www.arpnjournals.com

POPULATION METADATA DEVELOPMENT TO SUPPORT DATA INTEROPERABILITY BETWEEN GOVERNMENT AGENCIES IN INDONESIA

Lintang Yuniar Banowosari[1], I Wayan Simri Wicaksana[2], Lily Wulandari[3], Detty Purnamasari[3] and Pancat Setyantana[4]

[1]Information Management - Vocational Program in Information Technology Gunadarma University, Depok - West Java, Indonesia
[2]Information Technology - Doctoral Program Gunadarma University, Depok - West Java. Indonesia
[3]Information System - Computer Science And Information Tech. Faculty Gunadarma University, Depok West Java Indonesia,
[4]E-Goverment Department Ministry of Information and Commucation Technology Republic of Indonesia,Jakarta, Indonesia
E-Mail: lintang@staff.gunadarma.ac.id

ABSTRACT

E-Government is currently being actively implemented in Indonesian government agencies; only the benefits of the service in the public community still had not been optimal. A country with a population of hundreds of millions of people, it requires metadata in the demographic data management. If the population/demographic metadata are not developed, then there are some problems that will be encountered. Suppose diversity in terms of databases, is a constraint when government agencies will perform the information exchange. The public sector has the characteristics of the various items of data and information that has diverse meaning or understanding, as in the fields of law, regulation, public services, administrative processes and various other documents. A basic problem of the huge data is to use a service in a cross-sector and regional data exchange within a country. One obstacle is the diversity of information systems, both in the use of hardware, operating systems, database up to information understanding. This article presents a solution to overcome one of diversity, i.e. the database diversity by building data interoperability between government agencies in Indonesia by developing population metadata. Population metadata was developed by looking at demographic data elements that already exist, and choosing important elements of the population data. The government agency that already has the data or database applications do not need to make changes, but only need to add a middleware to serve as a converter to take advantage of the population metadata mapping. Population metadata user is all government agencies, especially for public services and who need for the exchange of data related to population.

Keywords: metadata, demogaphic, population, data interoperability, Indonesia.

INTRODUCTION

Currently the use of information technology in Indonesia is not only found in the business and education sectors, but also appears at the government by the existence of e-Government. Government agencies in Indonesia utilizing information technology with the purpose of improving service delivery to the public. However, most agencies are just developing information systems that provide only local benefits, particularly for unit owners of the information systems development budget. As a consequence, heterogeneity occurs in the many components, one of them is in the database system used. The issue began to arise when there is a need for data or information which is cross sector. If it would like performed data exchange, it must go through the long process of data copying and adjusting and it was time-consuming. Interoperability issue arises, which is how the different systems communicate with each other properly. One of the types of interoperability which is generally to be achieved is data interoperability, which is how an information system can communicate and exchange data with other information systems properly so that the transfer of data from the source to destination can be done without concern for diversity (the difference) of hardware platforms and software were used.

This article discus about the developments of demographic metadata to overcome one of the diversity that there exists in the implementation of e-Government in each government agency in Indonesia. This diversity is the diversity of database. The presence of the population/demographic metadata, it will make the interoperability occurred among government agencies in Indonesia.

Metadata is generally defined is data about data, or structured data about the data (Liu, 2004), (Sicillia, 2006). It could also be said metadata is data that describes a data or data that provide detailed information on the data, so that the population metadata is data that describes and gives information of each individual element on the biographical data and the information from each element in the family data and individual data.

Metadata was originally implemented for the creation of library catalogue. Several metadata schemes have been standardized are: (Taskforce on Metadata, 2004):

1. Data Documentation Initiative

Used for data sets in the social sciences. The idea of this scheme is to make the data more sharable, interoperable, and machine analyzable.

2. Dublin Core

Designed for all disciplines has 15 elements: title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage and rights.

3. Encoded Archival Description

Standards for encoding which helps in the search archives using XML

4. Machine Readable Cataloguing

Comprising three main components, those are: the leader, directory, and variable fields.

