Data Integration Governance in Public Sector: Identifying Dimensions through Preliminary Study

Noor Hasliza Mohd Hassan*, Kamsuriah Ahmad, Hasimi Salehuddin
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
43600, Bangi, Malaysia
nhasliza948@gmail.com, kamsuriah@ukm.edu.my, hasimi@ukm.edu.my

Abstract

Towards digitalization era, data plays an important role especially in public sector. Providing reliable and irredundant data are among the requirement to be filled by the government. This objective could be achieved using an efficient data integration initiative in public sector. However, to ensure the success of data integration initiative in public sector, few considerations should be taken care of. One of them is to build a comprehensive data integration governance framework comprising dimensions and factors influencing it in public sector. This study aims to gather the public sector data integration experts' knowledge in exploring the dimensions influenced the data integration initiative in public sector through a preliminary study. One of the best methods to gain tacit knowledge is through interview. Hence, a series of experts' interview was conducted to explore the possible dimensions influencing data integration governance in public sector. Seven experts were identified, and interviews were conducted in semi-structured format, based on the input of content analysis done before hand. The preliminary study has been done in six-month period. All the seven experts agreed on the three pre-determined dimensions which are; people, process and technology, with some additional recommendations. Determining the dimensions will help to further explore the factors involved and building a comprehensive data integration governance framework in public sector.

Keywords: data integration, data governance, preliminary study, interview, qualitative method

1. INTRODUCTION

In this digitalization era, governments across the world run towards the digital government adoption. Thereby, data has become the key factor that assures the successful implementation of the digital government (Hassan, Yusof, & Ahmad, 2019). However, relying on a single source of data without integrating and profiling the data is known as one of the factors of system failure in digital government (Belle et al., 2015). Data integration stimulates progress in the government's digitalization initiative by playing the role as data provider in supporting the digital government initiative. Many countries have realized that data is the key to their digitalization plan (Mohd Hassan & Ahmad, 2019).

This is supported by the studies done across the world. For example, in Russia, a research identified that without the data integration element, it is almost impossible to accomplish the "Digital Economy of Russian Federation" plan (Akatkin & Yasinovskaya, 2019). As in Malaysia, The Half Term Report of The Eleventh Malaysia Plan in 2018 asserted that the government will be emphasizing on data integration in order to ensure all relevant data will be accessible by the citizens (Malaysia, 2018). While in European Union (EU), they introduced General Data Protection Regulations (GDPR) to regulate the data usage in EU in ensuring data as an asset is being governed efficiently (Costa et al., 2018; Seabolt, Kandogan, & Roth, 2018). These examples verified the importance of data integration initiatives in public sector across the world.

———
*Corresponding author. Tel.: +60109010237
E-mail: nhasliza948@gmail.com

Hence, considering the importance of data integration in public sector, a proper data governance must be imposed (DAMA, 2017). Issues such as data duplication, accuracy, quality and accessibility which involved data integration could be solved by an extensive data governance (Alhassan, Sammon, & Daly, 2019; Mohabeer, Santally, & Sungkur, 2019; Yulfitri, 2016). This agreed by previous studies that identified the effect of deficiency of governance to the data integration implementation (Al Qassimi & Rusu, 2015; DAMA, 2017; e Vasconcelos & Brás, 2015; M.Yusof, Ismail, Ahmad, & M.Yusof, 2012; Manda, 2017; Moat KA, Wicks M, & Wilson MG, 2016; Mohabeer et al., 2019; Mohd Hassan, Ahmad, & Salehuddin, 2020; Prasser, Kohlbacher, Mansmann, Bauer, & Kuhn, 2018; Sayogo et al., 2015).

