MASTER DATA MANAGEMENT IMPLEMENTATION IN DISTRIBUTED INFORMATION SYSTEM CASE STUDY DIRECTORATE GENERAL OF TAX, MINISTRY OF FINANCE OF REPUBLIC OF INDONESIA

Aris Budi Santoso[1], Yoga Pamungkas[2], Yova Ruldeviyani[3]

Faculty of Computer Science
University of Indonesia
Jakarta, Indonesia

E-mail: aris.budi@ui.ac.id[1], yoga.pamungkas81@ui.ac.id[2], yova@cs.ui.ac.id[3]

Abstrak

Direktorat Jenderal Pajak (DJP) memiliki arsitektur sistem yang tersentralisasi dengan data yang terdistribusi. Permasalahan utama terdapat pada replikasi data master dan referensi pada database di beberapa aplikasi, mencakup juga struktur yang berbeda-beda dan proses sinkronisasi yang tersebar. Untuk itu perlu dilakukan master data management (MDM) dalam mendukung sistem informasi terdistribusi. Pengukuran tingkat kematangan pengelolaan master data di DJP telah dilakukan menggunakan MDM maturity model (MD3M) Spruit dan Pietzka dengan hasil topik Data Protection, Data Quality dan Maintenance berada pada level 3 atau tahap defined process, sedangkan topik Data Model, Usage and Ownership berada pada level 2 atau tahap repeatable. Untuk meningkatkan master data maturity, MDM, penanganan permasalahan pada MDM mengacu pada Data Management Body of Knowledge (DMBOK). Permasalahan pada MDM ditemukan berkaitan dengan kebijakan dan arsitektur yang belum mendukung. Kebijakan dan pendekatan arsitektur sistem MDM terpusat diperlukan untuk mengatasi permasalahan tersebut. Salah satu solusi permasalahan arsitektur adalah dengan menerapkan Data Virtualization, sehingga memungkinkan DJP menggunakan arsitektur sistem MDM yang terpusat tanpa harus mengkonsolidasikan seluruh data master dan referensi dalam satu database khusus.

Abstract

Information system architecture of Directorate General of Tax (DGT) is centralized with distributed data. The main problems are replication of master and reference data which spread among applications which vary on data structure. Beside that, the synchronization jobs that spread in many locations and not well managed. Therefore, Master Data Management (MDM) needs to be implemented with accordance to characteristic of centralized distributed information system. Master data management maturity evaluation is conducted using MDM maturity model (MD3M) Spruit dan Pietzka, the result is Data Protection, Data Quality and Maintenance topic have maturity level 3 or defined process stage, while Data Model, Usage and Ownership topic have maturity level 2 or repeatable stage. Solutions to solve MDM issues and to enhance the master data management maturity level are proposed based on Data Management Body of Knowledge (DMBOK). DGT's MDM issues are related to insufficiency of policy and architecture of MDM system. Policy and architectural approach of centralized MDM system is required to solve that issues. Proposed solution involves the use of data virtualization to enable implementation of centralized system of MDM without consolidate all master and reference data into new database.

Keywords: *Master Data Management, Distributed Information System, DMBOK, Data Virtualization*

1. Introduction

Information system with distributed type of architecture has consequence of the spread of master and reference data in many databases. Each application requires entity of master and reference data in same structure or with some modification. Hence, integration and interoperability are needed to make the master and reference could be shared and reusable.

Information system of Directorate General of Tax (DGT) is centralized with distributed data.

Service Oriented Architecture (SOA) is applied by implements SOA middleware which gives capability to develop atomic services, and to compose complex services through modelling tools. All services which represent the main business function of taxation are published through Enterprise Service Bus (ESB).

MDM system development project has been initiated in 2013, centralized database was created for manage master and reference data, but, limited only for entities that related to taxpayer, it is called DB Master. Updating process and synchronization to core tax system (SIDJP) has been defined. However, in 2014 the development of SIDJP is changed from single big system to modular and distributed system. Each of new module of SIDJP are developed separately, so that each application has their own database. Common master and reference data from DB Master are accessed through web services.

Replication of master and reference data still has some problem on data synchronization between main database and specific application database. The synchronization process not in centralized system, so it is hard to monitor those background processes. That condition leads to master data inconsistence problem in several applications.

