

A framework design to develop integrated data system for smart e-government based on big data technology

Fahmi Dzikrullah ^{a,1,*}, Muhammad Angga Rinjani ^{b,2}

^a PT. Gamatechno Indonesia, e-Government Division, Yogyakarta, Indonesia

^b Department of Electrical Engineering and Information Technology, Engineering Faculty, Universitas Gadjah Mada, Yogyakarta, Indonesia

¹ fahmi@gamatechno.com *; ² angga.ti14@mail.ugm.ac.id

* corresponding author

ARTICLE INFO

Article history

Received September 14, 2017

Revised October 9, 2017

Accepted November 1, 2017

Keywords

e-Government

Interoperability

Integrated data system

Open government data

Big data

ABSTRACT

The rapid development of information communication and technology (ICT) in the field of governance or public service has shifted from the age of information to the age of data. The public sector is becoming increasingly aware of the potential value of data, where governments generate and collect large quantities of data (volume), rapid growth data (velocity) and various type of data (variety) through their services. Meanwhile, the government agencies keep constructing the various database, information system or application with different data sources and platform. Therefore, the interoperability has become the important requirements in electronic Government (e-Government) infrastructure that progresses towards higher levels of integration among government levels and branches. In this paper, we proposed a framework design for the development of Integrated Data System for e-Government that we call Smart e-Government, which integrates traditional information systems that is combined with an intelligent system based on big data technology. With this system, the government agencies could provide services environment with greater ease and deliver public value through open government data initiatives in a Smart e-Government context. Moreover, this would also be very useful for the government executives to monitor public services which are increasingly prestige, institutional transparency, as well as accountability. Such integration is to improve the quality of public services through information technology.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. Introduction

The information communication and technology become the backbone of e-Government to reorganize the public service agencies who are still using traditional techniques so that the quality of service to the community could be faster and easier. Governments around the world have been implementing information and computer technology on e-Government initiatives for more than ten years [1]. Electronic Government (e-Government) is the use of electronic communication devices, computers and the Internet to provide public services to citizens. The benefits and values of e-Government are to enhance the quality of public services, transparency, accountability, cost-effective service provision and government operation, reduced corruption, citizen services, optimization of public policies for better outcomes and integrated government processes. The service delivery also involves many interrelated parties. It requires integration of the various related activities to the delivery of services.

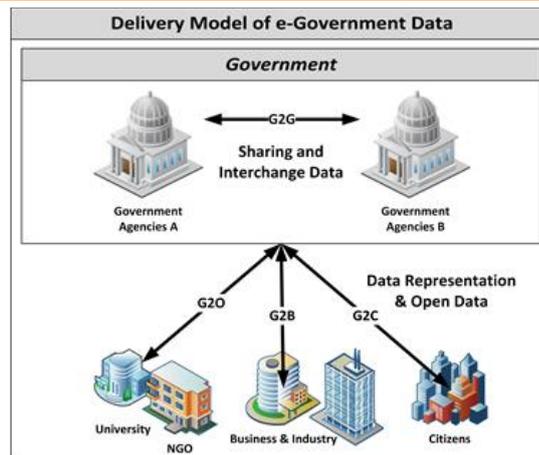


Fig. 1.e-Government delivery model

According to J.C Hai [2], the term e-Government consists of the digital interactions delivery model; the e-Government delivery model is described in Fig. 1, namely:

1. Governments to other Government agencies (G2G), the purpose is “to reduce cost by reducing paper clutter, staffing cost, or communicating with private citizens or public government”.
2. Government to internal Employees (G2E), which are “online tools, sources, and articles that help employees maintain communication with the government and their own companies”.
3. Government to Citizens (G2C), offers “variety of ICT services to citizens in an efficient and economical manner in order to strengthen the relationship between government and citizens using technology”.
4. Government to Businesses (G2B), is a “non-commercial interaction between local or central government and the commercial business sector with the purpose of providing businesses information and advice on e-business 'best practices”.
5. Government to Organization (G2O), is a “non-commercial interaction between non-government organizations such academic institutions”.

Currently, the rapid development of information technology in the e-Governments has shifted from the age of information to the age of data. The public sector is becoming increasingly aware of the potential value of data, where governments generate and collect large quantities of data (volume), rapid growth data (velocity) and various type of data (variety) through their services. Meanwhile, the government agencies keep building various applications with different data sources, programming platform, operating system, hardware, and network specification. Therefore, the interoperability has become the important requirement in e-Government infrastructure. That progresses towards higher levels of integration among government levels and branches. The interoperability is described in Fig. 2.