Among the governance issues arise in data integration implementation are; policies, organizational setting, political influence, trust and privacy as well as the security aspect (Manda, 2017; Yang & Maxwell, 2011). Meanwhile, there were also some concerns about the human resource and financial limitation being the governance issues in data integration implementation (Eom & Kim, 2017; Mohammed et al., 2015). From the issues collected, (Mohd Hassan et al., 2020) suggested a further study to classify the issues into a couple of dimensions to ease the problem solving. Besides, based on the previous studies, an efficient data integration governance should not only focus on certain dimensions such as technical, but also to explore non-technical dimensions as well (Eom & Kim, 2017; Fan, Zhang, & Yen, 2014; Manda, 2017; Negara, 2017). Examining the dimensions through the right sources will help to form a comprehensive data integration governance. The aim of this study is to pre-identify the dimensions influencing data integration governance in public sector in practical environment. Data will be collected from the public sector itself.

2.   METHODOLOGY

This section discusses the material and method of the study. It will begin with the determination of research question followed by the research methodology. Methodology adopted is based on the previous studies and tailored to the research question.

2.1   Research Questions

A research question has been assigned to explore the dimensions of data integration governance in public sector based on the theories and studies reviewed earlier, namely "What are the dimensions affecting data integration governance based on practical environment?".

2.2   Methods

This study adopts the qualitative method in finding the answer for the research question. A qualitative method is chosen in a study when we need to delve into the issue to find a clearer understanding on it (Creswell & Poth, 2018). Therefore, qualitative study uses the researcher as the main instrument to dig further into the issue during the data collection session in ensuring the right explanation and interpretation (Sharan B. & J.Tisdell, 2016).

In referring to the research question, which aiming to collect data in practical environment, expert interview was conducted to pre-determine the dimensions and to be matched with suitable theories for data integration governance in public sector. Interview has been recognized as one of the best methods to gain insight from the experts. It applies two ways communication that giving advantages for both sides; the informant and the researcher. On account of that, interview has been described as the main data collection method in qualitative studies as to get reliable data direct from the source (Sharan B. & J.Tisdell, 2016). The pre-determined dimensions will be verified by the content analysis of previous studies afterward.

The selection of an expert as the "informant" is dependent on their ability to comprehend the culture, and capability to synthesis and communicate the situation for the researcher (Sharan B. & J.Tisdell, 2016). Experts' interview conducted as part of triangulation process to coalesce the findings from both method, content analysis and experts review. This will enhance the trustworthiness of the results concluded (Akter et al., 2019).

2.3   Interview Phase

There are few types of interviews, namely; structured, semi-structured and unstructured interview. Semi-structured interview provides guided questions but keeps the discussion open to the panel to introduce new points (Dadzie et al., 2018; Sharan B. & J.Tisdell, 2016). The semi-structured interview is agreed as the one that have gained more attention and are commonly used (Flickr, 2019). This motivation is related to the belief that the opinions of the panel are more likely to be conveyed in an open-minded interview situation than in a structured

interview or questionnaire. Additionally, semi-structured interview has been recommended for more flexibility in the study (Guest, Namey, & Mitchell, 2013). Thus, in this study, semi-structured method is adopted during the experts' interview session to gouge more information from the experts. A number of open-ended questions are prepared within the guided theme as per Fig. 1.

There are few phases in interview protocol, regardless the type of the interview. The protocol adapted from (Azlin, Mukhtar, & Yahya, 2014) and (Creswell & Poth, 2018) as shown in Fig. 1 has been setting up in this study.
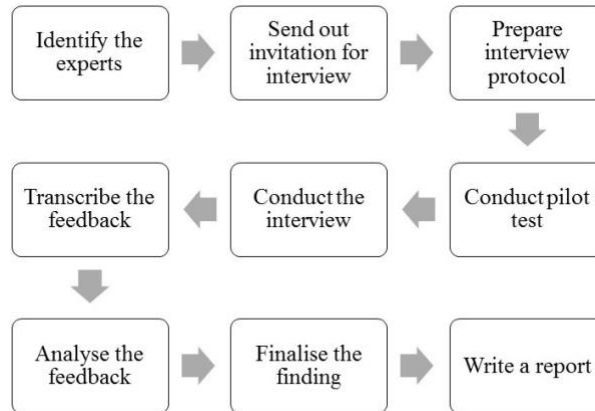

