies gain deeper insights and succeed in competition. Big data is usually a collection of data that is large enough to be retrieved, managed, and processed in a reasonable amount of time by standard software.

Big data can play an important role, especially in developing countries. This article focuses on the potential of big data to be used in the developing country of Iran. It is worth noting that the security, protection and trust of users are among the challenges that exist in enjoying the benefits of big data.

Due to the importance of big data, many countries have launched their programs and initiatives on big data and its applications. Therefore, by using big data in the development of e-government and its application, long steps can be taken. On the other hand, big data in governments and especially in governments of developing countries can play an important role and reduce the knowledge and technology gap of these countries with developed countries. But its realization requires field-based studies. Governments can use big data to better serve citizens.

In this paper, in order to measure big data, three subscales of volume, speed and variability based on Patel et al. [1] questionnaire have been used. The validity and reliability of this questionnaire has been confirmed due to its standard. So that the total reliability is equal to 0.76. The method of analysis and fitting of modern utility is based on SMART PLS software. The inferential statistics of this research are measured using the analyzes related to modern validity and reliability, and considering the desirability of the modern structural framework, and based on that, the research hypotheses are tested. It is worth mentioning that the beta path coefficient and T-

Doi:

statistic have been used to test the hypotheses, which according to the obtained results, the confirmed hypotheses should be higher than 1.96. The statistical population of this study is the managers of different levels of Islamic Azad University in Khorasan Razavi province who have been selected by simple stratified random sampling method. Considering that the statistical population of this research is 235 people, Cochran's formula has been used to determine the sample size.

Big data, characterized by its vast volume, velocity, and variety, has the potential to revolutionize industries and societies. In developing countries like Iran, big data can be a powerful tool for addressing pressing challenges and accelerating development. By harnessing the insights derived from large datasets, governments can enhance public services, improve decision-making, and foster innovation.

This paper explores the potential of big data to transform e-government in Iran. We examine the specific benefits of big data, such as improved service delivery, enhanced transparency, and more efficient resource allocation. To measure the impact of big data, we employ a quantitative research methodology, utilizing a validated questionnaire and structural equation modeling (SEM) techniques.

The growth and development of e-government communities are recognized as a sovereign asset in developing societies. This serves as a means for governments to utilize new technologies, enhance public access to government information and services, improve the quality of these services, and facilitate citizen participation in democratic processes and activities. The aim of this study is to assess the feasibility and significance of employing metadata in the execution of e-government services.

This research, conducted over a six-month period, involved 146 senior, operational, and middle managers from Islamic Azad University in Khorasan Razavi province, selected through stratified random sampling. The research instrument was a questionnaire, whose validity was confirmed by experts and reliability established through Cronbach's alpha coefficient.

Using partial least squares analysis (PLS), a model fitting method based on validity and reliability assessments, the results indicated that macro data volume, macro data speed, and macro data variability significantly and positively impact three criteria for e-government implementation: the core of e-government, the extended layer of e-services, and the outer layer of e-services, as evidenced by a t-statistic greater than 1.96.

The remainder of this paper is organized as follows: Section 2 provides a literature review on big data and its applications in e-government. Section 3 outlines our research methodology, including the data collection process and statistical analysis techniques. Section 4 presents the results of our empirical analysis. Finally, Section 5 concludes the paper with a discussion of the findings and implications for policy and practice.

Related works

Big data analytics offers a plethora of opportunities to enhance the quality of services in distance education. By analyzing vast amounts of data generated from learners' interactions with learning platforms, educators can gain valuable insights to tailor educational experiences and optimize learning outcomes. Here's a comprehensive breakdown of the methods employed:

By analyzing learners' historical data, preferences, and performance, AI-powered systems can suggest tailored learning materials, exercises, and activities. These systems create customized learning journeys for each learner, adapting to their unique pace, style, and strengths. Analyzing learners' study patterns allows for the dynamic adjustment of content delivery schedules to accommodate individual preferences and optimize learning outcomes. Machine learning algorithms can forecast future learner performance based on historical data, enabling early identification of at-risk students. Proactive measures can be implemented to support struggling learners, preventing academic setbacks. By tracking learners' progress in real-time, educators can provide timely feedback and adjustments to the learning plan. Natural Language Processing (NLP) techniques can be used to analyze textual content, identifying areas for improvement in terms of clarity, relevance, and engagement. By examining learner-system interactions, educators can identify areas where learners may be struggling and design more effective learning experiences. Data analysis can reveal common challenges faced by learners, allowing educators to develop targeted interventions.