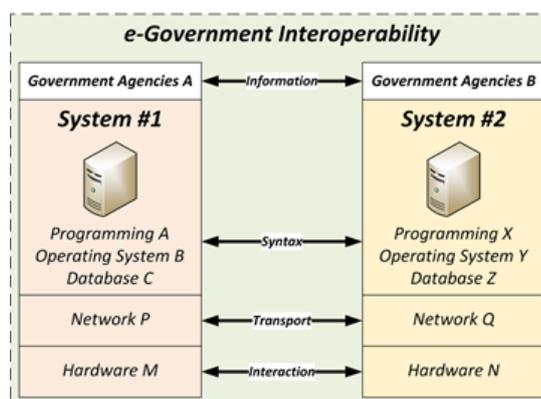


Fig. 2.e-government interoperability

In this paper, the use of information technology is to develop a platform that aims to integrate information systems by designing framework of integrated data systems based on big data technology for e-Government, called Smart e-Government, which integrates traditional information system of government agencies and an intelligent system based on big data technology. Through this system, the government agencies could provide services environment with greater ease and deliver public value through Open Government Data (OGD) initiatives in a Smart e-Government context. Moreover, this is also very useful for the government executive to monitor public services which are increasing prestige as well as institutional transparency and accountability. The integration of information technology is to improve the quality of public services through information technology.

2. Related Research

Designing and implementing an e-Government framework intend to simplify contacts between government bureaucracy and citizens, because it is of this kind of organizations that implement technologies with bureaucracy, with the aim at enhancing public administration function and service to citizens; the technology generates the new horizon for a data-driven economy [3]. The public sector is becoming increasingly aware of the potential value of data, Fahmi D. et al. governments generate and collect large quantities of data (volume), rapid growth data (velocity) and various type of data (variety) through public services. However, implementation of e-Government in developing countries as Indonesia is easier said than done. Some studies [4], [5], have already been done to investigate challenges and strategies in implementing a local e-Government; then it was found the gap raises when the government agencies are trying to scale up the infrastructure. Proposed study about infrastructure and design for integrated SKPD in Manado, Indonesia. Meanwhile, the agencies keep constructing various database and application with a different platform for the sake of agencies' sectoral ego that could lead to failure in implementing an e-Government [6]. Therefore, the interoperability has become the important requirement in e-Government infrastructure to extensive information sharing among governmental entities [7]. According to IEEE Glossary [8] "Interoperability is the ability of two or more systems or components to exchange information and to use the information that has been exchanged. Review of several models of interoperability in e-Government services [9], and the proposed study about SOA-Based Approach for e-Government Interoperability to solve the traditional problem for the integrating system in public services platform [10]. Furthermore, it could integrate with Open government and data sharing and intelligent system [11]. Data analytics becomes the fundamental of an intelligent system, by examining large amounts of data to uncover hidden patterns, the goal is to turn data into information, and information into insight, with today's technology, it is possible to analyze large amounts of public service data and integrated with intelligent or machine learning system. In our previous research, we implemented a machine learning system with Hadoop platform using Smart Card Automatic Fare Collection System (SCAFCS) dataset on public transportation system to analyze passengers temporal pattern and provide insight information to support for public transportation management [12]. The purpose of providing information from organizations to citizens is to gain trust and transparency between citizens and government, in line with open data initiatives [13]. The open data initiatives facilitate data-driven decision making, so that all e-Government delivery models could be accomplished. In this paper, we try to fill all technological gaps above by proposing a framework design for the development of Integrated Data System for Smart e-Government, which integrates traditional e-Government information system and an intelligent system based on big data technology.

3. System Framework

3.1. The Interoperabilities in e-government

e-Government systems grow so rapidly and generate large of data. Interoperability among the e-Government systems is very important for connecting and transmitting data between several systems. So that, the interoperability has become the critical requirement in e-Government infrastructure. Based on paper review [9], there are several types of interoperability in e-Government, namely:

1. e-Government interoperability using Semantic Oriented Architecture
2. e-Government interoperability based on interlinking application layer
3. e-Government interoperability using data integration.

4. e-Government interoperability based on ontology model

According to European Commission definition, interoperability is the ability of information and communication technology (ICT) systems to support exchange data of business processes and will to enable sharing of information and knowledge. There are three aspects of interoperability are identified, namely:

1. Interoperability of organizational, an aspect which deals with cases where the organizations of cooperate have differences in structure and their business processes. Therefore, it needs interoperability solution in the organizational side.
2. Interoperability of services, an aspect where the information exchanged between organizations is interpreted differently by each side of systems in a service. Therefore, it case needs interoperability to solve the different interpretation of information.
3. Interoperability of semantic, this issues about how to connect various computer systems and services. Other problem of this aspect such as interconnection, not close interfaces, data integration and middleware, data presentation and exchange, accessibility. Then, another problem in this aspect is security of the services [14].

