

Building of Alumni Data Warehouse Using NDS Architecture for Study Tracer Report in the Form of Dashboard at Muhammadiyah Yogyakarta University

(Building of Alumni Data Warehouse Using NDS Architecture for Study Tracer Report in the Form of a Dashboard at Muhammadiyah Yogyakarta University)

JUANITA YUSANTI, SLAMET RIYADI, ASRONI

ABSTRACT

Setiap makalah harus dimulai dengan suatu abstrak dengan panjang antara 200-250 kata dan Muhammadiyah University Yogyakarta is one of the universities in Indonesia. Every college must have graduates or commonly called alumni. Alumni have an important role for the development of educational institutions. Currently alumni are used as one of the requirements for the university accreditation process by BAN-PT as evaluation material, therefore the importance of alumni monitoring or information, UMY has a tracer study to monitor alumni information using a questionnaire. However, to obtain data from the questionnaire, it is still having trouble because the data is not yet integrated. Integrated data is obtained by building a data warehouse. this research builds a data warehouse using NDS architecture, There are some noise was found that need to be cleaned, such as null data and duplication data. data warehouse alumni has been built and it can meet the requirements of the institution that are displayed in a report to make it easier to analyze reports, the report is made into a dashboard form that is enhanced using Power BI applications

Keywords : Alumni, Accreditation, Data Warehouse, Dashboard

INTRODUCTION

A. Background

Muhammadiyah University Yogyakarta is as one of the institutions of higher education having graduates or commonly referred to as Alumni, Alumni have an important role for the development of educational institutions, because alumni act as a mirror of quality and quality improvement of an institution itself.

Since 2011, the Directorate General of Higher Education of the Ministry of National Education has used study tracking as a monitoring tool for the adaptation of alumni from universities in Indonesia when it enters the workforce. This alumni trace research is not just for the needs of the university, but also as a national and global interest. Therefore tracking the study is considered important because it becomes the material for evaluating the performance of PT and is now used as one of the requirements for completeness of accreditation by the National Accreditation Board of Higher Education (BAN-PT)

At present the study tracking of UMY applies data management using a database (database), but the data applied is still in the form of data that has

not been properly integrated from filling out the questionnaire submitted by UMY alumni, resulting in the possibility of data differences and data noise generated . In addition, tracking study studies will also experience difficulties in making data reports and making decisions or policies, because there are too many different information filled by the alumni.

Based on this description to obtain integrated data, database development is carried out, database development using normalized architecture as reporting with the alumni study store data dashboard form has never been done, therefore the author will use normalized data store architecture in building databases with final results in the form of a report in the form of a dashboard that is published. So the author will make the title of the research on " Building of Alumni Data Warehouse Using NDS Architecture for Study Tracer Report in the Form of a Dashboard at Muhammadiyah Yogyakarta University"

B. Formulation of the problem

According to the background described above, the formulation of the problem in this study is the alumni data from the study tracking study is still in the form of data that has not been integrated and the data is not feasible to be used as a report to meet the accreditation requirements at Muhammadiyah University of Yogyakarta.

C. Goal

Build data warehouse data of alumni Tracer study UMY by using NDS data architecture that can display the data needed by the study tracking study related to its development for UMY Alumni data reporting as one of the requirements for completing accreditation and Producing reporting forms that are easy to understand by using a dashboard power BI.

THEORETICAL

A. Literature Review

research related to data warehouse has been done several times. some references for this research are:

- research with the title Perancangan Data Warehouse Alumni Untuk Mendukung Kebutuhan Informasi Business Placement Centre Universitas AMIKOM by Arik Sofan Tohir and his two friends Kusrini and Sudarmawan (2017) in this design the method used is the Kimball method with the "nine-step methodology" methodology, for database design using the On-Line method Analytical Processing and Nine-steps method, the purpose of using this method is to be able to display multi-dimensional data quickly. The purpose of this design is to produce information on the data of alumni who have worked, both working as entrepreneurs and working as employees. The results of this design obtained an alumni fact information, where the data from the Extract Transform and Load (ETL) process of transactional data is displayed in pivot form .
- According to Dinda Wilanda (2016) her research entitled " Pembangunan Data Warehouse pada PT.PUPUK ISKANDAR MUDA ". The researcher used the Fact constellations modeling method and OLAP (Online Analytical Processing) design, while for the data processing and analysis the researcher used the method with Roll-Up and Drill-Down techniques. The results of the study show that the method of modeling the Fact Constellations scheme and OLAP design are accurately applied to integrate and analyze the company's sales and production data. The aim is to help companies integrate data and analyze sales and production of goods.

