Architecture Data Warehouse Procurement For Data Communication in Organizations

Kunkun Nur Fauzi[1]
Departement Management of Informatics
Polytechnic Piksi Ganesha, Bandung, Indonesia
Jl. Jendral Gatot Subroto No.301 Bandung, West Java-Indonesia,
Phone.+6222-87340030, Fax.+622287340086.
Email: kunkun.fauzi@gmail.com; website:http://piksi-ganesha-online.ac.id

Abstract---- Data warehouses are designed to help the processing of procurement process and providers in order to deliver the accurate power to facilitate the decision making process in organizations. Suitable to be applied in data warehouse neighborhood due to the development of government procurement emerging technology makes procurement process must be run in an effective, efficient, transparent and non-discriminatory. The design that are used include data collection, analysis of information systems, data analysis and design of data warehouse architecture. In the architectural design stage of the design are the dimensional data model, staging areas / ETL, data warehouse construction.

Keywords: data warehouse, model dimensional, staging area, procurement

I. INTRODUCTION

The data warehouse used for organizations with many branch and also a different geographical location so we need a centralization of data so that the information required will be much easier. With the rapid technological development of the data warehouse has been implemented throughout the organization that do not have branches, it is driven by the need to get information quickly and accurately in a amount very much. Data warehouses can integrate all the data scattered in the organization, although with different platforms.

Government procurement are to obtain procurement activities within the organization that started the process of planning needs until completion of all activities to obtain procurement. Electronic procurement is a procurement conducted by using information technology and electronic transaction that has been adapted to the needs. Enabling electronic government procurement auctions will run with effective, efficient, transparent, competitive, non-discriminatory, open and accountable. Implementation of electronic government procurement produce many benefits, both from the user and the provider of stuff.

II. RESULT AND DISSCUSSION

A. The basic concept of data warehouse

The data warehouse is a database that is read-only analysis and used as the foundation of decision support systems [12]. The data warehouse is a database designed to support decision making within an organization, where the data is in the production of operational data is copied to the data warehouse so that queries can be done without disturbing the performance or stability of the current system [15].

The data warehouse has the characteristics[7]:

1. Subject Oriented. Data warehouse is designed to analyze the data based on specific subjects within the organization, rather than on the specific application or function.
2. Integrated. Data Warehouse can store data that is derived from sources apart into a format that is consistent and integrated with one another.
3. Time-variant. Data warehouse can be said to be accurate or valid at any given time. To view the time interval used to measure the accuracy of a data warehouse, we can use the most simple way is to present the data warehouse at regular intervals, for example, between 5 to 10 years into the future.
4. Non-Volatile. Data warehouse is not updated in real time but in the refresh of the operating system on a regular basis, new data being added as a supplement to the operational database.

B. Elements of the data warehouse

Elements of the data warehouse by federico[5]:

Phase	Basic Elements	Definition
Data Sources	System Resources	Operating system that serves to capture business transactions.
Staging Area	*Staging Area*	Storage area and a set of processes of cleaning, transforming, combining, eliminating duplication, archiving and storage of data for use in the data warehouse.
Integrated	*Data Mart*	Selected data are summarized from the data

		warehouse organization.
	Operational Data Storage	Integrating different systems in the organization at the operational level.
Data Warehouse	*Presentations server*	Target machine where the data is physically stored for the organization and accessed by end users, publishes reports and query usage.
	Dimensional Model	Subject-specific data modeling as an alternative for Entity-Relationship models.
	Relational Online Analytical Process (ROLAP)	RDBMS expansion of the operational mapping multidimensional data is a standard relational operations.
Data Warehouse	*Multidimensional Online Analytical Process (MOLAP)*	Specific database server that retrieves data based on the relational transaction systems and physically store them in a special format to enhance query access.
	Metadata	Metadata or "data about data" is used not only to notify the operator or user of the data warehouse and data status yag information in it, but also the integration of the data from which it comes and to update his tool.
Construction Dimensions	*Online Analytical Process/ Online Analytical Process Cubes*	A type of processing to manipulate and analyze large volumes of data of various multidimensional perspective.
Tool Data Analysis and Application / User for Information Exploration	*Business Process*	Organization's set of business activities that give meaning to the data warehouse business user
	Applications for users	Set of device requesting, analyzing and displaying the desired information
	Data Access Control Tool for End Users	Client data warehouse can be simplified as ad-hoc query system or as complex and sophisticated data mining or modeling application.
	Tools for queries	Special devices to access the data for the end user to access his query, manipulate relational tables and functions directly.
	Modeling applications	Sophisticated devices with the capacity to change shape analysis or understanding of data warehouse output.