Metadata development undertaken by India was metadata for the land codification region. This metadata to identify the data elements associated with the generic land codification region, and standardize their formats to meet requirements of interoperability for vertical / horizontal exchange of data between various domains of e-Governance applications. (Expert Committee on Metadata Standards India, 2008)

RELATED WORKS

Research related to the application of e-government as a tool that can be used to improve the dual function, i.e. to make government information easier to find and manage, as well as interoperability have been widely carried out. Among these studies are as follows.

According to Pardo *et al*. [12], the combination of inter-organizational information has become key factor to the digital government. The proposed article describes two processes which share complex information such as: practical processes and governance processes. In a practical process system designers and developers must regularly exploit dilemma related to viability of various platforms, various database design and data structures, the quality of data that is constantly changing, and the network infrastructure that cannot be reconciled. In the government's view, the practical use of new work processes, restrictions on asset mobilization, and develops inter-organizational interactions.

Scholl *et al*. [13] have presented the interoperability of electronic government (or digital government) and nine constraints that affect the electronic government integration and interoperability. Nine of these constraints is as follows:

(a) Constitutional / legal constraints: Integration and interoperation may be illegal because it requires a democratic constitution that would authorize force (alienated) in levels and branches of government that are separate.

(b) Jurisdiction constraints: the government's constituent and non-government operate independently of each other and they have their own information and business processes, integration, interoperation and information sharing cannot be forced upon them.

(c) Collaboration constraint: Organizations those are diverse in terms of character and willingness for cooperation and interoperation with others. Past intimacy, socio-political organization, and leadership style affect the level of willingness and ability of potential interoperation.

(d) Organizational constraints: In these constraints the integration and interoperation may be difficult to establish organizational processes and resources that may differ between organizations.

(e) Information constraints: Transactional information may be more flexible than the strategic and organizational information. Standard of information quality appears when the information is collected, integrated and shared across multiple domains. Finally, information management encourages the use of information sharing, which in turn encourages managers to provide information.

(f) Managerial Constraints: interoperation becomes really complicated coupled with parties with the diverse interests and needs. Therefore, the need for management-related tasks may exceed the capability of interoperating partner management.

(g) Cost constraints: Integration and interoperation between diverse constituents may often limit in terms of availability of funds. However, information sharing initiatives have certainly helped to control costs.

(h) Technological constraints: Heterogeneity of electronic Government policy and network capabilities can limit the interoperation system standard becomes relatively low.

(i) Performance constraints: the higher the number of interoperating partners, generally the lower the performance of the system in the terms of response time. This article proposes a way to eliminate the limitations of the nine e-Government interoperation. This proposed work will help to achieve e-Government operating and services which efficient, in favor of citizens, accountability, and clear. Integration of the government process and information resources, in order to achieve interoperation of autonomous information systems, is very important to achieve that goal. In contrast, most of the integration and interoperation attempts to prevent challenges and limitations.

Chen *et al*. [6] have presented the stages of a mature e-Government model. Maturity model of e-government provides users information and services on a large scale in various dimensions of e-government. There are three types of mature e-government models as:

(a) The first model featuring several aspects, policies, technologies, data, and organizational concerns that must be specified for the organization for growth to the level of more mature e-government for the equality between government organizations and end users. This requires high technology and density of the organization to achieve the level of e-government is more established.

(b) The second model identifies four stages of e-Government integration of the following: (1) online catalogue, catalogue presentation and download forms, (2) transactions with the online forms and services, database work, and support online transactions, (3) vertical integration with the local system is connected to the higher-level system and the related functions, and (4) horizontal integration with integrated systems across

different functions and the real one-stop shopping for citizens.

(c) Third Model intensify the level of integration of the data needed for proper transformational e-government, but warned that the integration of such data raises significant concerns of confidentiality when the data involve personal information. This article is useful for students and practitioners to identify the phase of improvements to interoperability in e-Government and help to prepare for future improvements in the e-government research.

Peng Liu *et al*. [7] have proposed a trust model based on the interests and information sharing protocols for the successful issue of information sharing among government agencies. The proposed protocol is integrated in a policy group to share information, along with the information exchange and trust negotiation, and depend on each other. In addition, the protocol is implemented by utilizing new technology that XML Web Services.

Jing Fan *et al*. [10] have presented a conceptual model for the exchange of information in electronic government infrastructure (or digital). They know that the information sharing model of Government to Government (G2G) will help in understanding the contribution of G2G information sharing and will assist decision makers in formulating decisions with regard to the contribution in G2G information sharing.