According to related research [14], interoperability model consists of several government agencies. The model of the interoperability approach is described in Fig. 3.

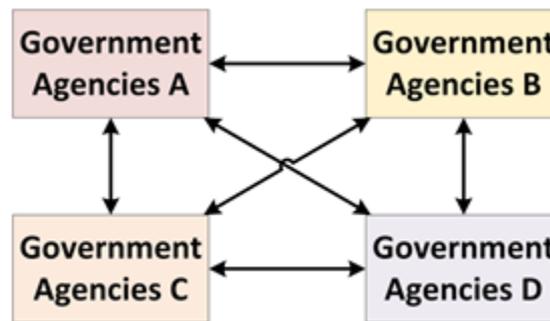


Fig. 3. Bilateral model interoperability [14]

This model is called SmartGov, the most straightforward approach towards achieving interoperability. The basic idea behind this model is that a direct connection to send e-Government transaction services from each government agencies to all other government agencies that it needs to interoperate, this model called SmartGov, an approach for interoperability model that of direct bilateral or multilateral communication between government agencies is preferred when implementing e-Government services. The approachment can be implemented and this is the better way to solve interoperability challenges. However, this model which is described in the research is limited only number of interoperability patterns and implementation models. In that case, e-Government information systems are sociotechnical and not purely technical systems. Therefore interoperability challenges related to legal, management, cultural, ethical and other social issues should be more deeply investigated [14].

In 2008, Ministry of Information and Telecommunication (MCIT) of Indonesia published interoperability guideline [15]. Current situation of e-Government implementation shows that government agencies implement an information system, but it is not interoperable with other agencies. This happens because ICT solution depends on a proprietary product. Shared services become a challenge in the situation. The government agencies cannot deliver their service effectively. Thus, the government mandates interoperability guidelines to get benefit from government agencies.

1. Data management and access delivery become easier.
2. The government agencies can deliver their service effectively.
3. The information system can provide accurate information decision making.
4. Sharing information among different agencies becomes possible.

Based on the problems above, MCIT developed an application to enable data exchange between various government agencies called MANTRA (Management of Information Integration and Data Exchange). MANTRA acts as an information system interoperability framework that integrates and provides data exchange through the Internet using web service technology based on Service Oriented Architecture (SOA) and Government Service Bus (GSB). This framework is to provide integrated information system securely, accurately, and efficiently.

Web services are developed as an API (Application Programming Interface). The API connects to applications using RPC (Remote Procedure Call) technology. The API output is in the standard data formats such as JSON, or XML that have a similar structure; the extraction of data or information is in accordance to public service agency needs in information systems respectively. Fig. 4 shows the web service sequence diagram.