Tab.1 Elements of the data warehouse

C. Method Design Architecture Data Warehouse Procurement

The method used is an architectural design process stages of procurement data warehouse.

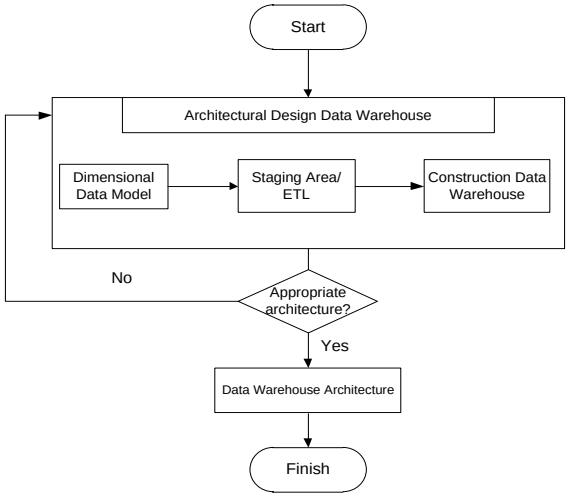

Fig.1 Method Design Architecture

D. Perancangan Arsitektur

The design of procurement data warehouse architecture consists of several stages that determine the dimensional data model, staging areas / ETL, and data warehouse construction. The data source that is built on the same platform and operating system use the same database format. The cornerstone is to define a data source for the data warehouse jobs. On this project must be able to retrieve data from the source system that can be used as important to the decision-making process, followed by the extraction process using software that can extract data stored in the staging area.

1. Dimensional Data Model

Dimensional data model that will be designed referring to the scheme of the data that has been running in the procurement of information systems. The design of data warehouse is used based on the nine stages by Ralph Kimball.

a. Choosing a process

The process used to design data warehouse architecture procurement:

a) Procurement. Procurement is the activity of obtaining items. Documents used for the auction process is the procurement documents.
b) Provider. Provider is an entity that provide the goods at auction process. Documents used in the document provider is a partner.

b. Choosing a grain

Grain is the fact that the data of the candidates can be analyzed, by selecting the grain means determining what is described by the records in the fact table.

Grain	Description	Related Business Processes
Work carried auction package	The amount of packets that have done work on the electronic auction every year.	auctions
Partners who frequently participate in the auction	Amount of auctions has been followed by every partner in every electronic auction.	auctions
Business fields frequent auctions	Amount of businesses that have been done on the electronic auction every year.	auctions
Amount of auctions based on time	Amount of electronic auction every year.	auctions
Origin Partners who participate in the auction	Amount of associates who have been following the auction by area of origin partner.	providers
Long standing partners who participate in the auction	Age of founding partners who have been following the auction.	providers
Amount of partners who often do rebuttal	Amount of objections that have been made by each partner.	providers
Amount of questions based on auction	Amount of questions / aanwijzing on every electronic auction.	auctions
Jumlah pertanyaan berdasarkan rekanan	Amount of questions based on partnership	providers

Tab. 2 Determination of grain

c. Identify and Making the appropriate dimensions

Identify and adjust dimensions of the fact table. The following are the required dimensions:

Dimension	Description
dimension Auction	Auction is conducted electronically, the data needs to save the project name, project work packages and value.
dimension Partner	Partners who participate in the auction in electronic format to store the data needs of the company name, city and date of certificate of incorporation.
Dimension Line of Business	Types of business to be conducted electronic auction, for storing data line of business names.
dimensions of Participation	Participation status of each partner in an electronic auction, to the need to save the data status of the auction.
Disclaimer dimension	Refutation of the electronic auction committee, to answer the needs of storing data rebuttal the auction committee.
dimension aanwijzing	Question and answer process between the partners committee, to question data storage requirements.

Tab. 3 Determination of Dimensional Tables

Each dimension has been linked to the process of determining the grain fact table. The following is a dimension associated with the grain on the fact table, including the following:

1) Dimensions of the procurement process is used for the auction, partners, businesses, participation.