Fillia Makedon *et al*. [9] have presented a negotiation-based sharing system called SCENS: Secure Content Exchange Negotiation System developed at Dartmouth College with the help of many interdisciplinary experts. SCENS is multilayer scalable system that guarantees the safety of transactions through many security mechanisms. It is based on a metadata description of heterogeneous information and applied to a number of diverse domains. They show that the government information which vulnerable and distributed, possibly bringing the deal on condition of sharing information by means negotiation.

Xin L. [8] has made a model of a distributed information sharing and also considers the technique standard support of the model. He concludes that the cost of managing information exchange and cooperation among government agencies will be reduced by improving the capability and efficiency of collaborative agencies and provide solutions for secure e-Government information sharing.

Rachmawati, U.A. [24] also propose metadata model in her research concerning electronic government in Indonesia. She built the initial model for e government topology and she did not elaborate deeply in service metadata model and ontology.

United Nation Development Programme [26] report a comparative study regarding Government Interoperability Framework which is also propose to build the metadata inside that framework, and the study showed that Indonesia government not yet have the GIF which contain metadata.

E-GOVERNMENT METADATA IN SOME COUNTRIES

Based on the articles that have been read, the following are review on the metadata of e-Government in some countries.

Framework for UK e-Government Metadata Standard (eGMF) published in May 2001, as a result of several months of consultation and planning that began in 1999. It is based on Dublin Core and has six additional elements that include a description and management for the purpose of e-Government. These six elements are: disposal; preservation; audience; location for the purpose of records management and archiving requirements and; accessibility and status for the purpose of resource discovery. [14]. in 2003, further work has been carried out and the e-Government Metadata Standard (eGMS) version 2.0 was published, containing four additional elements of records management: Addressee; digital signature; mandate; and aggregation.

Development of metadata standards The Australian Government Locator Service (AGLS) began in December 1997, with the AGLS metadata standard version 1.0, published in 1998. It is designed to improve the usability, accessibility and interoperability of government information and services through the provision of a standard web-based resource description [15]. AGLS metadata standard based on Dublin Core and has four additional elements designed for the Australian context; the function and availability for the purpose of government information and service find ability, and the audience and the mandate for records management

NZGLS metadata standards recommended in 1998 by the NZ Discovery Metadata Standard Level Working Group (NZMSWG), which was established to advise the general policies, standards and rules that must be used in all government agencies to improve the discovery of New Zealand information and government services. NZMSWG recommends that the Australian Government Locator Service (AGLS) are used in NZ e-Government with the some changes, such as the obligation element, refinement and encoding schemes As AGLS, NZGLS has four additional elements: function, availability, audience and mandate. NZGLS was issued in 2001. [15].

Irish Public Service Metadata Standard (IPSMS) was developed in 1999, as a result of recommendations made by Web Publication Group (WPG). The group recommended the use of Dublin Core with the two additional elements: service descriptor and descriptor life events [16]. A metadata working group (MWG) was established to determine metadata that will be in accordance with the e-government of Ireland, and in 2002, MWG agreed metadata standards proposed. It is based on Dublin Core, without the addition of new elements.

The Canadian government metadata framework established a strategy for the development of metadata within departments or federal agencies. The Government on-line Metadata Working Group was established to adopt a common metadata standard to be used on federal web.

www.arpnjournals.com

The group agreed on the Common Look and Feel Metadata Standard (CLF), which is based on the Dublin Core with the two additional elements: the audience for records management purposes, and keywords for resource discovery. This appeared in 2002. [17]

THE CONDITIONS OF DEMOGRAPHIC DATA IN INDONESIA

According to the Indonesia Central Bureau of Statistics (BPS), as of 2011 the number of Indonesia's population of more than 237 million people. So with the large populations, personal data must be managed properly. Significant large amounts of data should not be viewed as a problem; it should be as an asset if it is managed properly.

In addition, the development of information technology / information systems can be used for managing the large number of demographic data, so for the optimally utilization of information technology / information systems, it requires development of population/demographic metadata according to the needs and functions.