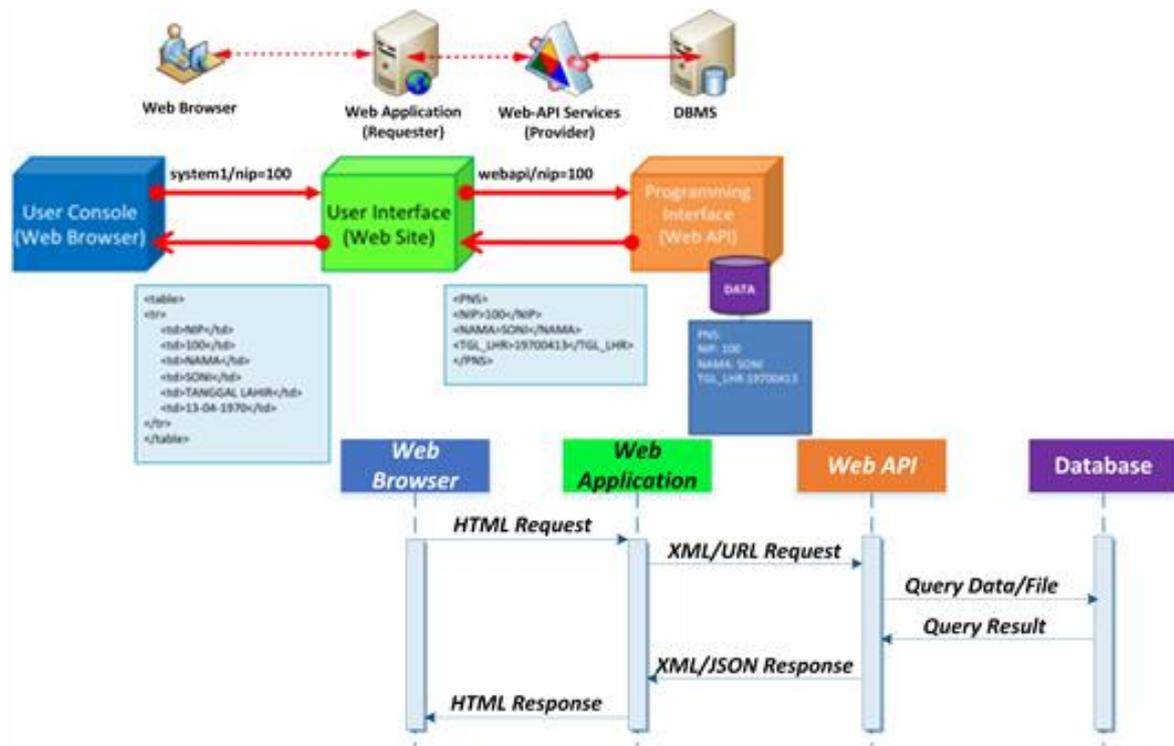


Fig. 4. Web service sequence diagram [15]

MANTRA, in general, has two interaction concepts that can be implemented, namely: (1) Point to point web-API. The concept of MANTRA point to point interaction between applications is middleware that connects MANTRA and database; the concept is on the layer of the web-API, that supplies data and the output from a database source, and communicates between API and the web application directly. There are authorizations between requester (who requires data) and provider (provider of web-API) as the form of security offered by these systems; MANTRA sets authorization requested based on registered user-agent. Fig. 5 describes the point to point Web-API. However, application access with point to point method will cause a loss of service availability if there is changing at domain name service, so that the application access on API web service should be mediated with a HUB/API service bridge using Government Service Bus (GSB).

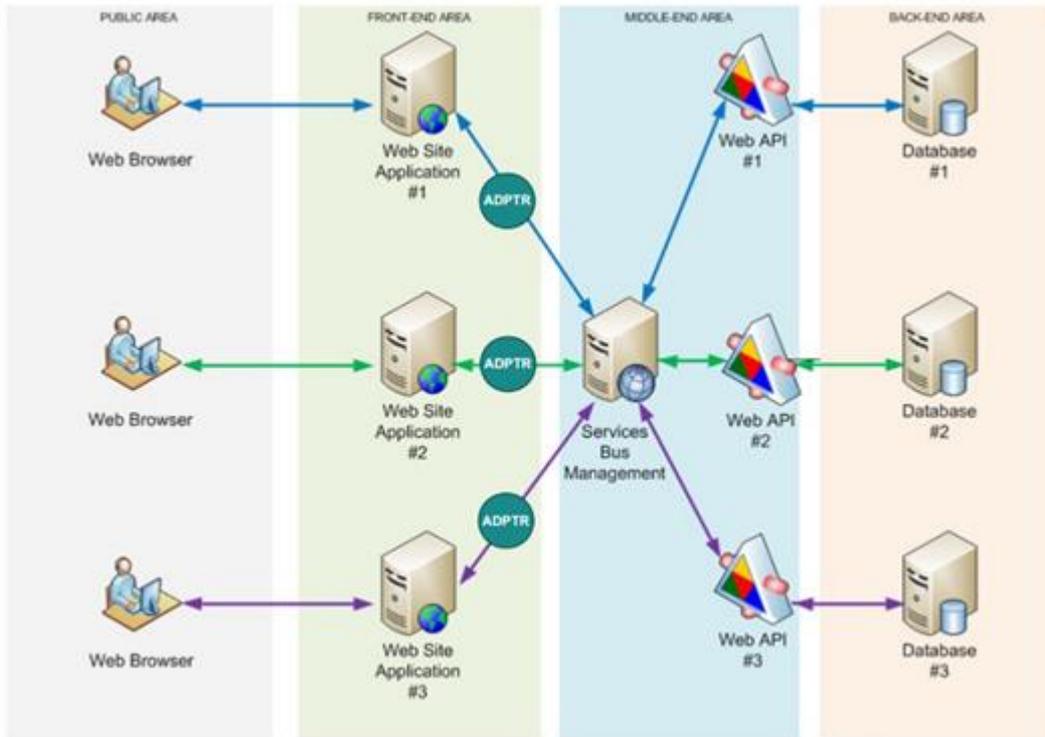


Fig. 5.MANTRA Point to Point Web-API [15]

(2) Government Service Bus (GSB) is the concept of interaction adopting services oriented architecture (SOA). The solution offered by GSB is to provide integrated description management services via Web-API Provider. Agent management is known as Universal Description Discovery and Integration (UDDI). Communication between UDDI can support access to information between agencies using channel based SOA or Service Bus Management. Because of MANTRA is intended to support interoperability between government agencies; the interaction concept is called Government Bus Services (GSB). UDDI is managing access to each web-API. UDDI agent will give access to applications or systems that want to utilize data via MANTRA authorization application. GSB enables open integration process which does not need to rely on application, operation system, and database. The messages are exchanged in XML form which is retrieved from a local database of a certain system. Fig. 6 describes GSB as a channel that stores various lists of web-API from many data sources and systems.