Dimension / Grain	auction	Partner	Line of Business	participation
Work packages are often performed auction	√	√		√
Partners who frequently participate in the auction	√	√		
Business fields frequent auctions	√	√	√	
Amount of auctions a based on time	√		√	

Tab. 4 Table Relation with grain dimension to auctions

2) The dimensions are used to process the provider is a partner, auction, disclaimers, aanwijzing.

Dimension / Grain	auction	Partner	Disclaimer	Aanwijzing
Origin Partners who participate in the auction	√			
Long standing partners who participate in the auction	√			
Amount of partners who often do rebuttal	√		√	
Amount of questions based on auction		√		√
Amount of questions based on partnership	√			√

Table 5 Table Relation with grain dimensions for providers

d. Choosing the Facts

Here's a fact tables are used for the design of data warehouse architecture and procurement of goods, that is:

a) Procurement facts

Attribute	Data types	Dimension
auctionID	Int	dimension auction
memberID	Int	dimension partners
businessesID	Int	line of business dimension
participatedID	Int	dimensions of participation
amountofworkpackagesauctions	Num	dimensional auctions, dimensions partners, participation dimension
amountofpartnersauctions	Num	dimensional auctions, dimensions partner
amountofieldofbusinessauction	Num	dimensional auctions, dimensions partners, line of business dimension
amountoftimeauction	Num	dimensions of the auction, line of business dimensions

Tab. 5 Procurement facts

b) Providers facts

Attribute	Data types	Dimension
auctionID	Int	dimension auction
memberID	Int	dimension partners
disclamersID	Int	dimension disclamers
aanwijzingID	Int	aanwijzingID
OriginAuctionsPartner	Int	dimension partners
StandinglongAuctionsPartner	Num	dimension disclamers
AmountPartnerDisclaimer	Num	Dimension partnership, dimension disclaimer
AmountAuctionsQuestions	Num	Dimension auctions, dimension aanwijzing
AmountPartnerQuestions	Num	dimension disclamers, dimension aanwijzing

Tab. 6 Providers Facts

e. Determining Prekalkulasi Data from Table Facts

Prekalkulasi for procurement data warehouse are:

a) Procurement facts:

1) amountofworkpackagesauctions is the amount package of work that has been included in the auction based group pekerjaan.Data package names package name is taken from the work of the auction table.
2) amountofpartnersauctions is a amount of associates who have been following the auction. The data is taken from the table partner.
3) amountofieldofb usinessauction is the amount area of the business that had been included in the auction. The data is taken from areas of line business table.
4) amountoftimeauction is the amount of auction on a weekly, monthly, yearly. The data is taken from the auction table.

b) Providers facts

1) OriginAuctionsPartner total of auction is based on a weekly, monthly, yearly. The data is taken from the auction table.
2) StandinglongAuctionsPartner is to calculate the date of the founding partners up to date following the auction. The data is taken from the table partner.
3) AmountPartnerDisclaimer is a total refutation based counterparts. The data can be retrieved from the table rebuttal.
4) AmountAuctionsQuestions is the question that follows total auction based business field. The data can be retrieved from the table aanwijzing and line business.
5) AmountPartnerQuestions is the total of the following questions based partner auctions and by line of business of each partner. The data can be retrieved from the table aanwijzing, businesses and partners.

f. Completing the Dimension Table

Table dimensions are used for the design of data warehouse architecture procurement:

a) Dimension Auction

Atribute	Data types	Constraint
auctionID	int	Primary key
nameoftheproject	Char	
usernameservices	Char	
workpackages	Char	
projectvalue	Int	
dealsdoc	Char	Not null

Tab. 7 Dimension auction

b) Dimension partner

Atribute	Data types	Constraint
participatedID	Int	Primary Key
username	Int	Notnull
password	Int	Not null
companyname	Char	
nobusinesslicense	Char	
companyaddress	Char	
city	Char	
phone	Int	
email	Char	
datedcertificateof incorporation	Char	
nameofnotary	Char	
periodbusinesslicense	Char	

Tab. 8 Dimension partner

c) Dimension line of business

Atribute	Data types	Constraint
businessesID	Int	Primary key
namelineofbusiness	Char	Not null

Tab. 9 Dimension Line of Business

d) Dimensions of participation

Atribute	Data types	Constraint
memberID	Int	Primarykey
statusauction	Int	

Tab. 10 Dimensions of participation

e) Dimension disclaimer

Atribute	Data types	Constraint
disclaimerID	Int	Primary key
disclaimerdate	Date	
disclaimerquestion	Text	

Tab. 11 Dimension disclaimer

f) Dimension aanwijzing

Atribute	Data types	Constraint
aanwijzingID	Int	Primary key
questionaanwijzing	Text	
answeraanwijzing	Text	

Tab. 12 dimension aanwijzing

g. Choosing The duration of the Data Base

Based on data analysis for data warehouse procurement needs, then the data can be analyzed to date are pengadaan_elektronik database. For that data can be entered into the data warehouse is the data acquisition is for 2 years.

h. Keep track are slowly changing dimensions.