By analyzing user behavior, systems can be designed to provide a more intuitive and personalized user experience. Clear and concise information presentation, along with streamlined navigation, can minimize cognitive load on learners. Gamification elements and personalized feedback can increase learner engagement and motivation. Data analysis can help identify underutilized resources and optimize their allocation. By analyzing learner demand, educational resources can be dynamically allocated to meet varying needs. Learners can be recommended specific resources based on their individual learning goals and preferences.

By leveraging big data analytics, educational institutions can create more personalized, effective, and engaging learning experiences, ultimately improving student outcomes.

The proliferation of technology and the increasing adoption of e-learning have led to an exponential growth

in educational data. This data, often referred to as big data, presents unprecedented opportunities to enhance the quality and effectiveness of online learning. In recent years, researchers have focused on leveraging big data analytics to personalize learning experiences, improve student outcomes, and optimize educational resources.

Over the past five years, a significant body of research has explored the applications of big data analytics in e-learning. Many studies have highlighted the potential of big data to:

A. Personalize learning

By analyzing learners' interaction data, researchers have developed intelligent tutoring systems and recommendation engines that can tailor educational content to individual needs and preferences. For instance, [2] proposed a personalized learning framework that utilizes machine learning algorithms to recommend optimal learning paths for students based on their learning styles and performance.

B. Predict student performance

Predictive models have been developed to forecast student performance and identify at-risk students. These models can leverage various data sources, including historical grades, assignment submissions, and engagement metrics. For example, [3] employed a combination of machine learning techniques to predict student dropout rates in online courses.

C. Enhance instructional design

Big data analytics can inform the design of effective instructional materials and activities. By analyzing learner interactions with instructional content, researchers can identify areas where students struggle and make necessary adjustments to the curriculum. [4]demonstrated how learning analytics can be used to optimize the design of online courses.

D. Improve student engagement

Big data can be used to identify factors that contribute to student engagement and disengagement. For instance, [5] investigated the relationship between social interaction, learning activities, and student satisfaction in online learning environments.

E. Optimize resource allocation

By analyzing data on resource utilization, institutions can allocate resources more efficiently and effectively. For example, [6] explored how data analytics can be used to optimize the scheduling of online tutoring sessions.

Challenges and Future Directions

While the potential benefits of big data analytics in e-learning are significant, several challenges remain. These include data privacy concerns, the need for robust data infrastructure, and the development of advanced analytical techniques. Future research should focus on:

A. Developing more sophisticated algorithms

Researchers should continue to develop advanced machine learning and data mining techniques to extract meaningful insights from large and complex datasets.

B. Addressing ethical concerns

Ensuring the ethical use of student data is crucial. Researchers should develop guidelines and frameworks to protect student privacy and avoid biases in data analysis.

C. Integrating human expertise

While data analytics can provide valuable insights, human expertise is still essential for interpreting results and making informed decisions.

D. Scalability and interoperability

As institutions collect increasing amounts of data, it is important to develop scalable and interoperable data infrastructure.

Big data analytics has the potential to revolutionize e-learning by providing educators with the tools to create more personalized, effective, and engaging learning experiences. While significant progress has been made in recent years, there is still much work to be done. By addressing the challenges outlined above, researchers can unlock the full potential of big data analytics to improve the quality of education.

In Keoduangsine et al.'s research [7], a new model was reviewed in order to identify the most influential factors based on Ruff theory analysis and it was found that infrastructure investments are considered as the most effective factor in the development of e-government readiness. In [8] a study is presented based on qualitative analysis approaches using content analysis that created a coherent structure in decision-making level of education in school systems. In this study, three structural changes including training, evaluation and accreditation are considered as the most important educational platforms for data production in educational institutions that can change the decision-making methods in educational institutions.

In [9] an approach which was conducted in the period of one year 2017-2018, 400 managers of active companies in the field of private service sector in China participated. In this research, partial least squares analysis (PLS) method was used to fit the model. The results showed that increasing the level of dynamism in recognizing technological management styles and big data analysis in line with the dynamic capabilities of companies can help to develop and effectiveness of decision-making level and cause the level of quality services to be developed. Shonhe et al. [10] is focused on

information technology participation. The analysis method in their study was basic data analysis based on three stages of open coding, selective coding and axial coding. The results showed that the use of social media, use of update networks and websites with useful information content for consumers can increase the level of interaction of the audience.