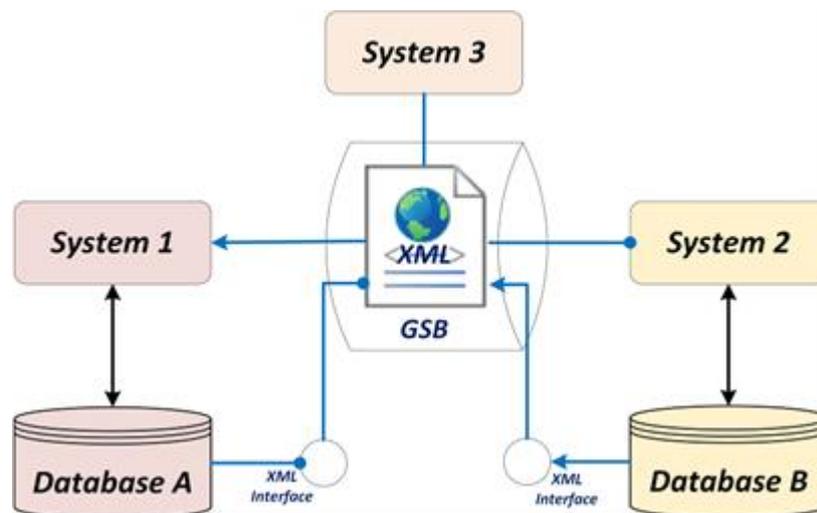


Fig. 6.Government Service Bus

3.2. Integrated Data System for Smart e-Government

In this research, we separated the system framework into two tier system and seven layers; the front-end system is the interface between user or application that interacts directly with users and back-end system is data access and communication layer between the system.

1) Smart e-government front end system

The development of integrated information system for Smart e-Government is more complex than developing the local system because we need to standardize the meta-data from front-end so that it can be processed in the back-end.

The Smart e-Government front-end application needs to be centralized for identity verification of all citizens to access the certain application. Numerous governments around the world are utilizing ID based smart cards [16]. In Indonesia, this is called e-KTP. In this study, we also adopted the application based smart card e-KTP as a single identity verification control system (Single-ID) with access privileges and stores value for use in various front-end application layers based on e-Government services. The classification according to the services provided is shown in Fig. 7 [17].

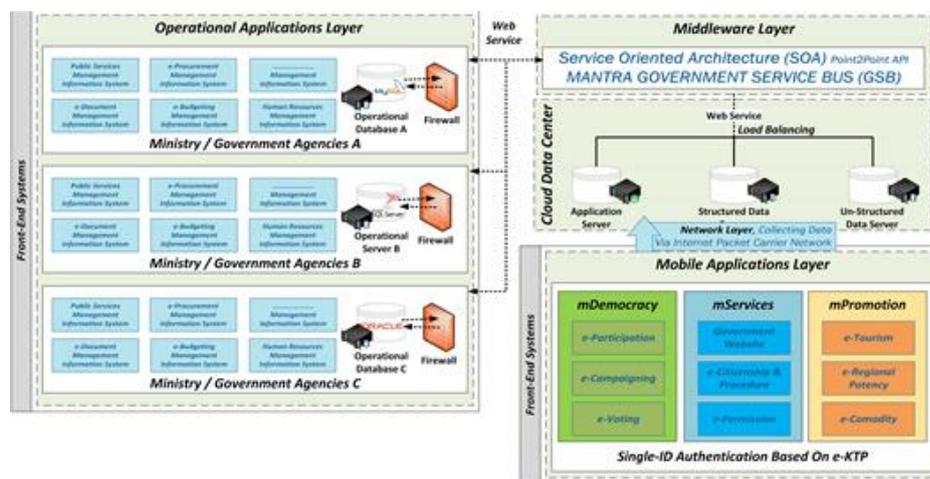


Fig. 7. Smart e-Government front-end system

1. e-Management: The use of ICT is to improve the efficiency of public administration and management as eProcurement, which improves the management of government purchases. Another application is the management of data for public administration through ICT.
2. Single-ID Based on e-KTP: Citizen electronic identity (e-ID) contactless smart card based on single identification number, In Indonesia also known as e-KTP, the purpose is to perform citizen Single-ID authentication with high security and accuracy, by verifying and registering user identity using NFC reader, where inside the e-ID card contained electronic identification/e-ID and digital signature/e-Signature).
3. mDemocracy: An e-Government mobile application portal called mAspirasi to promote democratic mechanisms by implementing e-Participation (for participation in service assessment and government decisions) e-Campaigning and e-Voting (electronic voting through ICT).
4. mServices: an e-Government mobile application portal called mPerizinan that makes information regarding its management and simplification of public services such as e-Citizenship and e-Permission.
5. mPromotion: An e-Government mobile application portal called mCity that makes information about e-Tourism (local culture) e-Regional Potency and e-Commodity (strategic place, economic commodity and local weather) available to the public to promote the city.