- Research by Aulia Paramita (2016) entitled " Perancangan Data Warehouse pada Perpustakaan Yayasan Lentera Insan ". In this study, researchers developed databases in libraries where the source still uses a manual method to connect each separate modules. The need to do this research is adjusted to the needs of library management, this study has the results of data where the data becomes more structured and integrated so as to facilitate the library in making decisions.

B. Data warehouse

Data Warehouse is a place where historical data is differentiated based on the subject as a support for the decision making of an analysis or organization (Malik 2005)

Data Warehouse is a data collection that is analytical, used to be the core of a decision making, Data Warehouse can be said as a relational database because the design of the database refers to the query and analysis of the transaction process, so the Data Warehouse facilitates the analysis work to compile database design methods

C. Characteristic of Data Warehouse

The characteristic of Data Warehouse are:

1. Subject-Oriented

Subject-oriented data warehouse is a database designed to analyze organizational data based on certain subjects, used as decision making and analyze related to the history of the main subjects. But not based on the function or process of the application, because the database is used as data storage to make decisions, not for data-oriented applications.

2. Integrated Data

An integrated data warehouse is a database that can store a lot of data, coming from many different or separate sources in a format that is consistent but can be integrated from one another, thus there is no confusion in the data.

3. Timeline

Data Warehouse Timeline is a database that contains the entire data, from the present value to the historical value, the data will be made later to be made in decision making or analysis, the simplest way is if you will present a database at a certain vulnerable time, for example daily, weekly, monthly, even yearly, for example between 1 to 10 years.

4. Non-Volatile

The Data in an operationa database will periodically move into data warehouse according to the schedule that had been determined, daily, weekly, monthly and the others

D. Extract, Transforming, Loading (ETL)

ETL is a process to produce a Data Warehouse consisting of extracting, transforming, loading, and several processes that need to be carried out before being published into the Data Warehouse. ETL is a system for processing data from one data store, changing data forms, integrating data, and saving to other data stores

E. On-line Analytical Processing (OLAP)

OLAP works with data in the form of multidimensional, usually 3-dimensional shapes which are often referred to as cubes (cube), cube is the main part of OLAP. Cube contains a collection of data that was previously united so that it can quickly find the results of queries. OLAP is also a subject oriented system, its own function is preferred in supporting decision making in databases, so OLAP data cannot be replaced, edited or deleted.

F. NDS + DDS Data Flow Architecture

In the existing architecture on fig 1, there are three data stores namely stage, NDS and DDS. This architecture has similarities with single DDS, the difference is only in this architecture there is data normalization before being loaded into DDS, the goal is NDS can integrate data from several system sources.

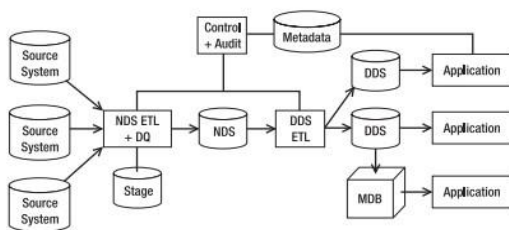


Fig 1. NDS + DDS data architecture (Rainardi, 2008)

The advantage of this NDS + DDS is the making of DDS will be very easy, because NDS runs as a master from the data store which contains complete data sets. This can be used to create a separate static data store for the purpose of analyzing. But the disadvantage is that it takes more extra than a single DDS

architecture because the data from the source entered into the stage needs to be loaded into the NDS before it is loaded into the DDS and then two ETL processes are needed in this architecture.

METHOD

A. Software

This research uses some software in its implementation, that :

- SQL Server Management 2014
- SQL Server Data Tools for Visual Studio 2013
- Microsoft Excel 2016
- Power BI Dekstop

B. Research Procedure

The steps that will be taken in this research are:

- 1) *Determine the subject of data warehouse* : One characteristic of database design is subject-oriented. In this study using the alumni data subject in the tracer study
- 2) *Analyze Needs* : Needs analysis is needed so that the design of this data warehouse has a clear purpose. There are two activities in this needs analysis, namely:

a. Document Analysis

In this research, an analysis is done by outlining the documents and reviewing the documents so that they can draw a conclusion

b. Interview

In this study, interviews were conducted with the Student and Alumni Development Institution (LPKA) which was the user of data tracking studies related to the management of alumni data in the accreditation process

- 3) *Build Data Warehouse*: There are some steps to build a data warehouse. First designing the architecture and ETL process. NDS+DDS architecture design use three data store, which is:

- a) *Stage* : is an internal data store used transforming and preparing the data obtained from the source systems, before the data is loaded to other data stores in a data warehouse

- b) Normalized Data Store (NDS) : is an internal master data store in the form of one or more normalized relational databases for the purpose of integrating data from various source system captured in a stage, before the data is loaded to a user-facing data store
- c) Dimensional Data Store (DDS): is a user-facing data store, in the form one or more relational databases, where the data is arranged in dimensional format for the purpose of supporting analytical queries.