Torrecilla et al. [11] was attended by 104 IT managers and R&D managers in the automotive industry. In this study, which was conducted over a period of one year, the method of experimental analysis using the experimental and control groups was used. One group was examined from the perspective of big data analysis theory only and one group was examined after passing a 6-month course. The results clearly showed the difference in the effectiveness of learning between the experimental and control groups, so that the group that had received practical training used more knowledge-based approaches in analyzing the problems of the organization. Patel et al.'s method [1] was based on simulation analysis, examined more than 30 scenarios at specific time intervals based on characteristics such as the type of governance, how to manage resources, the effectiveness of optimal allocation, and so on. The results showed that optimizing the data storage process in the form of a comprehensive database while increasing the speed of information flow increases the level of transparency. It was also found that big data analysis is very important to gain insight and predict the future of government governance, but extensive infrastructure investments need to be made in this regard.

In [12] a standard questionnaire was used to collect research data and participants Research 117 small company managers had an activity life of less than 5 years. The results showed that there is a direct relationship between technical training, knowledge training and human resource skills training with the effectiveness of IT implementation, but it turned out that investing in IT infrastructure in small companies, at least for the short term, does not affect the provision of effective technological services in these companies.

Study	Focus	Methodology	Key Findings
[2]	Factors influencing e-government readiness	Ruff theory analysis	Infrastructure investments are crucial for e-government development.
[3]	Decision-making in education	Qualitative content analysis	Training, evaluation, and accreditation are key for improving decision-making in education.
[4]	Big data analysis and decision-making in the service sector	Partial least squares analysis (PLS)	Increasing the level of dynamism in technological management and big data analysis can enhance decision-making and service quality.
[5]	Information technology participation	Basic data analysis (open, selective, and axial coding)	Social media, updated networks, and informative websites can increase audience interaction.
[6]	Impact of big data analysis training	Experimental analysis	Practical training in big data analysis improves knowledge-based approaches to problem-solving.
[1]	Optimizing data storage and information flow	Simulation analysis	Optimizing data storage and increasing information flow enhances transparency and enables better future predictions.
[7]	IT infrastructure and service effectiveness in small companies	Questionnaire survey	Technical training, knowledge training, and human resource skills training are more important than IT infrastructure investments for short-term service effectiveness.

Research Method

The validity of the questionnaires of this research was standardized using the opinions of experts and specialists. To evaluate the reliability of the research, Cronbach's alpha method was used according to the following equation:

$$r_{\alpha} = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum \sigma_j^2}{\sigma^2}\right) \quad (1)$$

where x represents the reliability coefficient of the whole test, k represents the number of test questions, x 2 is the variance of the scores of question j, and x 3 is the variance of the scores of the total questions. If this coefficient is greater than 0.7, the test has acceptable reliability. The Cronbach's alpha value is as shown in Table. 1. Therefore, considering that the alpha value for all variables of the questionnaire is greater than 0.7, the reliability of the questionnaire is confirmed. Table.1 shows the Cronbach's alpha coefficient of research variables.

Table 1:Cronbach's alpha coefficient of research variable

Variable	Number of questions	Cronbach's alpha coefficient
Volume	4	0.76
Velocity	4	0.73
Variety	4	0.80
The core of e-government services	5	0.92
Extended layer of e-government services	5	0.93
External layer of e-government services	4	0.89
Total number	26	0.83

In this study, stratified random sampling method was used. In this method, based on the number of each class and the percentage of the specified sample volume, the sample is determined. According to the statistics and information announced by the Ministry of Science of Iran, there are 12 branches of the Islamic Azad University in the province of Khorasan Razavi, whose total number of managers is equal to 235 people at different levels. Based on Cochran's formula, the sample size required for this study was 146 samples.

Table 2:Descriptive statistics related to education variable

Respondents education	Frequency	Valid relative frequency percentage	Valid relative CPF
Bachelor	31	21.23	21.23
Master	51	34.93	56.17
Ph.D	64	43.83	
Not answered	0	100%	100%
Total sample size	146		

According to the hypotheses presented and the statistical method used, the research findings are presented in the form of descriptive statistics tables and inferential statistics. For this purpose, first the descriptive findings are expressed, then in order to test the hypotheses, the partial least squares analysis (PLS) method is used, which is based on model fitting based on the validity and reliability. Demographic statistics are used to explain the data collected in the research. The frequency distribution of the respondents based on their education, age and managerial level is shown in Table, 2 - 4, respectively.