Poor documentation of application specification is one of factor that causes difficulty in identifying the occurrence of master and reference data replication. Those condition will make synchronization process more complex when master and reference data is changed. Hence, the master and reference data need to be managed in centralized manner.

DGT has concern about those MDM issues. MDM becomes one of application platform of SIDJP in DGT's IT Blueprint for 2015 – 2019 [1]. MDM implementation is not only about technology, but requires comprehensive approach which cover process, governance policy, and implementation of related technology [2]. Based on it, the research question is "how to approach master data management implementation at Indonesian DGT?"

This study aimed to propose an approach to enhance the MDM in DGT which cover governance policy and appropriate architecture of MDM system that support master, reference and transaction data integration and interoperability in distributed system environment. Therefore, it could improve the DGT's MDM capacity. MDM maturity model (MD3M) from Spruit and Pietzka is used to evaluate DGT's MDM maturity level [3]. While activities to enhance maturity level, MDM, and solution are proposed based on Data Management Body of Knowledge (DMBOK) [4].

2.   Literature Study

A.   Master Data Management Definition

Master Data Management (MDM) is set of processes to maintenance, to integrate, and to align master data, therefore the information produced from information system become consistent. MDM is control over value of master data to enable consistent and common use across multiple systems. While Reference Data Management is activity to control dictionary of terminology which include standard definition, business process object codification, relationship of references across domain, and consistently shared use of references from multiple sources [4].

Principles of reference and master data management activities based on DMBOK are below[4]:

1) Reference and master data are shared data which used across business functions, therefore, provision of those data need to be well managed;
2) Reference and master data are owned by the organization, and don't belong to certain application or unit. Because it has broader usage, including all functions in organization, so it should be managed appropriately and based on standard;
3) Business data steward is responsible for controlling and monitoring quality of reference data;
4) Master data must be real time and updated. Every change in reference and master data must be authorized, logged, and communicated before it is applied;
5) Replication of master data allowed only from defined source. Some reference system is needed to enable shared use of data.

Recently, master data management has become part of enterprise information management, that is driven by need of better quality of data [5], especially in distributed information system environment. Master data is category of data which related to shared business entities that used across business processes. Master data is representing primary entities which have important role to fulfil strategic business requirement [5].

B.   Master Data Maturity Assessment

Master data maturity assessment required by organization to measure the level of master data integration across multiple units or areas. Maturity model is an artefact which aimed to defining current, target and improvement of capability status of organization.   MDM maturity assessment requires some model which could be used to measures all of the MDM processes, including view of data and operational aspect [3]. Maturity assessment could help organization to improve

capabilities of data management system component and adapting changes.

DAMA [4] also emphasizes about importance of maturity assessment of each data management section. The need to accommodate changes in regulation, standard, organization, business and technology are the reasons why maturity assessment is important in data management.

Literatures have been reviewed to choose the MDM maturity model which will be used to assess DGT's MDM maturity level in this study. Spruit and Pietzka propose model for assess MDM maturity level, it is called MDM maturity model (MD3M). Those model uses 5 maturity levels which adopted from COBIT [3], the levels are below:

1) Initial, means there are awareness about MDM in operational level;
2) Repeatable, means MDM activity has been conducted to solve individual problems, still in operational level;
3) Defined-process, means there are collaboration in tactical level;
4) Managed and measurable, means best practices has been used to implement MDM. Processes is clearly defined in tactical level;
5) Optimized, means MDM processes has been optimized, affecting to organization efficiency improvement.

Assessment including 5 main topics which consist of 13 MDM areas. The topics which play as indicator of MDM maturity are: 1) Data model, 2) Data quality, 3) Data ownership and usage, 4) Data protection, 5) Data maintenance.

Zuniga et al. on 2018 conducted literature review to define MDM maturity model for micro finance in Peru. Those study comparing the existing models, they are MD3M which proposed by Spruit dan Pietzka, Dataflux's model, Standford University's model, Oracle's model and others model from previous research. All topics of MD3M is adopted on the model resulted from those study [6].

MD3M that proposed by Spruit dan Pietzka is fundamental, comprehensive, and general, so it could be used in many business sectors. The journal of MD3M has high H-index and categorized as Q1-best quartile. Hence, this study will use MD3M to conduct MDM maturity assessment and to get insight about MDM processes in DGT.