Fig. 1. Interview Phases

2.3.1 Experts Identification

Experts were identified based on their experience and knowledge in research area, which is for this study, the focus is on either data integration, data management or data governance. This expert identification criteria has been suggested and adopted by (Azlin, Mukhtar, & Yahya, 2014; Azlin, Mukhtar, & Yahya, 2018; Davis, 1992; Flick, 2009; Grant & Davis, 1997; Mohd Hassan et al., 2020; Rubio et al., 2003). Seven experts from Malaysian public sector that meet the criteria mentioned had been identified. Experts list is as per Table 1.

Table 1. Expert List

Expert ID and Job Title	Work Experience
Expert 1 Senior Assistant Principal Director of Government Data Exchange Team	20 years; 5 years of experience in data integration
Expert 2 Senior Assistant Principal Director of Government Data Exchange Team	18 years; 4 years of experience in data integration
Expert 3 Senior Assistant Director of Data Management Team	Operation-16 years; 7 years of experience in data management
Expert 4 Senior Assistant Director of Data Management Team	Operation-13 years; 8 years of experience in data integration
Expert 5 Senior Assistant Principal Director Enterprise Architecture Team	Operation-20 years; 6 years of experience in data management
Expert 6 Senior Assistant Director of Enterprise Architecture Team	Operation-13 years; 6 years of experience in enterprise architecture
Expert 7 Senior Assistant Director of ICT Technical Committee	Operation-17 years; 6 years of experience in ICT project management

Experts were invited to the interview session via email. Upon their willingness, interview dates were set up in both party's consent. The interviews were done in face to face method to get more personal contact with the panel and having the advantage of observing the emotional reaction of the panel (Flick, 2009). However, in some cases, there

would be some challenges to the researcher if the panel imposed a shy or introvert personality (Creswell & Poth, 2018). Hence, it is suggested regardless of the method of interview, the researcher as an interviewer must establish an atmosphere as relaxed as possible. In this study, as agreed by the experts, the venue of the interview chosen was in the office of each panel to provide a familiar and conducive environment for them.

2.3.2   Interview Protocol

In qualitative study, the researcher use the open-ended questions designed by the researcher themselves and do not depend on others' templates produced before (Creswell & Poth, 2018). Meanwhile, for semi-structured interview, normally theoretical-driven or hypotheses-directed questions will be asked and directed to the literature review and theoretical assumptions of the researcher (Flick, 2009). Questions imposed in the interview sessions in this study were based on the literature review done earlier. Theoretical-driven assumptions as mentioned by Flick (2009) were also included in the questions prepared.

Meanwhile, before going to the ground, the researcher is suggested to conduct a preliminary interview to test the questions and practice their interviewer's skill (Sharan B. & J.Tisdell, 2016). The trick for having good results from interviews is asking good questions. Preliminary interview could be the base to evaluate your questions. In this study, the session with panel number 1 and 2 were declared as the preliminary interviews. However, as they were giving the real data for the answers of the questions, the analysis of the sessions was included in the final report.

During the interviews with the seven panels, the researcher used an electronic recorder to record the conversation. At the same time, the researcher observed the panel's emotional and physical reaction during the interview session. Important points were jotted down to ensure the researcher did not miss the point during the transcribing session. Through the widespread use of information technology in qualitative analysis, further attention is likely to be given to the collection and preservation of qualitative data, whether using observation, transcribing or rough jotting (Creswell & Poth, 2018). Feedback from the interviews were transcribed and analysed to gather the fundamental opinion on the dimensions of data integration governance in public sector. Sample of interview protocol adapted from (Creswell & Poth, 2018) that was prepared for this study is as per in Fig. 2.