2) Smart e-Government Back-End Systems

The Smart e-Government back-end system is divided into six layers and illustrated in Fig. 8.

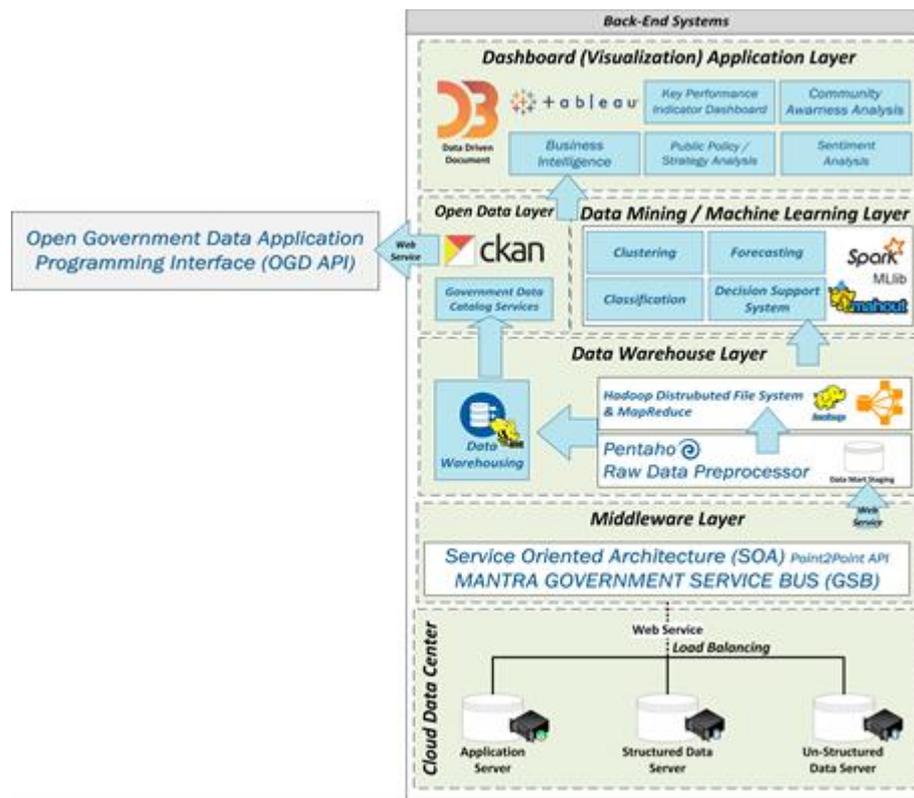


Fig. 8. Smart e-Government back-end system framework

1. Cloud Data Layer. It is a set of physical databases on the cloud servers for the non-transactional database collected from MANTRA API from different public agencies.
2. Middleware Layer. It is used for implementing interoperability between database and applications. We use MANTRA, an application to enable data exchange between different government agencies developed by Ministry of Information and Telecommunication (MCIT) of Indonesia. MANTRA works without looking at the application, database or operating system of various systems. It employs Government Service Bus (GSB) technology and Web-API (Application Programming Interface). The interoperability feature of MANTRA is described in Table 1.

No	Feature
1	The framework composed of web service component, which is loosely coupled, it allows to integrate with application that diverse implemented in various operating system, programming, platform, databases and network.
2	The middleware approach used in the framework is GSB. GSB responsible for integrating legacy systems, which provides interoperability with various old legacy applications, platform and database with the framework
3	Services in the process can be changed without affecting other services, as far as input/output types of the service are unchanged.
4	The services access can be from the Internet as an open network as well as from the private governmental network.

3. Data Warehouse Layer. This is used for reporting and data analysis. For raw data preprocessor and data integrator from large scale and multisource database, we used Pentaho Data Integration [18], which is integrating data in both structured and unstructured database format then stored in data warehousing system (i.e. Hive) that used to processing a large number of datasets. It has three key functions like summarization of data, query, and analysis In term of processing large and unstructured data using Hadoop Platform, Hadoop has two essential

components, namely: Hadoop Distribution System File (HDFS), it is used for storing and retrieving unstructured data,

MapReduce that responsible for processing jobs in distributed data processing and the aggregation of output [19].