C. Master Data Management Activity

DAMA [4] provide guidelines for define steps in implementing MDM. When implementing MDM, organization will arrange it in step by step and iteration processes. Each organization will have different beginning maturity level of MDM, therefore, the approach will vary depending on theirs beginning condition. According to DMBOK, MDM activities for each category are below:

1. P – Planning
   - Identify driver and requirement;
   - Evaluation and assessment of data sources.
2. D – Development
   - Define architectural approach;
   - Define data model;
   - Implement data sharing/ integration services mechanism.
3. C – Control
   - Define stewardship and maintenance process;
   - Establish governance policy.
4. O – Operational
   - Implement data sharing/ integration services mechanism.

MDM system architecture is important part of solution. Architecture is basic composition of system which represent its components, relationship of components, and principles to create system design [4]. Defining an architectural approach is a part of MDM development activity. MDM system architecture will be defined based on result of data sources evaluation and identification of driver and requirement. DAMA provides detailed explanation about factors that need to be considered in defining architecture type of MDM system, that is: (1) Business strategy, (2) Platform of existing data sources; (3) Characteristic of data, especially about its lineage and volatility, (4) Implication of latency level.

D. Recent Technology Concept which Could Improve MDM

DAMA published DMBOK V.2 as a refinement to respond to technological changes. Those version of DMBOK added one section about data integration and interoperability. That section explains data movement and consolidation both in same location or between different locations. Data integration and interoperability is important things for data warehousing, and also important for reference and master data management, because in providing data to user or systems, they have requirement to transforming data and integrate data from multiple sources [4].

Data virtualization is a new concept that is introduced in section data integration and interoperability on DMBOK V.2. It is a mechanism that could be implemented in MDM and data quality management projects and will help to gain benefit from that projects. It enables to provide single layer from multiple data sources in different platform. Since the data sources seems in single layer of virtual database, so data transformation, data ownership administration, and data access

control could be centralized managed without move data to new location.

Implementing concept of data-service driven organization, the organization need to be transformed, in order to build capability in providing data access over multiple heterogeneous data sources in timely and conveniently way. Data distribution or data access provision is part of data integration architecture. It refers to process of organization in provide business view of data in virtual layer [7]. Therefore, data will become services which serves organization in achieving their mission.

All of literatures mentioned above will be used to build MDM solution in this study. The theoretical framework is depicted in Figure 1.
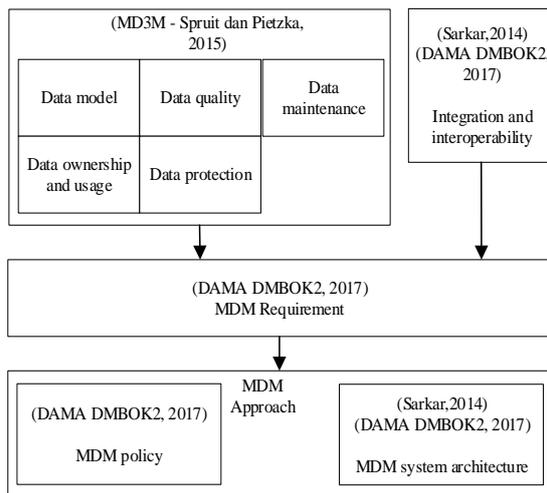
Figure. 1.   Theoretical Framework

3.   Research Methodolgy

This study uses primary data which collected through interview and observation. The data is used to assess the beginning MDM maturity level, and to get insight about existing MDM system that has been implemented by DGT. Literature review also conducted to find appropriate model for MDM maturity evaluation, and to propose solution of MDM issues.

List of question or check list of condition is arranged for each topics and areas according to MD3M, then interview is performed based on it. Observation is also conducted to collect information about data sources and the existing integration architecture that has been implemented in DGT. Requirement of MDM system solution is defined according to information about business process, existing data sources, existing data integration architecture, and master data issues which identified.

MDM policy and refinement of MDM system architecture will be proposed based on MDM system requirement, data model, and integration pattern. Existing MDM system architecture and MDM issues which identified are considered in proposing appropriate MDM system architecture.