Pilot Interview

Date:

Time:

Venue:

Informant:

Job position:

Objective:
1. Exploring current situation of data integration in public sector.
2. Exploring issues in data integration governance in public sector.
3. Exploring dimensions affected data integration governance in public sector (in practical environment)

No.	Interview Question	Theme	Response
Informant's Work Experience			
	How long have you been involved in data integration?		
	What is your position in data integration team/project?		
As-is (General)			
1.	How is the current situation of data integration implementation in public sector?	General	
2.	How data integration in public sector being implemented?	General	
3.	What are the guidelines being followed for data integration implementation in public sector?	General	
4.	What are the policies being followed for data integration implementation in public sector?	General	
As-is (Dimensions)			
5.	What is the current situation/enforcement for data integration in public sector in terms of:	Issues	
a.	policy	Issues	
b.	technology	Issues	
c.	process	Issues	
d.	people	Issues	
e.	cultural	Issues	

Fig. 2. Interview Protocol Sheet

3.   RESULTS AND DISCUSSION

In this section, the findings from the experts' interview will be tabulated and discussed. A summary of interview findings on the dimensions of data integration governance is shown in Table. 2.

Table 2. Summary of Interviews' Finding on the Dimensions

Experts	Findings			
	People	Process	Technology	Others
Expert 1	✓	✓	✓	Legislation
Expert 2	✓	✓	✓	Legislation
Expert 3	✓	✓	✓	
Expert 4	✓	✓	✓	External influence
Expert 5	✓	✓	✓	Legislation
Expert 6	✓	✓	✓	
Expert 7	✓	✓	✓	

The questions which were formed in semi-structured interview were derived from the dimensions discussed earlier by Mohd Hassan, Ahmad, & Salehuddin (2019) and theory set up by DAMA (2017). Mohd Hassan, Ahmad, & Salehuddin (2019) disclosed five dimensions through the systematic literature review namely; data, environmental, organizational, policy and technology, while DAMA suggested three dimensions namely; process, people and technology. The researcher constructed the interviews' questions based on the discussion of these two content analysis findings. During the interviews, all the experts agreed that the five dimensions as mentioned in the protocol sheet could be discussed further and re-clustered to three main dimensions which are; people, process and technology. The above summary as per Table 2 was concluded by doing the analysis on the interview's transcription. All the experts agreed with the three dimensions mentioned; people, process and technology, with some additional dimensions suggested, such as legislation and external influence.

People is referred to human resource and their behavior towards data integration governance. Conjointly, people are known as the object that execute the activities using the tools and technology based on the objectives and principal (DAMA, 2017). The scope of this dimension includes organizational behavior such as cultural, leadership, communication, skills and trust. Expert 1 to 7 agreed that without a substantial people dimension, data integration will not be governed efficiently. For example, without the support from top management and good leadership skill in an integration project, data integration might succumb to failure. However, Expert 4 also highlighted the roles of vendor in data integration governance. The expert suggested that a good vendor management and negotiation skill should be included as of the component in people dimension.

Meanwhile, for process dimension, it is a identified as a set of inter related activities possessing input, procedures and specific output (Prodan, Prodan, & Purcarea, 2015). Process is automated by technology and steered by people. During the interview, all the experts conceded process includes factor such as law and regulation, policy and standardization play an important role as the pillar for data integration. This dimension should be identified and analyzed first before embarking into data integration initiative. Expert 1 and 2 also pointed out the needs of checking the law and regulation as well as the policy in process dimension upon choosing the tools and technology for data integration. They also pointed out that in a few cases of data integration failure in the public sector, the incompliance of law and policy were among the main triggers. Setting up the user requirement based on the law and policy embedded is definitely an important step to be taken before embarking into data integration initiative.

The third dimension established is technology. Tools include software or application systems and other technologies that support the objectives of the data integration initiative (DAMA, 2017). In referring to Leavitt's Diamond Model, it is mentioned that technology provides tools and strategies to interact effectively (Prodan, Prodan, & Purcarea, 2015). Technology incorporates the information management systems and as well as the design, hardware and software used. In data integration, technology is supported by people and allows process to be executed smoothly. During the interviews, all the experts agreed on technology as the other dimension spearheading data integration in public sector. They quoted that in public sector, the choice of tools and technology normally influenced by the cost allocation and the policy governed. Some agencies which do not have enough budget will go for open source software or technology smart partnership with private sectors or inter public sector agencies. However, it is crucial to deploy the best fit for purpose tools and technology to ensure the successful implementation of data integration in public sector.