Table 3: Descriptive statistics related to the age variable

Age of respondents	Frequency	Valid relative frequency percentage	Valid relative CPF
Less Than 30 Years	0	0	0
Between 30 And 40 Years	78	53	53
Between 40 And 50 Years	66	45.20	98.2
More Than 50 Years	2	1.8	100%
Not Answered	0	0	
Total Sample Size	146	100%	

Examining the average of the calculated variables, it was found that the highest average is related to the big data diversity variable (4.39), which indicates that the existence of different levels of data in the form of technological software can help improve the level of analytical efficiency and cause Provide more reliable and reliable information for users. It was also found that the core of government services (4/29) is considered as one of the consequences of implementing e-government services, which can reflect the nature and philosophy of services that can be provided and the need for it in the community. The highest rate of standard deviation is related to the speed of big data, which with a deviation of about (0.95) indicates that the participants in this study had different views on the speed of big data in the analysis and service delivery, and this issue It has increased the level of standard deviation.

Table 4: Descriptive statistics related to the management level variable

Respondents' Management Level	Frequency	Valid Relative Frequency Percentage	Valid Relative CPF
Senior Managers	24	16.43	16.43
Middle Managers	79	54.10	70.53
Operational Managers	43	29.47	100%
Not Answered	0	0	
Total Sample Size	146	100%	

One of the tests used to determine the normality of data distribution is the Kolmogorov-Smirnov test, which is a non-parametric test. In this test, if the significant level of Z statistic value is greater than 5% (Sig <0.05), the statistical null hypothesis that the distribution of the variable under study is normal is accepted with 95% confidence, but otherwise the hypothesis is confirmed that States that the data is not normally distributed. Examination of the results of Kolmogorov-Smirnov test showed that the significance level of Z statistic of all research variables is significant at the 5% error level (Sig <0.05), hence the percentage assumption that the data is normal is rejected and the acceptance hypothesis is rejected.

Table 5 : Descriptive Statistics

	External service layer	Extended service layer	Government services core	Bigdata diversity	Bigdata Velocity	Bigdata volume
Average	4.22	4.19	4.29	4.39	4.15	4.28
Standard Deviation	0.91	0.91	0.75	0.74	0.95	0.78
Minimum	2.25	2.40	2.80	2.50	2.25	2.25
Maximum	5.00	5.00	5.00	5.00	5.00	5.00

According to the results of Table. 6 and considering that the structural equation modeling method with partial least squares (Smart-Pls) approach has been used to analyze the research data, and in this method, the absence of normal data distribution does not create a limit. The model is modeled in the form of validity and reliability of dimensions. It is noteworthy that the software that uses structural equation modeling based on this statistical method is compatible with conditions such as the alignment of independent variables, the normality of the data and the small size of the sample.

Table 6: Experimental results for normality of research variables**.**

Variable	Absolute Value	Positive	Negative	Z	Significant Level
External Service Layer	0.10	0.10	-0.14	1.18	0.001
Extended Service Layer	0.09	0.09	-0.15	1.07	0.000
Core Of Government Services	0.10	0.10	-0.16	1.14	0.005
Big Data Variety	0.07	0.07	-0.11	0.84	0.012
Big Data Velocity	0.07	0.07	-0.09	0.78	0.000
Big Data Volume	0.17	0.17	-0.19	2.03	0.0001

Table. 6 shows the correlation matrix of the main components of the research model. Based on this matrix, since the correlation between the model variables is less than 0.9, so the non-overlap between the research model variables in the form of differential validity is also confirmed.

Table 7: Correlation matrix for main dimension of research model.

	Big Data Volume	Government Services Core	External Service Layer	Extended S1rvice Layer	Big Data Velocity	Big Data Variety
Big Data Volume	0.84					
Government Services Core	0.52	0.91				
External Service Layer	0.63	0.47	0.88			
Extended S1rvice Layer	0.55	0.47	0.52	0.86		
Big Data	0.54	0.57	0.53	0.53	0.88	

Velocity Big Data Variety	0.46	0.48	0.48	0.50	0.53	0.85

Table 8 :Summary of the hypothesis test.