Proposed solution for improvement MDM will be validated by checking availability of relevant technology such as middleware which could fulfil the requirement. The working prototype will be built based on it. Steps in this research are refer to literatures such as journal and DMBOK. Summary of research methodology is depicted as Figure 2.
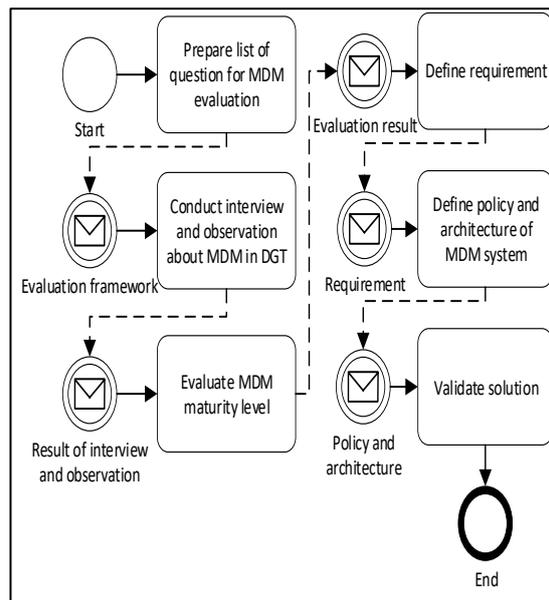

Figure. 2.   Research Methodology

4.   Results

A.   Master Data Management Maturity Assessment

The MDM maturity assessment is conducted through interviews with 5 key persons who involved in the project of MDM on 2013, the respondent profile is shown in Table 1.

Table. 1. RESPONDENTS OF MDM MATURITY ASSESSMENT

Respondent	Role	Experience
A	Data Warehouse Developer	8 years
B	Data Warehouse Developer	8 years
C	Database Designer	3 years
D	Database Designer	5 years
E	Manager at Application Development Unit	>10 years

We conduct the interview with all respondents guided by question list that arranged according to MD3M. The answer from all respondents then reconcile, if contradiction occurred then we clarify by observe the document and the information system. Based on the results of those process, the maturity level of MDM is calculated according to MD3M. It is found that data protection, data quality, and data maintenance are at level 3 or defined process stage, while data model, and data ownership and usage are at level 2 or repeatable stage. Details of assessment result for every areas of each MDM topic are shown in table 2.

Table.2. RESULT OF MDM ASSESMENT IN DGT

Area	Level
Data Model	**2**
Data landscape	2
Definition of master data	3
Master data model	3
Data Protection	**3**
Data protection	3
Data Quality	**3**
Assessment of data quality	2
Impact on business	3
Improvement	3
Reason/sources for poor quality	4
Maintenance	**3**
Data lifecycle	2
Storage	4
Usage and Ownership	**2**
Data access	3
Data ownership	1
Data usage	4

B. Identification of Sources and Contributors of Master Data

Identification of sources and contributors of master data in DGT is conducted through interview and observation. Information about databases which stored master data and business processes that contribute to master data are collected in this step, the results are below:

1. Taxpayer registration
Master data and reference which related to this business process such as taxpayer, region code, business field classification code, taxpayer type and others have been managed in DB Master. Any data changes that occur in e-registration will also updated to this database directly.

2. Staffing and organization
Master data related to staff, office, organization unit and theirs reference are managed in DB SIKKA which is part of Information System of Finance, Staffing, and Asset (SIKKA).

3. Finance and assets management
Master data related to this business process such as chart of account and government asset managed in DB SIKKA.

4. Object of property tax administration
Master data related to this business process are managed in SIDJP-NINE property tax module. This data has not been consolidated into DB Master.

C. Identification of MDM Problem

According to the result of MDM maturity assessment, it has level between 2 and 3. Main problems of MDM in DGT which identified during the assessment are as follows:

1. Replication of reference and master data across applications which not well managed cause maintenance difficulties;
2. Master data is spread over 3 different locations, that are DB Master, DB SIKKA, and Integrated Data Warehouse. Because there is still poor solution to integrate the master data, so the application developer still replicates master data when needed;
3. Data quality examination still performed in each data sources, so quality monitoring of master data become difficult;
4. Data access management still performed in each application, that makes difficult to manage privilege of master data consistently;
5. Metadata of master data has not consolidated, it is spread over data sources, that makes difficult to understand all the master data;
6. Policy and procedure in application development have not consider MDM, so there are chance that master data entity is fully replicated to another database, even though it could be created with inheritance, and the main entity is reused;
7. Policy and procedure of access right to master data have not aligned with job requirement of function, unit or position in organization;
8. Stewardship of master data has been defined, unit which responsible for manages database especially DB Master is pointed to be a data steward. However, because there are no regulations regarding master data, so several master data are still managed by certain business unit.