Attributes of dimension tables do not always have a fixed value or is relatively static. Changes in the dimension attribute values may occur in a long time. Therefore, it is necessary to update it if necessary to maintain consistency and accuracy data.Perubahan all of the dimensions in the dimension tables can be done in 3 ways:

a) The changed dimension attribute rewritten directly. Example: when the project value is changed, then the data value has changed projects directly rewritable.
b) Establishment of a new record for any new changes. Example: The value of the change project will establish a new record in the table the value of the project rev.
c) Change the new column of data that make up the different. Examples: Value addition of a new project on the table nilaiproyekbaru to see changes in the value of the project, so the project can also note the value of the old.

i. Determine priorities and modes of query

Determine priorities and modes of query is to consider the effect on the physical design, such as the existence of summary and the summation.

2. Star schema design

The form of a star schema data warehouse that is on the table consists of facts related to the dimension tables. Here is illustrated for the star schema requirement data warehouse procurement.

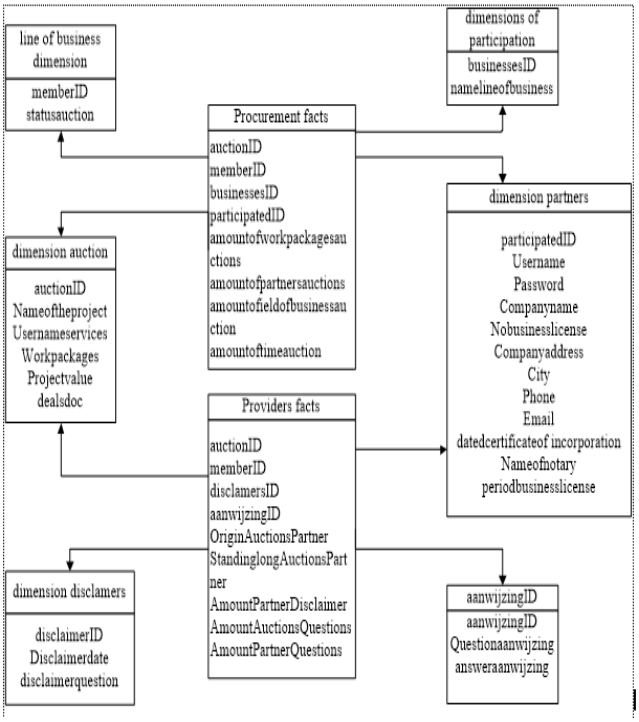

Fig. 2 Stars scheme Data Warehouse Procurement

3. Staging Area

Staging area is a storage area as well as a collection of processes to perform data extraction, cleaning the data, data transformation, data loading and data standardization. Staging areas are used to facilitate the integration and cleansing of data so as to produce quality data. Because there is a process in the Staging area

for data extraction, cleaning the data, data transformation, data loading and data standardization.

Performers ETL	Time	Description
Server	Conducted every day	ETL be done automatically by the server

Tab. 13 ETL Processing

In the ETL process is to extract data from data sources. Data warehouses can combine data from different sources with separate systems that use different data formats. Data extraction is the process of converting data into a format that is useful to the process of transformation. In this process, data is collected in two ways. The first information is downloaded into a separate database. The second way, an application designed to be associated with the old system. For the extraction process using a software that can extract data from many tables that exist in the source system.