4. CONCLUSION

This study managed to pre-identify the dimensions influencing data integration governance in public sector in practical environment using the experts interview technique. Through the semi-structured interview session, all the seven experts involved collectively agreed on three dimensions suggested based on theories derived from content analysis namely; people, process and technology. Other than that, the factors identified for each dimension through these interviews will be listed as the base point for further study in finalizing the factors. Determining the dimensions and factors influencing data integration governance in public sector will assist in developing a comprehensive framework for influencing data integration governance in public sector.

ACKNOWLEDGMENTS

REFERENCES

Akatkin, Y., & Yasinovskaya, E. (2019). Data-Centricity as the Key Enabler of Digital Government : Is Russia Ready for Digital Transformation of Public Sector. In *International Conference on Electronic Governance and Open Society: Challenges in Eurasia* (pp. 439–454). Springer International Publishing.

Akter, S., Bandara, R., Hani, U., Fosso, S., & Foropon, C. (2019). Analytics-based decision-making for service systems : A qualitative study and agenda for future research. *International Journal of Information Management*, *48*(January), 85–95.

Al Qassimi, N., & Rusu, L. (2015). IT Governance in a Public Organization in a Developing Country : A Case Study of a Governmental Organization. *Procedia Computer Science*, *64*, 450–456.

Alhassan, I., Sammon, D., & Daly, M. (2019). Critical Success Factors for Data Governance: A Theory Building Approach. *Information Systems Management*, *36*(2), 98–110.

Azlin, W., & Mukhtar, M. & Yahya, Y. (2014). Validating the Contents of a Social Content Management, The 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI).

Azlin, W., Mukhtar, M., & Yahya, Y. (2018). Developing the Dimensions of a Social Content Management Framework. *International Journal of Advanced Science Engineering Information Technology*, *8*(5), 1850–1857.

Belle, A., Thiagarajan, R., Soroushmehr, S. M. R., Navidi, F., Beard, D. A., Najarian, K., Arbor, A. (2015). Big Data Analytics in Healthcare, *2015*. BioMed Research International.

Costa, N. A., da Silva, J. M. V.-B., Möhring, M. M., & da Almeida, I. D. (2018). Definition of key drivers for project success regarding the General Data Protection Regulation (GDPR), *25th Annual EurOMA Conference* (pp. 1–10).

Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry & Research Design, Choosing Among Five Approaches* (Fourth Edi). SAGE.

Dadzie, J., Runeson, G., Ding, G., & Bondinuba, F. K. (2018). Barriers to adoption of sustainable technologies for energy-efficient building upgrade-Semi-structured interviews. *Buildings*, *8*(4).

DAMA, I. (2017). *DAMA-DMBOK2 Data Management Body of Knowledge*. (D. Henderson, Ed.) (2nd Editon). Technics Publications.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*(4), 194–197.

e Vasconcelos, A. F. F. C., & Brás, T. J. G. (2015). Information Reference Architecture for the Portuguese Health Sector. In *2015 IEEE 17th Conference on Business Informatics Information*.

Eom, S., & Kim, J. H. (2017). Information Sharing Success in Korean Metropolitan Governments : Combining Multi-level Factors with Fuzzy-set Analysis. In *Proceedings of dg.o conference, Staten Island, New York USA, June 2017 (dg.o'17)* (pp. 279–288).

Fan, J., Zhang, P., & Yen, D. C. (2014). Information & Management G2G information sharing among government agencies. *Information & Management*, *51*(1), 120–128.

Flick, U. (2009). An Introduction To Qualitative Research (5th Editio). SAGE.

Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health*, *20*(3), 269–274.