D. Define Master Data Management Requirement

Based on problems identified in previous section, MDM requirement will be defined. The requirements to improve MDM in DGT are as follows:

1. Requirements related to governance policy
   a. Policy and procedure to support integration of master data;
   b. Procedure that govern MDM must be included in application development policy;
   c. Data access right must be included in jobs description, therefore access control over

master data could be defined clearly and implemented;

d. Unit that responsible for stewardship of master data must be determined and included in jobs description.

2. Requirements related to MDM system architecture
   a. Integration of data sources of master data which spread in several databases;
   b. Data access provision for target applications which use master data, therefore those application could access master, reference and transaction data simultaneously, just like in the single database;
   c. Data quality examination could be performed centrally;
   d. Metadata of master data could be managed and accessed centrally.

E. Data Model Identification

Sources and entities relationship of master data are identified through interview and observation. Locations of master data spread in 3 different databases, however master data entities have clear relationship. There are master data that still replicated in different structure, such as entity of taxpayer, tax office, and property tax object. Conceptual model of master data in DGT is as depicted in Figure 3.

F. The Main Points of MDM Governance Policy

Governance policy is one of key requirements to improve MDM in DGT. Main points of governance policy proposed for DGT are as follows:

1. Determine business data stewardship, in this case unit which responsible for analysis and design information system is the most appropriate unit pointed as data steward for master data. The reason for that determination is because those unit has authority in application and database design, other than that, it has knowledge and capability to understand whole business processes in DGT;
2. Procedure of System Requirement Specification (SRS) definition must be including identification of master data;
3. Logical layer or data virtual layer should be implemented in providing access to databases, so users or applications don't need to know complexity of physical database, they only need to connect to logical layer according to their needs;
4. Operational of MDM should centralized, database development activities related to master data requirement are performed by unit which has authority in database development in Directorate of Transformation of Information and Communication Technology. Centralized MDM could improve quality and availability of master data;
5. Access rights of master data granted to users based on job description and authority of theirs position and unit. While access rights for applications granted based on business scope of each application.
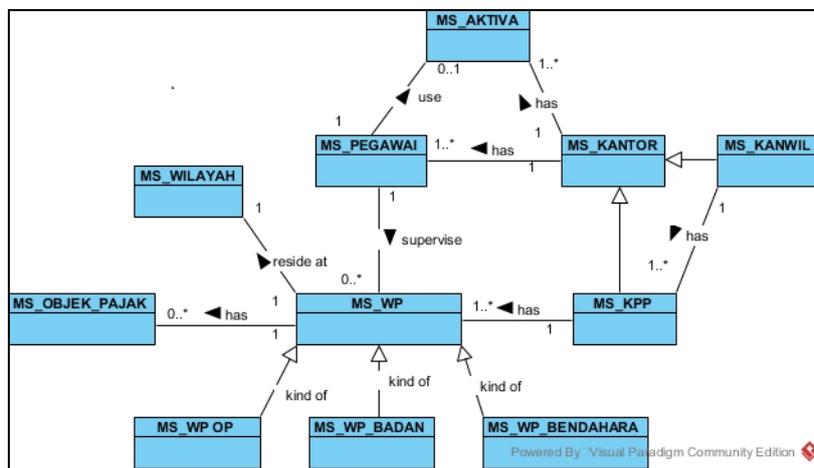

Figure. 3. Conceptual model of master data of DGT

G. Define Architectural Approach for MDM System

DGT's MDM system is toward centralized system and apply Service Oriented Architecture (SOA) to distributing master data to target applications. However, master data related to staffing, finance and assets still managed in SIKKA which is the leading system of that business

processes. MDM with leading system pattern is frequently encountered in practice [5].

According to theoretical framework as shown in Figure 1, in proposing architectural approach for MDM system, this study will consider requirements of MDM, MDM policy, and technology advancement which could improve the effectiveness and efficiency of MDM. Other than that, the advantage and disadvantage of every architecture pattern of MDM is also taken as consideration, refers to the study about decision model for MDM architecture which has been conducted in previous.