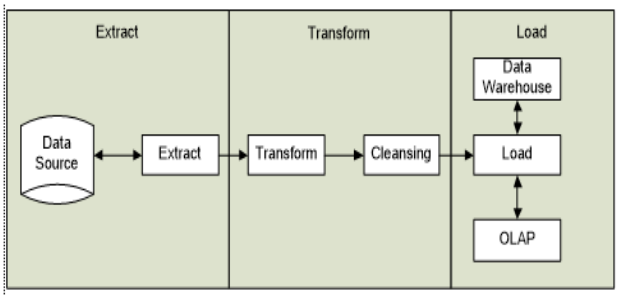

Fig. 3 Architecture ETL

In general, the structure of the database is a good organization, there are only a few things that need to be fixed. The most important thing in this process is to examine the tables that need to be cleared from the information systems of government procurement. Database storage in e-procurement information system using MySQL as a whole to the data warehouse on the procurement is also using MySQL data storage format is the same as operational.

data warehouse ETL process from the data source procurement:

a. Dimension line of business
   Table name : dimensionlineofbusiness
   Description : data and engaged in electronic procurement

Atribute	Data type	Field	Transform
businessesID	Int	businessesID	Copy
namelineofbusiness	Char	namelineofbusiness	Copy

Tab. 14 ETL dimension line of business

b. Dimension Auction
   Table name : dimensionauction
   Description : auction data on electronic procurement

Atribute	Data type	Field	Transform
auctionID	Int	auctionID	Copy
nameoftheproject	Char	nameoftheproject	Copy
usernameservices	Char	usernameservices	Copy
workpackages	Char	workpackages	Copy
projectvalue	Int	projectvalue	Copy
dealsdoc	Char	-	New

Tab. 15 ETL dimension auction

c. Dimension disclaimer
   Table name : dimensiondisclaimer
   Description : disclaimer data on electronic procurement

Tabel 16 Tabel ETL Dimensi Sanggahan

Atribute	Data type	Field	Transform
disclaimerID	Int	disclaimerID	Copy
disclaimerdate	Date	disclaimerdate	Copy
disclaimerquestion	Text	disclaimerquestion	Copy

Tab. 16 ETL Dimension disclaimer

d. Dimensions of participation
   Table name : dimensionsofparticipation
   Description : data on the status of the auction electronic procurement

Tabel 17 Tabel ETL Dimensi Keikutsertaan

Atribute	Data type	Field	Transform
memberID	Int	memberID	Copy
statusauction	Int	statusauction	Copy

Tab. 17 ETL Dimensions of participation

e. Dimension partner
   Table name : dimensionpartner
   Description : partner data in an electronic procurement

Tabel 18 Tabel ETL Dimensi Rekanan

Atribute	Data type	Field	Transform
participatedID	Int	participatedID	Copy
username	Int	username	Copy
password	Int	password	Copy
Companyname	Char	companyname	Copy
nobusinesslicense	Char	nobusinesslicense	Copy
companyaddress	Char	companyaddress	Copy
city	Char	city	Copy
phone	Int	phone	Copy
email	Char	email	Copy
datedcertificateof incorporation	Char	datedcertificateof incorporation	Copy
nameofnotary	Char	nameofnotary	Copy
periodbusinesslicense	Char	periodbusinesslicense	Copy

Tab. 18 ETL Dimension partner

f. Dimension aanwijzing
   Table name : dimensionaanwijzing
   Description : aanwijzing data on electronic procurement

Atribute	Data type	Field	Transform
aanwijzingID	Int	aanwijzingID	Copy

questionaanwijzing	Text	questionaanwijzing	Copy
answeraanwijzing	Text	answeraanwijzing	Copy

Tab. 19 ETL Dimension aanwijzing

g. Procurement facts
Table name : Procurementfacts
Description : procurement data facts on electronic procurement

Atribute	Data type	Field	Transform
auctionID	Int	auctionID	Copy
memberID	Int	memberID	Copy
businessesID	Int	businessesID	Copy
participatedID	Int	participatedID	Copy
amountofworkpackagesauctions	Num	amountofworkpackagesauctions	Create New
amountofpartnersauctions	Num	amountofpartnersauctions	Create New
amountofieldofbusinessauction	Num	amountofieldofbusinessauction	Create New
amountoftimeauction	Num	amountoftimeauction	Create New

Tab. 20 ETL Procurement facts

h. Providers facts
Table name : providersfacts
Description : data providers facts on electronic procurement

Tabel 21 Tabel ETL Fakta Penyedia

Atribute	Data type	Field	Transform
auctionID	Int	auctionID	Copy
memberID	Int	memberID	Copy
disclamersID	Int	disclamersID	Copy
aanwijzingID	Int	aanwijzingID	Copy
OriginAuctionsPartner	Int	OriginAuctionsPartner	Create New
StandinglongAuctionsPartner	Num	StandinglongAuctionsPartner	Create New
AmountPartnerDisclaimer	Num	AmountPartnerDisclaimer	Create New
AmountAuctionsQuestions	Num	AmountAuctionsQuestions	Create New
AmountPartnerQuestions	Num	AmountPartnerQuestions	Create New

4. Construction Data Warehouse

Procurement data warehouse architecture used is a centralized data warehouse architecture. This shape looks like a functional form of the data warehouse, but the first source of data collected in a centralized place, then spread the data into their respective functions, as needed organization. The advantage of this form is completely integrated data is due to the high consistency.