A country with a population of hundreds millions people needs metadata in their demographic data management. If the population/ demographic metadata are not developed, then there are some problems that will be encountered, namely:

1. There still exists that Indonesia people having more than one ID card (*KTP*), and if this continues then some problems difficult to handle such as the difficulty of Indonesia's population data collection, security of the country and their society.
2. The difficulty in determining what demographic data need to be collected and the demographic data that are important
3. The interests of some agencies / institutions both government and private on population/demographic data, which when metadata is not developed so there will be many variations / forms of population data which resulted in the duplication of the population data. Frequently also there is a difference perception or the population value quantity and quality from various agencies.

POPULATION METADATA DEVELOPMENT IN INDONESIA

Population Metadata in Indonesia was developed within government agencies for the purpose of interoperability. The underlying is the application and implementation of *e-KTP* (e-ID card) all over Indonesia and based of this model then the generic population metadata model were developed.

Additionally, demographic metadata is not only an important for the information integration and interoperability but it is also for new systems development related to a particular field. This allows the e-Government to be more flexible in dealing with changes according to the environment and needs.

Population metadata developed to establish interoperability among government agencies without requiring changes the database applications, but only require adding a middleware to function as a converter to take advantage of the population metadata mapping.

The phases that carried out in population metadata development in Indonesia are shown in Figure-1.
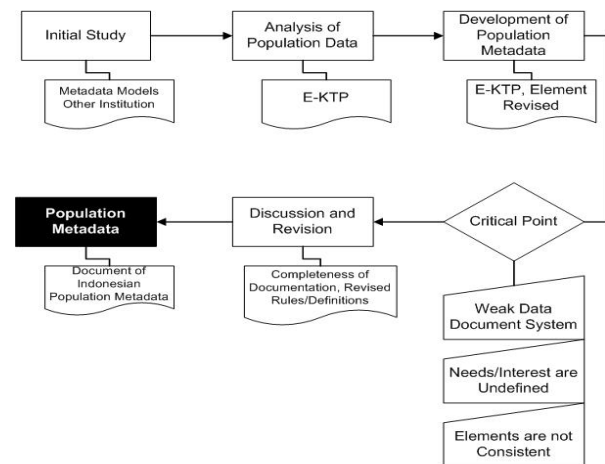

Figure-1. Phases of Development Population Metadata.

Development of population metadata that has been done is shown in Figure-1 begins with a preliminary study by studying metadata models, identifying metadata models which have similarity from the majority metadata which is owned by government agencies. Phases followed by an analysis of demographic data on the e-KTP and selected data elements by considering: a). it will be followed by and has always been a reference point for all government agencies in Indonesia, b). representation of metadata elements in e-KTP relatively similar with which available in metadata elements from other agencies.

After performing the analysis (shown in Figure-1), the next stage is to develop population metadata by selecting the demographic data elements that are important from existing e-KTP data, and then compare the demographic metadata with data owned by other agencies. In the course of the development of metadata population is not as easy as imagined. Some of the factors that become constraint are:

1. Not well defined its needs / interests/ requirements of the institution. This happens because of the difficulty to involve agencies / other agencies in conducting the analysis at the stage before the development of population metadata.
2. Data system associated with population do not have adequate documentation related to its data scheme, this is reflected in inconsistency between one data with other data viewed from the relationship between the content of metadata elements with the rule elements itself.

At the critical point of phase with the problems were found, conducted discussions with government agencies associated with metadata population, and make revisions to complete the documentation, rules and definitions of the metadata that has been developed.

In the *e-KTP (e-ID card)*, there are 27 population/residents data elements stored, some of whom are resident registration number, name, address, gender / sex. Analysis was performed on 27 data elements of the population and is also look at the various forms used by various government agencies such as form fields that are used during the establishment of birth certificates, death certificates, tax ID, bank accounts, credit cards, and school enrolment. Then the compression is performed by selecting elements that stated as an important element to be created its metadata. The selection of essential elements cannot be separated of the Law of the Republic of Indonesia No.23 of 2006 (*Undang-undang No.23 tahun 2006*) [25] regarding Population Administration which managed by the Ministry of the Internal Affairs, and one of the articles/chapters mentioned the population of data elements which are open. The results of the analysis that has been conducted, the population of the 27 data elements are determined 19 essential/important population data elements and to be created its metadata. Figure-2 is an example of metadata for population data element "Population Full Name ".