The entire architectural pattern has advantage and disadvantage. Centralized, leading system and SOA pattern are good in data quality, but poor in flexibility and interoperability. While connection hub and registry pattern are good in flexibility and business optimization, but poor in data quality [5].
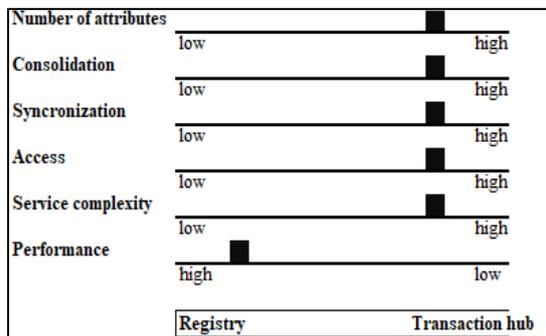
Figure. 4.   MDM Architecture Selection

MDM architecture should be chosen based on requirement as selection criteria. At least there are six criteria that must be considered when choosing MDM architecture style[8]. Based on it, we analyse the requirement of MDM in DGT as depicted in Figure 4.

Almost all the criteria except performance lead the decision to transactional hub. However, transactional hub has limitation in performance because single consolidated master data repository often raises performance issue[8]. Other limitation of transactional hub is about poor data quality[5]. The complexity of data consolidation and data distribution which involving data movement between locations increase the risk of incomplete and inconsistence data on target application.

According to advantage and disadvantage of transactional hub style and consider performance

and data quality requirement as mentioned before, we prefer to use federative approach as MDM architecture style for DGT. Federative approach more focus on data interoperability[14], the use of shared attribute will reduce the need of moving data into another location. Data virtualization also need to be used to enable centralized maintenance and implementation of data governance.

Based on requirement related to MDM system architecture described in previous section, it can be understood that centralized MDM system needs to be implemented in DGT to ensure the data quality, and to govern data ownership, data access, and metadata of master data that's used across unit and function in organization. Besides that, there are also requirement to integrate master data with transaction data with simple and timely mechanism. The use of Extract Transform Load (ETL) for moving data to a new data source has high complexity and can causes delay. ETL job failures becomes a risk in the timeliness of data provision to target applications.

Data virtualization could make multiple data sources with different platform seem to be in one layer, therefore data transformation, data ownership management, and control over data access could be performed centrally through those data virtual layer, without having to move data to new location. Hence, this study will involve the use of data virtualization concept as part of architectural approach for MDM. The proposed MDM system architecture for DGT is as depicted in Figure 5.

Each virtual database has data schemas which arranged based on requirement of each application. Every tables or views in virtual database represent the logical model of tables in multiple and heterogeneous data sources. Connection to virtual database could be created in several mechanisms, there are query execution through JDBC or ODBC protocol, and web services such as SOAP, REST, or OData. Security policy could be implemented in virtual database in form of authentication and authorization which integrated with enterprise authentication provider or Active Directory (AD) that is used in DGT. Caching could be used both in memory or materialized to improve performance while accessing reference data which rarely changed.
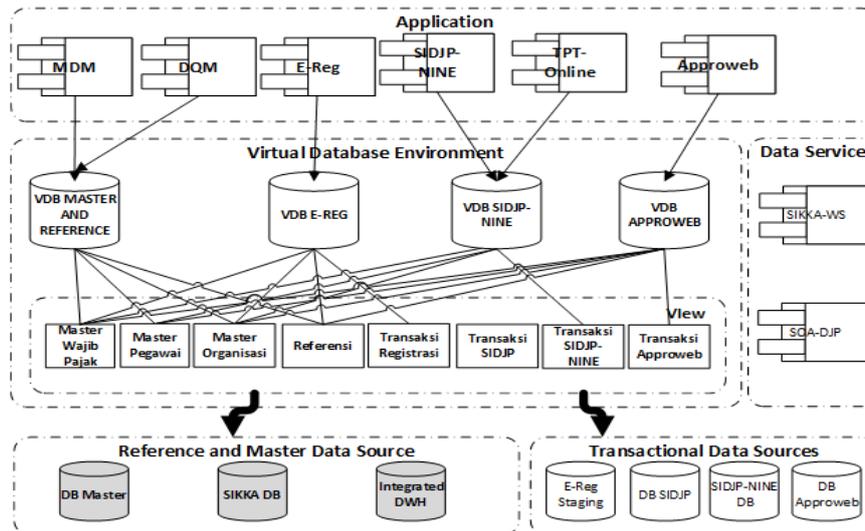
Figure. 5.   MDM Architectural Framework with Data Virtualization

5.   Discussion

Nowadays, data virtualization is supported by middleware from many credible principals. It has capability for data integration of multiple data sources in different format and enable system interoperability. It provides universal access to virtual database through JDBC and ODBC connectivity, and also provides web service channel with standard mechanism such as SOAP, REST and OData and standard data format like XML and JSON. Almost all middleware has centralized system administration such as facility for deployment management, metadata, versioning, data access monitoring and activity log.