The second step of building a data warehouse is ETL Process. The ETL process here is the process of selecting and retrieving data from the source that has been selected and moved into a stage table that previously defined the needs that will be needed, then carried out the conversion of data noise or data that is messy, so that the data form becomes a form that matches the data warehouse needs .

- 4) *Analyze the Data Warehouse use reporting* : The data warehouses that have been completed are then analyzed by creating a report or reporting in accordance with business needs, to make reporting on the use of cube and dimensions that have been made before, after being successfully created then proceeding with reporting, reporting made using the dashboard feature on the device Power BI software. In making this power reporting dashboard, the data is imported into the SQL Analysis Services Database, importing the data is the cube and dimensional data contained in SQL, then the diagram and information are arranged for the required data requirements.

- 5) *Data Warehouse testing* : The methods of data warehouse testing are :

- a) ETL Testing

This ETL testing warehouse data is very important because ETL brings data from the source system into the data warehouse. If an error occurs with ETL, then the data in the data warehouse is wrong. So that no matter how good the data model is and how good the application is, the data warehouse still cannot be used

- b) Functional Testing

Functional testing is done to ensure that the data warehouse that is made is in accordance with its business needs or not.

The steps taken are by analyzing the data in the data warehouse that has been made into the form of a reporting dashboard and then the results are matched to the needs at the stage of analyzing the needs.

RESULT AND EXPLANATION

A. Research Subject

The research subject refers to the data needs of the accreditation process regarding graduates, namely the UMY study tracking data. The subject of this study was focused on alumni tracking data from 2001 to 2014.

B. Needs Analysis

Before doing database design the stage that is done is defining the needs first. The user needs referred to in this study are the alumni graduate data needed in the accreditation process, to find out the user's needs for the accreditation process carried out the interview method. The interview method produces data requirements which can be seen in Table 1. There is data that is not available in the system, such as the waiting period to get a job, the data will later be calculated according to the compatibility between the graduation month and the start date of work.

Table 1. Interview results are data requirements

Category	type of information
Alumni Data	Alumni profile
	IPK Alumni
	Waiting period to get a job
	Salary Received (amount)
	Suitability of work study program

C. Building Data Warehouse

- 1) Stage

The results of the Needs Analysis have some data that is needed and not needed, so in this stage contains the data needed in the subsequent database creation process. The stage functions as the first data store in this architecture that contains data from data sources to local storage. Data entered into the

Stage has not been made data changes or transformation. Tables name change in Stage data source from source can be seen in Table 2

Table 2. Change the source name on the Stage

Data Source	Stage Data Source
dbo.fakultas	dbo.stage_fakultas
dbo.jenis_pekerjaan	dbo.stage_jenis_pekerjaan
dbo.prodi	dbo.stage_prodi
dbo.salary_range	dbo.stage_salary_range
dbo.send_jobs	dbo.stage_send_jobs
dbo.UserRole	dbo.stage_UserRole
dbo.user	dbo.stage_user
dbo.job_position	dbo.stage_job.position
dbo.alumni	dbo.stage_alumni

2) Normalized Data Store (NDS)

After the data collected in Stage, the next step is ETL Process into NDS. In this process the data cleaned noise in each table. Noise can include null data, duplicating data, and the others. There was a change of naming tables after ETL process that can be seen in table 3.

The next step is making a relation in NDS. NDS relation can be seen in fig 2 uses the nine tables that have been cleared of noise data

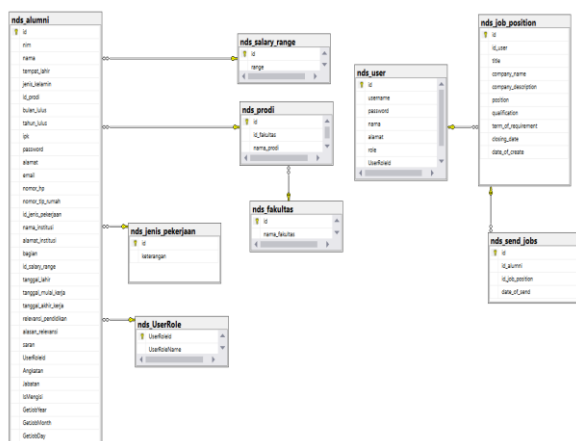


Fig 2. Relationship diagram on NDS

3) Dimensional Data Store (DDS)