The reason for choosing a centralized data warehouse architecture:

a. Facilitate the monitoring and maintenance of the existing data in the data warehouse because all data is integrated in a storage area.
b. Reduce data redundancy and improve the consistency of the data because the data is managed in a centralized repository.
c. The data stored in the data warehouse is the result of the integration of a variety of different sources so as to provide more reliable.

E. Conclusion

a. Design data warehouse provides an overview of the design of the data warehouse architecture ranging from the data source to create tables for the data warehouse.
b. Data on electronic procurement already have good data so that the data warehouse is only doing calculations in the fact table.
c. Data warehouse procurement scheme is a scheme that is easy to understand so as to facilitate the development process data warehouse, the scheme used for the procurement data warehouse is a star schema.
d. ETL processes in data warehouse just follow the data base used in the operation in order to further facilitate the ETL process.

F. References

[1]. Antonius, Henry dan Eka Widjaja. *Data Warehouse pada Rumah Sakit*. Seminar Nasional Aplikasi Teknologi Informasi. Yogyakarta. 2010

[2]. Arfaoui, Nouha and Jalel Akaichi. *A Data Warehouse Assistant Design System Based On Clover Model. International Journal of Database Management Systems (IJDMS).* Tunisia. 2010

[3]. Connoly, Thomas and Carolyn E. Begg. *Database Systems: A Practical Approach to Design, Implementation, and Management, 4th Edition. Addison Wesley, Longman Inc, USA*. 2005

[4]. Barquini, Ramon. *Planning and designing the warehouse, New Jersey, Prentice-Hall*. 1996

[5]. Domenico, J. A. *Definition of a Data Warehouse environtment in a High Education Institution. Santa Catarina, Florianapolis, Brazil*. 2001

[6]. Harjanto, Prabowo. *Data Warehouse Sebuah Konsep Sistem Manajemen Database untuk Mendukung Keputusan Eksekutif*. 1996

[7]. Inmon, William H. *Building The Data Warehouse 3th Edition, Wiley Publishing, Inc*., 2002

[8]. Jian-bo Wang, Fan Chong-ju. *Research on Airport Data Warehouse Architecture. International Journal of Business, Humanities and Technology. University of Shanghai for Science and Technology. P.R.China*.2012

[9]. Kimball, Ralph and Margy Ross. *The Data Warehouse Toolkit. Second Edition. Wiley. United States of America.* 2002

[10]. Mulyadi, Lubis. *Pengembangan Prototipe Sistem Pengadaan Barang/Jasa untuk proyek konstruksi*, Universitas Islam Indonesia. Yogyakarta. 2006

[11]. Peraturan Presiden Nomor 70 Tahun 2012. Peraturan Kedua Atas Peraturan Presiden Nomor 54 Tahun 2010 Tentang Pengadaan Baran/Jasa Pemerintah.

[12]. Poe, Vidette, et al. *Building A Data Warehouse For Decision Support 2nd Edtion. Prentice Hall PTR; 2 Sub edition,* 1996

[13]. Sahama, Tony .R and Peter R. Croll. *A Data Warehouse Architecture for Clinical Data Warehousing. Queensland University of Technology. Brisbane. Queensland.* 2007

[14]. Singh,Ajit, et al. *Data Warehouse Processes: An Multidimensional Meta Modeling Framework. IFRSA International Journal of Data Warehousing & Mining. India.* 2011

[15]. Singh,R.K, et al. *Simplified Approach for Data Warehouse Quality Measurement. International Journal of Computer Applications. India.* 2010

[16]. Sitanggang, Imas S., et al. *Implementasi Fuzzy Olap Pada Data Potensi Desa Di Provinsi Jawa Barat Tahun 2003 Dan 2006*. Jurnal Ilmiah Ilmu Komputer. IPB. 2008

[17]. Vercellis, Carlo. Business Intelegence : *Data mining and optimization for decision making. John Wiley & Sons, Chichester.* 2009