Data integration and interoperability are important in MDM [4]. Data virtualization could realize benefit of MDM project, because with those mechanism, master data in single source of truth could be combined with operational data in real time to get common understanding of business process over all units in organization[9]. Data virtualization also could facilitate data quality management by provide monitoring data quality centrally and data normalization in virtual layer. Besides that, it could be a catalogue of master data [9].

This study will validate the possibility of data virtualization usage as proposed solution to improve MDM in DGT. The validation performed by learning several middleware products of data virtualization both commercial and open sources.

Based information collected from product data sheets or product home page, this study will summarize capabilities of data virtualization middleware as follows [10][11][12][13]:

1. Data sources federation, that is capability to consolidate many data sources in different format, therefore it could be accessed in one location in form of data virtual;
2. Data transformation, capability to transform data into new structure without move the data into new location, so it make data transformation become easier;
3. Data expose, capability to provide reusable data services which could be accessed through different data access protocol such as JDBC, ODBC and web service through HTTP;
4. Performance improvement, capability to improve performance while accessing data by reducing resources usage in database server;
5. Availability, capability to maintain availability of data services;
6. Security control, capability to perform user authentication and authorization and implementation of rule-based access control;
7. Development tools, available integrated development environment (IDE) which support to develop and maintenance virtual database;
8. System management, facility for configuration, deployment, resources monitoring, access monitoring and activity logging.

Virtual database prototype will be developed in this study to validate MDM architectural approach. It will be built using open source product, that is JBoss Teiid. The prototype of virtual database stores metadata of data sources both master, reference and transaction data as shown in Figure 6.
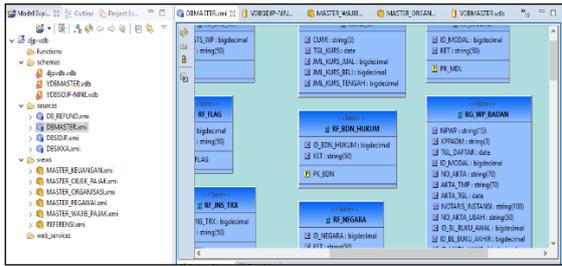
Figure. 6. Virtual Database Modeling (Teiid Designer)

Inheritance of master data entities could be accommodated by creating view. For example, if some entity is specialized from main entity, then it could be created with transformation in form of view at virtual layer, as shown in Figure 7.
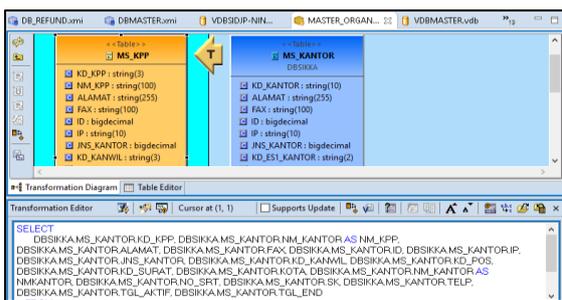
Figure. 7. Data Transformation

There is management console in web based which facilitate system administration of virtual databases including deployment, versioning, privilege, data access control and monitor, as shown in Figure 8.
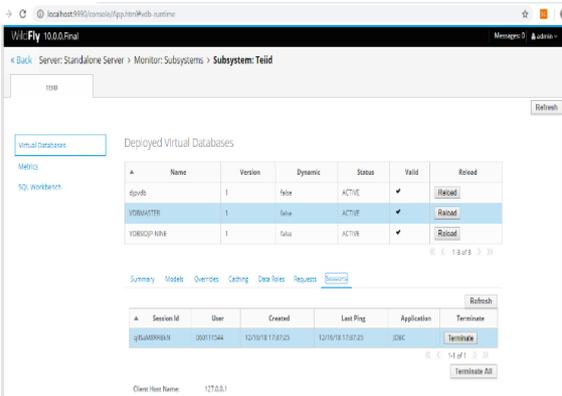
Figure. 8. Data Virtualization System Administration (Wildfly-Teiid)

6. Conclusion

MDM implementation is iteration and step by step process, it will take quite a long time be completed. DGT has started MDM project since 2013. Changes in business and organization affects the MDM that has been applied. Hence, MDM maturity level assessment is needed as first step in MDM improvement project. The assessment could give insight to MDM and become a starting point in MDM improvement activities in this study.