Causal Relationships Between Research Variables	Hypothesis Result
Big Data Volume Has A Significant Impact On The Core Of Big Data E-Government Services.	Confirmed
Big Data Velocity Has A Significant Impact On The Core Of Big Data E-Government Services.	Confirmed
Big Data Variety Has A Significant Impact On The Core Of Big Data E-Government Services.	Confirmed
Big Data Volume Has A Significant Effect On The Developed Layer Of Electronic Big Data Services.	Confirmed
Big Data Velocity Has A Significant Effect On The Developed Layer Of Big Data Electronic Services.	Confirmed
Big Data Variety Has A Significant Impact On The Developed Layer Of Big Data Electronic Services.	Confirmed
Big Data Volume Has A Significant Effect On The External Layer Of Big Data Electronic Services.	Confirmed
Big Data Speed Has A Significant Effect On The External Layer Of Big Data Electronic Services.	Confirmed
Big Data Variety Has A Significant Impact On The External Layer Of Big Data Electronic Services.	Confirmed

Conclusions

Big data diversity, known as the diversity of data analysis processes and their conversion into information, has a positive effect on the external layers of electronic services. Layers ranging from user satisfaction to culture and acceptance of the use of information technology also reduce overall social costs. In fact, e-government systems in Islamic Azad University have the necessary dynamism when the diversity of big data at the technological level of the university in a coherent manner and in line with the needs of analysis and information to create a level of satisfaction and satisfaction for users. Mere data analysis in the simplest possible way may not be very effective in increasing the level of social convergence of academic users and may not lead to the willingness of the community to use it effectively. Therefore, due to the sensitivities at the university level, it is necessary to provide a stronger infrastructure in this area to lead to more favorable social outcomes and increase the level of motivation.

Government use of information technology, especially the Internet, to increase the level of access of citizens, government agencies, public and private sector employees to online services and information has been referred to as the definition of e-government[13]. But it is noteworthy that communities that, even with the development of information technology in the last decade, still do not have a stable infrastructure in creating databases based on the desired features such as speed, feedback and information flow, etc., seek to create E-government can be. E-government as a communication between governments, customers and suppliers using electronic means and ICT to support government operations and service preparation, are other definitions in this area. According to another definition, e-government is defined as a way for governments to use new technologies, provide people with good access to government information and services, improve the quality of these services, and enable people to participate in democratic processes and activities. In the definition of government and governance, the provision of services is often the responsibility of the government and the decision to provide or not to provide services is often the responsibility of the government.

References

[1] Patel B, Roy S, Bhattacharyya D, Kim T-H (2017) Necessity of big data and analytics for good e-governance. International Journal of Grid and Distributed Computing 10 (8):11-20

[2] Maksimova M, Solvak M, Krimmer R Data-driven personalized e-Government services: Literature review and case study. In: International Conference on Electronic Participation, 2021. Springer, pp 151-165

[3] Altamimi MH, Aljabery MA, Alshawi IS (2022) Big Data Framework Classification for Public E-Governance Using Machine Learning Techniques. J Basrah Res(Sci) 48 (2):112-122

[4] McQuiston J, Manoharan AP (2021) E-Government and information technology coursework in public administration programs in Asia. Teaching Public Administration 39 (2):210-226

[5] Wong WH, Chapman E (2023) Student satisfaction and interaction in higher education. Higher education 85 (5):957-978

[6] Morariu C, Morariu O, Răileanu S, Borangiu T (2020) Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. Computers in Industry 120:103244

[7] Keoduangsine S, Robert G (2007) Measuring e-readiness assessment: The case of Laotian organisations. Literacy 68 (7)

[8] Zeide E (2017) The structural consequences of big data-driven education. Big Data 5 (2):164-172

[9] Shamim S, Zeng J, Shariq SM, Khan Z (2019) Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. Information & Management 56 (6):103135

[10] Shonhe L (2019) An assessment of the technology readiness of public librarians in Botswana. Global Knowledge, Memory and Communication

[11] Torrecilla JL, Romo J (2018) Data learning from big data. Statistics & Probability Letters 136:15-19

[12] Dyerson R, Spinelli R, Harindranath G (2016) Revisiting IT readiness: an approach for small firms. Industrial Management & Data Systems

[13] Layne K, Lee J (2001) Developing fully functional E-government: A four stage model. Government information quarterly 18 (2):122-136