Figure-2. Is an example of metadata for population? data element "Population Full Name".

In this paper, what is meant by the population metadata is demographic metadata which has been developed in Indonesia, resulting in the presentation of population metadata sample (Figure-2); it is in the Indonesian language in accordance with the original conditions in Indonesia. In Table-1 describes the meaning of existing words in Figure-1 which translated into English.

Table-1. Translation of Figure-1: Indonesian to English.

Indonesian Language	English
Properti	Property
Nama_LGKP/Nama_Lengkap	Full Name
Label Elemen	Element Label
Label e-KTP	e-KTP (e-ID card) label
Definisi	Definition
Nama: kata untuk menyebut atau memanggil orang secara lengkap bukan panggilan	Name: words for call or call people completely not nicknames
Klasifikasi	Classification
Terbuka	Open
Aturan	Rule
Dipecah adanya Nama Depan dan Nama Keluarga dengan panjang minimal adalah 2 kata (6 karakter) dan maksimal adalah 5 kata (40 karakter)	Person Name is broken into the presence of First Name and Last Name with the minimum length is 2 words (6 characters) and the maximum is 5 words (40 characters)
Contoh	Example
Upper Class	Upper Class
Biodata	Biographical data
Sub Class	Sub Class
Property	Property
Format	Format
Character	Character
Kunci Utama	Primary Key
Nilai Bawaan	Default Value
Keharusan	Imperative (must fill or not)
Pengkodean	Coding
Tidak	No/Not
Table Referensi	Reference Table
Acuan	Reference
Definisi: KBBI	Definition: KBBI (Indonesian National Dictionary)

ILLUSTRATION OF THE UTILIZATION OF POPULATION METADATA IN INDONESIA

Other cross agencies application examples such as immigration, and the Ministry of Internal Affairs as seen from the scope of e-KTP which are intended for

citizens residing within the country, by the population metadata usage, the traffic of citizens can be updated.

The following are another illustration for the use of population metadata in private institutions such as financial institutions which managing the credit card. To be able to propose credit card at a bank, customers do not need to actually rewrite the complete biographical data, but only need to write down the number of ID, or social security number. This is because the metadata can be retrieved from the metadata owned by Bank Indonesia referred to or connected to population metadata of a trusted institution which is currently held by the Ministry of Internal Affairs as the primary source of population metadata associated with element data of population biographical data. Bank A as credit card issuers will easily see the status of loans, transaction history, or other matters related to the needs of the credit card issuer.
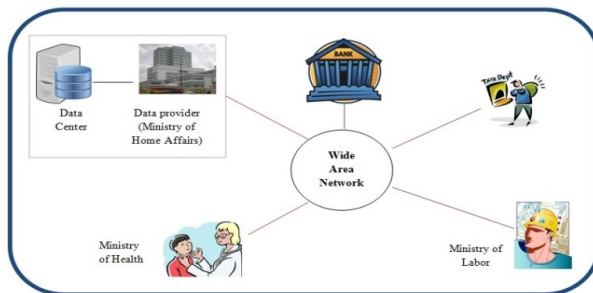

Figure-3. Data Interoperability between Institutions.

CONCLUSIONS

Population metadata is very important for Indonesia government. In this paper, the discussion is limited to the population metadata development in Indonesian government in order to support e-government. The process of preparation of population metadata is certainly expected to produce a high quality metadata. Metadata quality is a requirement for the success of the electronic government services (governance of e-services). In further research is expected to continue in the development of other metadata associated with population metadata, for example health metadata, employment metadata etc. If this can be done sustainably, it expected to be able to form a good quality Indonesia's population metadata which is measured by the completeness, integrity, accuracy, consistency, ease of access and not easily changed (timeless).

REFERENCES

[1] Task Force on Metadata. 1999. http://www.libraries.psu.edu/tas/jca/ccda/tf-meta3.html.

[2] Liu J. 2004. Metadata Development in China. D-Lib Magazine. 10(12). ISSN 1082-9873. China.