MDM enhancement could be arranged based on requirements. Master data maturity assessment provides systematic approach to get better understanding of current condition of MDM in DGT and to define the target condition of MDM which turn into requirements. Then the requirement will be a reference for define the governance and design of the MDM architecture.

Data virtualization is one of new concepts which introduced in DMBOK V.2 to supports system integration and interoperability. It could make multiple and heterogeneous data sources which spread in many locations seems in one database and accessed at single point, however, that is not replace data warehousing [4]. Hence, the use of data virtualization in MDM project, will makes DGT able to take advantages of centralized MDM system architecture, without having to consolidate all master and reference data into new database, and still able to take advantage from flexibility as in hub and registry architecture.

This study expected to give contribution to the organization by giving recommendation about steps in enhance the MDM. Several previous studies demonstrated MDM in different views. Most of studies focus on developing architecture and practices of MDM. Other studies are proposing MDM maturity model to assess the maturity level of MDM. This study gives more comprehensive view about MDM activities by combines both assessing maturity level and developing architecture and practice of MDM to help organization enhance their MDM system.

References

[1]   Direktorat Jenderal Pajak, *Cetak Biru Teknologi Informasi dan Komunikasi Tahun 2015 - 2019*. Jakarta: Kementerian Keuangan RI, 2018.

[2]   F. Haneem, R. Ali, N. Kama, and S. Basri, "Resolving data duplication, inaccuracy and inconsistency issues using Master Data Management," in *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*, 2017, pp. 1–6.

[3]   M. Spruit and K. Pietzka, "MD3M: The master data management maturity model," *Comput. Human Behav.*, vol. 51, pp. 1068–1076, 2015.

[4]   DAMA, *DAMA-DMBOK2 Framework*, V.2. DAMA International, 2017.

[5]   E. Baghi, S. Schlosser, V. Ebner, B. Otto, and H. Oesterle, "Toward a Decision Model for Master Data Application

Architecture," in *2014 47th Hawaii International Conference on System Sciences*, 2014, pp. 3827–3836.

[6]   D. V. Zúñiga, R. K. Cruz, C. R. Ibañez, F. Dominguez, and J. M. Moguerza, "Master Data Management Maturity Model for the Microfinance Sector in Peru," in *Proceedings of the 2Nd International Conference on Information System and Data Mining*, 2018, pp. 49–53.

[7]   S. Sarkar and S. Das, "A state level policy framework for integrating DFaaS with E-Governance," in *2014 International Conference on Parallel, Distributed and Grid Computing*, 2014, pp. 153–158.

[8]   D. Loshin, *Mater Data Management*. Burlington: Morgan Kaufmann Publishers, 2009.

[9]   S. Earley, "Data Virtualization and Digital Agility," *IT Prof.*, vol. 18, no. 5, pp. 70–72, Sep. 2016.

[10]   IBM, "IBM InfoSphere Classic Federation Server," *IBM InfoSphere Classic Federation Server*, 2018. [Online]. Available: https://www.ibm.com/us-en/marketplace/infosphere-classic-federation-server.

[11]   Cisco, "Cisco Information Server 7.0.5 Data Sheet," *Cisco Information Server 7.0.5 Data Sheet*, 2017. [Online]. Available: https://www.cisco.com/c/dam/en/us/produ cts/collateral/cloud-systems-management/data-virtualization/cisco-information-server-ds.pdf.

[12]   Informatica, "Data Service Data Sheet," *Data Service Data Sheet*, 2017. [Online]. Available: https://www.informatica.com/content/dam /informatica-com/global/amer/us/collateral/data-sheet/data-services_data-sheet_7010.pdf.

[13]   Redhat, "Redhat Jboss Data Virtualization Data Sheet," *Redhat Jboss Data Virtualization Data Sheet*, 2016. [Online]. Available: https://www.redhat.com/cms/managed-files/mi-datavirt-jdv-datasheet-inc0439167-201608at-us-en.pdf.

[14]   T. Dahlberg, A. Lagstedt, and T. Nokkala, "How to address master data complexity in information systems development - A federative approach," in *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018*, 2018.