Building DDS needs the ETL process from NDS. there are several tables that are not used in making this DDS process, because

the table is not used later for the analysis process based on user needs, of the nine tables, only 7 tables are used in this DDS process, those seven tables are turned into one factual table and six dimensional. the factual table consist of the measurements, metrics or fact of a business process. The following is the diagram relation of DDS on fig 3 the addition of the MasaTunggu column is calculated from the column bulan_lulus and tanggal_mulai_kerja.

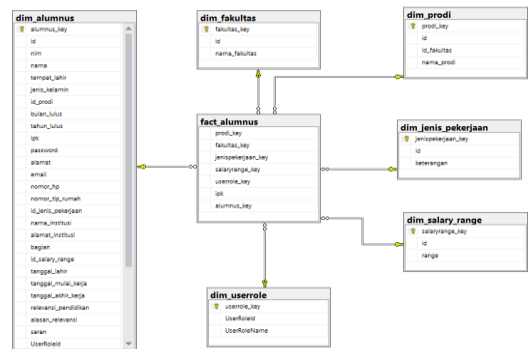


Fig 3. Relationship diagram on DDS

4) Cube Data Warehouse

In the data warehouse there are data that needs to be organized and stored in multidimensional forms, multidimensional data is when it can see a data from various points of view or dimensions, this multidimensional data makes it easy to retrieve data for OLAP

One way to view data with this multidimensional is Cube. Cube here is the main OLAP structure used to view data in the database. The data source of the cube structure can be seen in Fig.4 , there is a cube data source in the form of a star schema

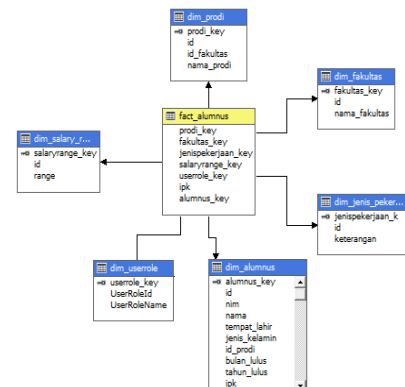


Fig 4. Display the cube data source

D. Analysis and Making of Data Warehouse reports

Database analysis in this study was done by making reporting, reporting is made to make it easier for users to understand and see the results of the data in the database. The information that will be analyzed at this database is presented in cube form on SQL Server Analysis Services.

In making an analysis report on this Data Warehouse Desktop Power BI is used because Power BI Desktop is a software application that can be connected directly to SQL Server Management Studio database, in this application its use is easier and varied, the data reporting process can also be easily published in the form of a dashboard

Information about UMY alumni can be seen from various dimensions, namely alumni, faculties, study programs, types of jobs, salaries, and so forth. As for the needs analysis, the information available on the alumni cube is as follows:

- The total number of alumni who filled out questionnaires from study tracking
- Alumni data (study program, faculty, etc.)
- Student IPK score
- Alumni waiting period to get a job
- Salary received (amount)
- Suitability of study programs with employment.

The report dashboard displayed is based on the user's needs, the report dashboard can be seen in the Fig 5.