[3] Sicilia M.A. 2006. Metadata, Semantics, and Ontology: Providing Meaning to Information Resources. Int. Journal Metadata, Semantics and Ontologies. 1(1). Spain.

[4] Land Region Codification Version 0.9. 2008. Expert Committee on Metadata and Data Standards. India.

[5] Chen Z., Gangopadhyay A., Holden S. H., Karabatis G. and McGuire P. 2007. Semantic integration of government data for water quality management. Government Information Quarterly. 24, 716-735.

[6] Peng Liu and Amit Cheta. L. 2005. Trust-Based Secure Information Sharing Between Federal Government Agencies. Journal of the American Society for Information Science and Technology. 56(3): 283-298.

[7] Xin Lu. 2007. Distributed Secure Information Sharing Model for E-Government in China. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing. 3: 958-962.

[8] Fillia Makedon, Calliope Sudborough, Beth Baiter, Grammati Pantziou and Marialena Conalis-Kontos. 2003. A Safe Information Sharing Framework for E-Government Communication. IT white paper from Boston University.

[9] Jing Fan, Pengzhu Zhang. 2007. A Conceptual Model for G2G Information Sharing in Egovernment Environment. 6th Wuhan International Conference on E-Business, Wuhan (CN).

[10] Achille Fokoue, Mudhakar Srivatsa, Pankaj Rohatgi, Peter Wrobel, John Yesberg. 2009. A decision support system for secure information sharing. Proceedings of the 14th ACM symposium on Access control models and technologies. pp. 105-114.

[11] Pardo T. A. and Tayi G. K. 2007. Interorganizational information integration: A key enabler for digital government. Government Information Quarterly. 24, 691-715.

[12] Scholl H. J. and Klischewski R. 2007. E-Government Integration and Interoperability: Framing the Research Agenda. International Journal of Public Administration, 30(8): 889-920. International Journal of Computer Science and Information Technology. 2(3), June 2010 91.

[13] Coles C. 2003. Metadata and Search. [online].

[14] Wilson A. 2002. Australian Government Locator Service. [online], http://www.egov.vic.gov.au/Documents/agls.ppt.

www.arpnjournals.com

[15] Booth K. 2002. Mediating for Metadata Standard: Competing Demands of e-Government, Archivists and Librarians for Resource Description in New Zealand. [online], ttp://www.vala.org.au/vala2002/2002pdf/05Booth.pdf.

[16] Booth K. 2002. Mediating for Metadata Standard: Competing Demands of e-Government, Archivists and Librarians for Resource Description in New Zealand. [online], ttp://www.vala.org.au/vala2002/2002pdf/05Booth.pdf.

[17] WPG. 1999. Recommended Guidelines for Public Sector Organizations. [online], http://www.gov.ie/webstandards/metastandards/background.html.

[18] GOMWG. 2005. Government of Canada Metadata Implementation Guide for Web Resources. [online], http://www.tbssct.gc.ca/im-gi/mwg-gtm/ts-sf/docs/2005/migwr-gpmrw/migwr-gpmrwtb_e.asp.

[19] Land Region Codification Version 0.9. (2008). Expert Committee on Metadata and Data Standards. India.

[20] Liu J. 2004. Metadata Development in China. D-Lib Magazine. 10(12). ISSN 1082-9873. China.

[21] Sicilia M.A. 2006. Metadata, Semantics, and Ontology: Providing Meaning to Information Resources. Int. Journal Metadata, Semantics and Ontologies. 1(1). Spain.

[22] Task Force on Metadata. 1999. http://www.libraries.psu.edu/tas/jca/ccda/tf-meta3.html.

[23] Ummi A. Rachmawati, Dana I. Sensuse, Heru Suhartanto. 2012. Initial Model of Indonesian e - Government Grid Services Topology. International Journal of Computer Theory and Engineering. 4(4).

[24] Undang-Undang Republik Indonesia No. 23 Tahun. 2006. tentang Administrasi Kependudukan (The Law of the Republic of Indonesia No. 23 of year 2006 regarding Population Administration).

[25] United Nation Development Programme. 2007. e-Government Interoperability: A Review of Government Interoperability Frameworks in Selected Countries. UNDP, ISBN: 978-974-13-1624-3, Bangkok, 2007.