Guest, G., Namey, E. E., & Mitchell, M. L. (2013). *Collecting qualitative data: a field manual for applied research*. Sage Publications, Inc.

Hassan, E., Yusof, Z. M., & Ahmad, K. (2019). Factors Affecting Information Quality in the Malaysian Public Sector. *International Journal of Advanced Science Engineering Information Technology*, *9*(1), 32–38.

M.Yusof, Z., Ismail, M. B., Ahmad, K., & M.Yusof, M. (2012). Knowledge sharing in the public sector in Malaysia : a proposed holistic model. *Information Development*, *28*(1), 43–54.

Malaysia, K. H. E. E. (2018). Kajian Separuh Penggal Rancangan Malaysia Kesebelas 2016-2020.Kementerian Hal Ehwal Ekonomi.

Manda, M. I. (2017). Towards " Smart Governance " Through a Multidisciplinary Approach to E-government Integration , Interoperability and Information Sharing : A Case of the LMIP Project in South Africa. In *International Conference on Electronic Government 2017* (pp. 36–44).

Moat KA, Wicks M, & Wilson MG. (2016). *Integrating Data Across Sectors for Public Service Improvement in Ontario*. Retrieved from https://www.mcmasterforum.org/docs/default-source/product-documents/citizen-briefs/data-integration-cb.pdf?sfvrsn=2

Mohabeer, P., Santally, M. I., & Sungkur, R. K. (2019). An Investigation of the Potential Benefits of Big Data in the Public Sector of Mauritius. *Journal of the Knowledge Economy*, *10*(3), 1230–1247.

Mohammed, M. A., Maroof, E. Y., Ali, T., & Huda, I. (2015). What are the Electronic Information Sharing Factors that Influence the Participation Behavior in Higher Education Sector ? *Procedia - Procedia Computer Science*, *72*, 49–58.

Mohd Hassan, N. H., & Ahmad, K. (2019). A Review on Key Factors of Data Integration Implementation in Public Sector, (July), 9–10.

Mohd Hassan, N. H., Ahmad, K., & Salehuddin, H. (2020). Diagnosing the Issues and Challenges in Data Integration Implementation in Public Sector, *10*(2), 529–535.

Negara, J. A. (2017). *Laporan Ketua Audit Negara Tahun 2017 Siri 2*. Jabatan Audit Negara.

Prasser, F., Kohlbacher, O., Mansmann, U., Bauer, B., & Kuhn, K. (2018). *Data Integration for Future Medicine (DIFUTURE). Methods of Information in Medicine* (Vol. 57).

Prodan, M., Prodan, A., & Purcarea, A. A. (2015). Three New Dimensions to People , Process , Technology Improvement Model Three new dimensions to People , Process , Technology improvement model. *Advances in Intelligent Systems and Computing*, (March).

Rubio, D. M. G., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifyng content validity: Conducting a content validity study in social work research. *Social Work Research*, *27*(2), 94–104.

Sayogo, D. S., Zhang, J., Luna-Reyes, L., Jarman, H., Tayi, G., Andersen, D. L., … Andersen, D. F. (2015). Challenges and requirements for developing data architecture supporting integration of sustainable supply chains. *Information Technology and Management*, *16*(1), 5–18.

Seabolt, E., Kandogan, E., & Roth, M. (2018). Contextual intelligence for unified data governance. *Proceedings of the 1st International Workshop on Exploiting Artificial Intelligence Techniques for Data Management, AiDM 2018*, (3).

Sharan B., M., & J.Tisdell, E. (2016). *Qualitative Research* (4th ed.). John Wiley & Sons Inc.

Yang, T., & Maxwell, T. A. (2011). Information-sharing in public organizations : A literature review of interpersonal , intra-organizational and inter-organizational success factors. *Government Information Quarterly*, *28*(2), 164–175.

Yulfitri, A. (2016). Modeling Operational Model of Data Governance in Government. *International Conference on Information Technology Systems and Innovation*, 1–5.