4. Open Data Layer, is “the Comprehensive Knowledge Archive Network (CKAN) is a web-based open source management system for the storage and distribution of open data.” CKAN has developed into a powerful data catalog system that is mainly used by public institutions to share their data with the general public, such as G2C, G2B and G2O [20].
5. Intelligent System Layer. An intelligent data analysis using machine learning algorithm after data go to preprocess stage. The government agencies can use machine learning for constituents and government employees (ex: campaign prediction, supporting operational management). Risk & security management as well as anomaly & threat detection are carried out by identifying anomalies or signatures to address proliferation of fraud, money laundering. The common machine learning platforms for handling a huge amount of data analysis and large scale in distributed computing system are Apache Mahout and Spark. There are several machine learning models that can be used for data analysis, namely:
 - a. Forecasting. Campaign prediction in particular states based on the history / time-series data or Bayesian method.
 - b. Clustering. A part of data mining task, for measuring public service performance by clustering public service data.
 - c. Classification. A part of data mining task applied on the social network dataset, such as semantic analysis. Several proven methods can be used for the classifier, namely: C4.5, Naive Bayes, K-nearest neighborhood, decision tree, Support Vector Machine (SVM).
 - d. Sentiment Analysis. Also known as opinion mining information from traditional data or scraping social media data (web, twitter feeds and etc.). The purpose of sentiment analysis is to help policy makers to prioritize services and be aware of citizens interests and opinions [3].
 - e. Decision Support System. A supporting system to policy makers for public service process, performance evaluation and assessment.
6. Dashboard Application Layer. This is used for showing a graphical presentation of the current status or historical trends of a government key performance indicators (KPI).

3.3. Framework Scenario

Fig. 9 describes the framework scenario of an integrated data system for smart e-Government based on Big Data technology.

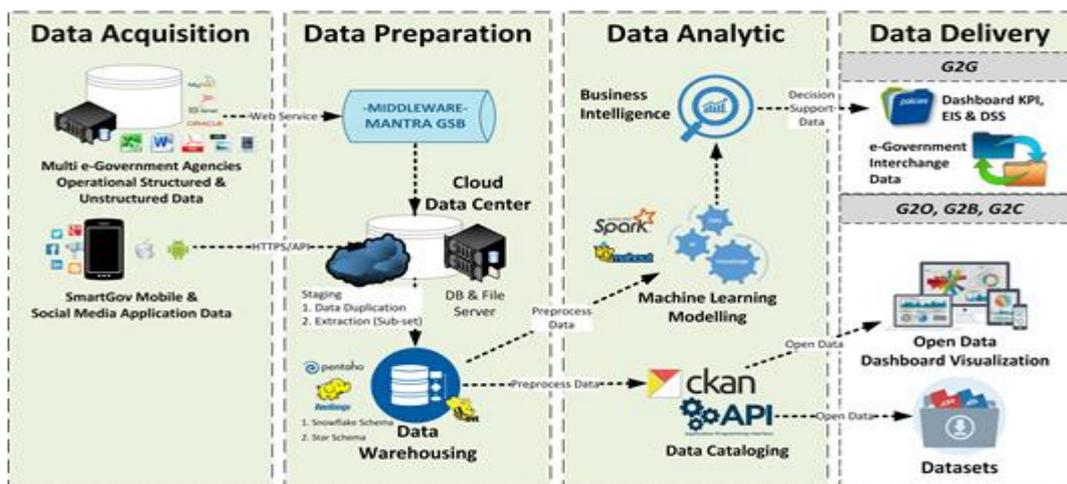


Fig. 9. Framework scenario

1. Data Acquisition. Integrate relevant data source between government agencies structured and unstructured data via MANTRA GSB and collected SmartGov mobile apps data in cloud data center.
2. Data Preparation. Preparing and merging data from different data source for reporting and data analysis preprocessing.
3. Data Analytic. Generating datasets using open data (i.e. CKAN) the output of datasets is files with rows and columns, these files can be used for data analytic and processing using machine learning platform (i.e. Spark, Mahout) for business intelligent in govt policy decision makers.
4. Data Delivery. Providing interchange data between govt agencies, open datasets and data representation for organization, business and citizen.

4. Conclusion

In this paper, we proposed a framework for developing Integrated Data System Smart e-Government which integrates traditional information system using MANTRA and Government Service Bus (GSB). The framework designed to achieve interoperability and flexibility that allows integrating diverse types of application, platform, operating system, database and network. The scalability that allows processing large and various data using Big Data technology, implementing Intelligent System or Machine Learning for decision support for policy-makers and providing Open Government Data to improve the quality of public services. Note that; we are on initial progress research and development; We are aware that the technology that we proposed is not a new product, but it is a new way of looking at how to integrate applications, database, coordinate resources, and manipulate information especially in a public service system or e-Government. For the future work, the framework will be separated into detailed implementation and validation progress, and we gratefully receive any suggestion for better research and development.