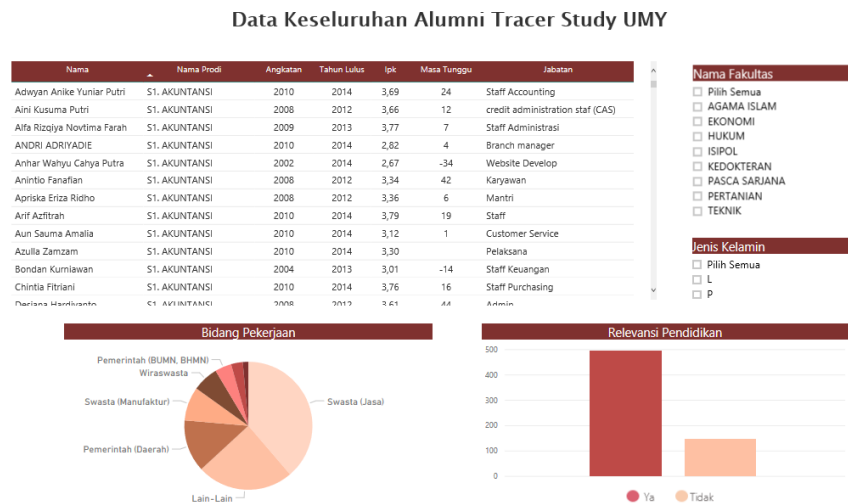


Fig 5. The overall data display is tracer study from 2001 to 2014

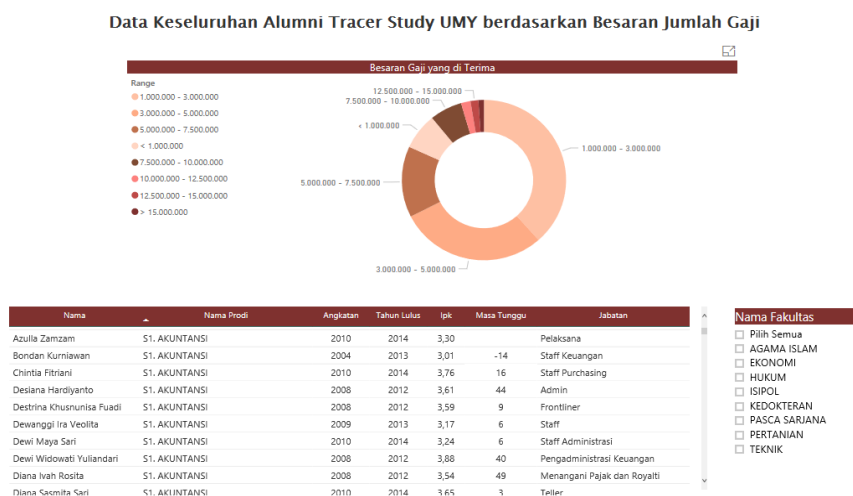


Fig 6. The overall data display of the tracer study is based on the amount of salary

E. Data Warehouse Testing

data warehouse require a test. the testing that has been done is :

1) ETL Testig

Testing on ETL Testing is done by comparing data results. After comparison the data will then produce valid data which means there is a match to the data value.

2) Functional Testing

To ensure that the data warehouse that is built is in accordance with user needs, this test is done by analyzing the dashboard which can be seen on fig 5 and fig 6, whether the data on the reporting dashboard meets the needs of users or not, the test results can be seen in the table

Table 3. Testing Results user needs

Kriteria Pengujian	Result
Displaying Alumni Profile	Fulfilled
Showing Alumni IPK	Fulfilled
Showing the waiting period to get the job	Fulfilled
Showing the salary received (the amount)	Fulfilled
Showing the suitability of the work study program	Fulfilled

CONCLUSION

From the results of analysis, design, and testing that has been carried out from research on Building of Alumni Data Warehouse Using NDS Architecture for Study Tracer Report in the Form of a Dashboard at Muhammadiyah Yogyakarta University, the following conclusions are obtained:

- The construction of a data warehouse tracking study of UMY alumni using the Normalized Data Store (NDS) architecture has been completed and successfully built.
- This data warehouse is used to report alumni data as a process to support

accreditation regarding graduates of Muhammadiyah Yogyakarta University

SUGGESTION

The suggestions submitted for both further research and data owners are as follows:

- In filling out the tracer study data, the information should be further expanded by the existence of socialization for students so that later when they have become graduates of data tracer study at CDC UMY it can be comprehensively covered by all graduates in each faculty.
- In the online data filling questionnaire, to fill in the data that is certain to have content provisions, the dropdown list selection menu is made so that the data recorded in the database is more consistent
- At the time of charging data is made alerts message to the data that must be filled in order to avoid null data or noise on the database

REFERENCE

- Malik, Shadan. 2005. *Enterprise Dashboards: Design and Best Practices for IT*. Hoboken, N.J: Wiley.
- Paramita, Aulia. 2016. "PERANCANGAN DATA WAREHOUSE PADA PERPUSTAKAAN YAYASAN LENTERA INSAN." *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)* 1 (1).
- Tohir, Arik Sofan, Kusri Kusri, dan Sudarmawan Sudarmawan. 2017. "Perancangan Data Warehouse Alumni Untuk Mendukung Kebutuhan Informasi Business Placement Centre Universitas AMIKOM." *E-Proceedings KNS&I STIKOM Bali*, 37-42.
- Wilanda, Dinda. 2016. "Pembangunan Data Warehouse Pada PT. Pupuk Iskandar Muda," Juni. <http://elib.unikom.ac.id/gdl.php?mod=browse&op=read&id=jbptunikompp-gdl-dindawilan-34658>.