Acknowledgment

The authors are thankful to Ministry of Information and Telecommunication Republic of Indonesia and PT. Gamatechno Indonesia that have given permission and supported this research.

References

- [1] S. Basu, "E- government and developing countries: an overview," *Int. Rev. Law, Comput. Technol.*, vol. 18, no. 1, pp. 109–132, 2004.
- [2] J. C. Hai and I. Ibrahim, *Fundamental of development administration*. Selangor: Scholar Press, 2007.
- [3] J. Domingue, N. Lasierra, A. Fensel, T. van Kasteren, M. Strohbach, and A. Thalhammer, "Big data analysis," in *New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in europe*, J. M. Cavanillas, E. Curry, and W. Wahlster, Eds. Springer Open, 2016, pp. 63–86.
- [4] W. Chutimaskul and V. Chongsuphajaisiddhi, "A framework for developing local e-government," in *IFIP International Working Conference on Knowledge Management in Electronic Government*, 2004, pp. 335–340.
- [5] C. G. Mkude and M. A. Wimmer, "Strategic framework for designing e-government in developing countries," in *International Conference on Electronic Government*, 2013, pp. 148–162.
- [6] A. A. Sinsuw, X. Najoan, and Y. D. Rindengan, "Architecture design of Manado e-Government ICT networks: The road to Manado Smart City," in *Telecommunication Systems Services and Applications (TSSA), 2014 8th International Conference*, 2014, pp. 1–6.
- [7] H. J. Scholl, "Interoperability in e-Government: More than just smart middleware," in *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference*, 2005.
- [8] A. Geraci *et al.*, *IEEE standard computer dictionary: Compilation of IEEE standard computer glossaries*. IEEE Press, 1991.

-
- [9] I. W. Ordiyasa, L. E. Nugroho, P. I. Santosa, R. Ferdiana, and W. Kumorotomo, "Interoperability model of e-Government services," in *Information Technology, Computer, and Electrical Engineering (ICITACEE), 2015 2nd International Conference*, 2015, pp. 177–182.
- [10] M. M. El Benany and O. El Beqqali, "SOA based e-government interoperability," in *2015 IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, 2015, pp. 1–2.
- [11] A. Kö, C. Leitner, H. Leitold, and A. Prosser, *Technology-Enabled Innovation for Democracy, Government and Governance: Second Joint International Conference on Electronic Government and the Information Systems Perspective, and Electronic Democracy, EGOVIS/EDEM 2013, Prague, Czech Republic, August 26-2*. Springer, 2013.
- [12] F. Dzikrullah, N. A. Setiawan, and S. Sulisty, "Implementation of scalable K-Means++ clustering for passengers temporal pattern analysis in public transportation system (BRT Trans Jogja case study)," in *Engineering Seminar (InAES), International Annual*, 2016, pp. 78–83.
- [13] G. V. Pereira, M. A. Macadar, E. M. Luciano, and M. G. Testa, "Delivering public value through open government data initiatives in a Smart City context," *nformation Syst. Front.*, vol. 19, no. 2, pp. 213–229, 2017.
- [14] E. Tambouris, N. Loutas, V. Peristeras, and K. Tarabanis, "The role of interoperability in eGovernment applications: An investigation of obstacles and implementation decisions," in *Digital Information Management, 2008. ICDIM 2008. Third International Conference*, 2008, pp. 381–386.
- [15] D. Sukyadi, *Kerangka acuan dan pedoman interoperabilitas sistem informasi instansi pemerintahan*. DEPKOMINFO, 2008.
- [16] F. Fioravanti and E. Nardelli, "Identity management for e-government services," in *Integrated Series In Information Systems*, Boston, MA: Springer, 2008, pp. 331–352.
- [17] B. S. de Abajo *et al.*, "eGovernment: real democracy in digital society," in *International Workshop on Human-Computer Interaction, Tourism and Cultural Heritage*, 2011, pp. 113–123.
- [18] Hitachi Vantara, "Pentaho data integration," 2016. [Online]. Available: <https://www.hitachivantara.com/en-us/products/big-data-integration-analytics/pentaho-data-integration.html?source=pentaho-redirect>. [Accessed: 29-Jun-2016].
- [19] simplilearn, "8 Essential Concepts of Big Data and Hadoop," *Big Data and Analytics*, 2016. [Online]. Available: <https://www.simplilearn.com/big-data-and-hadoop-8-essential-concepts-guide-pdf>. [Accessed: 29-Jun-2016].
- [20] ckan, "About CKAN," 2016. [Online]. Available: <https://ckan.org/about/>. [Accessed: 20-Jun-2016].
-