

Analysis and Design of Data Warehouse on Academic STMIK STIKOM Bali

Komang Budiarta^[1], Putu Agung Ananta Wijaya^[2], and Cok Gde Indrapartha^[3]
[1][2] Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University
[3] Department of Electrical and Computer Engineering, Udayana University
Email: komangbudiarta@student.unud.ac.id

Abstract—College accreditation by BAN-PT is one of the parameters in determining the quality of universities in Indonesia. As consideration to achieve the standard from BAN-PT, so they have an evaluation process itself in study program or college to be meet the standard universities when set by the BAN-PT. In carrying out the process of self evaluation, required data source that is used as the basis in assessing on a criteria. In most of the study program, all data spread on the system information and physical document that different, that is require more time and effort to integrate up to interpret. Data warehouse fight important in collecting data that spread and become an information. The process data warehouse with ETL used to integrate, extract, clean, transforming and reload into the data warehouse. With the existence of the data warehouse on Academic STMIK STIKOM Bali can make it easier for executives to get the information to support the standard accreditation standart three and can be used as a reference in decision making.

Index Terms— Data Warehouse, ETL, Accreditation, Standart BAN-PT

I. INTRODUCTION

The information is one assets of an organization or company where is the biggest influence for the survival the organization. STMIK STIKOM Bali is one of the organizations engaged in education, where have a lot of information such as Expense reporting students, lecturers, staff, curriculum and etc. To support the performance of the companies required support full information, quickly and accurately in take a decision.

In this case the most important role in the decision-making is a study program (known as prodi), where prodi is part of a large organization from a university which requires an information system that can be used to obtain, processing and disseminating information to support the daily operational activities. One of the information needs is in the preparation of Accreditation Borang especially on the three standards of students and graduates [1].

On the standard accreditation three have an information about the characteristics of applicants or new students who will be an entrance examination and information about the ratio of applicants to register and accepted on STMIK STIKOM Bali. In additional there is have information about the summary of students active, graduate and the average GPA of graduates.

The information that is required in the accreditation STIKOM STIKOM Bali is still spread on each of division such as new students in the Marketing division, the summary of active students and graduates in Academic division and cert student talent lies in the student bodies, so to collect this information by prodi became to long. Based on this they need a technology that can retrieve information from each of the location of that prodi get the information quickly in a timely and accurate.

The intended technology is using the data warehouse. The Data warehouse is a concept and the combination of the technology used to extract the important ingormasi from data that is spread on some system information into a centralized integrated storage and support the needs of the data history[2].

II. PURPOSE OF PAPER

The purpose of this research is with data warehouse in academic STMIK STIKOM Bali make each prodi can be easier in searching for information and make easy in taking a decision.

III. EASE OF USE

A. The Accreditation based on BAN-PT Standart

The accreditation is the determination of quality standards and the assessment of an educational institution in college by parties outside of the institution itself where assessed by the National Accreditation of Universities according to Law Number 20 Year 2010 the special handling college accreditation assessment. The assessment of the quality of the college based on some evidence such as documents related with the standard that has been agree by expert team when domination the growth of quality in university. There are 7 standard to become appraisers from the BAN-PT in assessing the quality of a program of study college, among other [3] :

1. Standard 1. The Vision, Mission aims and objectives and Attainment strategy.
2. Standard 2. Pamong governance, leadership, management system and guarantee the quality.
3. Standard 3. Students and graduates
4. Standard 4. Human Resources

5. Standard 5. The curriculum, Learning and Academic Atmosphere.
6. Standard 6. Financing, Facilities, Infrastructures and Information System
7. Standard 7. The Ministry of research or devotion to the community and cooperation.

Based on the standard there is a process that must be passed to get the accreditation of institutions BAN-PT . The process of BAN-PT accreditation can be show in figure 1.

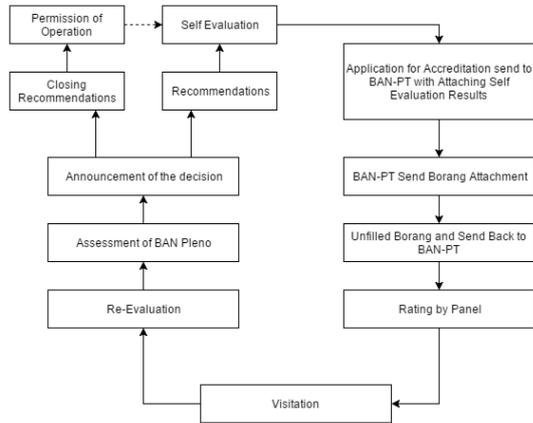


Fig. 1 The process of BAN-PT accreditation

B. Data Warehouse

warehouse typically contains data that represents the business history of a company. The data was collected from various existing applications, then restructured again to be stored in a Relational Database Management System (RDBMS). The data warehouse is the heart and foundation of all information processes because it has one source integrated with the right level of granularity. [4]

Data warehouses allow users to examine and analyze historical data in some form, but the data warehouse can not make a decision. Processes that occur in an operational database do not affect the data warehouse because they have separate databases. In the data warehouse there are some that take the required data, collect, prepare (trans-forming, cleaning, integrating, decoding), store (loading) and provide data for users or applications that are query or reporting [5]. The process to make a data warehouse can be show at figure 2.

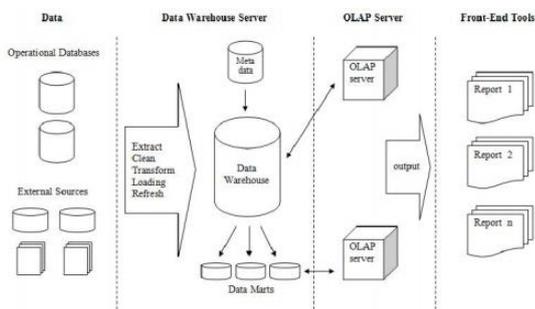


Fig. 2 Process to make a Data Warehouse

Data warehouse and OLAP built from based on multi-dimensional data model. In this model required the fact table and dimension table. The fact table contains the numerical facts that have characteristics: length, lean, and large, and often changed and useful to measure. While the dimensions of the table contains a column that is deskriptif, small short, and the width of which is useful for filtering and based on the attribute dimensions [6]. In dimensional modeling, there are some approaches that are used to create data warehouse, namely:

1. Star scheme [7], this scheme follows the form of stars, where there is a table of faktadi star center with some tables dimensions (dimensional tables) round about. All the table dimensions related to the table of the fact. Primary key on the table dimensions will be key on the fact table or can be Aceh also that fact table has a combination of key from the table dimensions.
2. The snow scheme [7], scheme of pieces of snow is the expansion of the star scheme where this scheme also have one or more dimensions. Only on snowflake, table which relate to the fact table table only main dimension, while the other table connected on the main dimension table.

IV. RESEARCH METHOD

The methods used in the development of the data warehouse on the academic STMIK STIKOM Bali can be seen in figure 3.

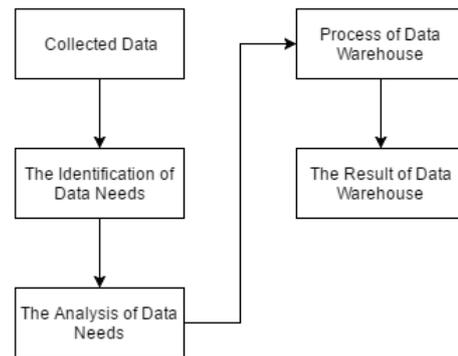


Fig. 3 Research Method

There are have several steps in this research, there is :

1. The process of collecting data, the process of collection data to find information and that can be used in making data warehouse.
2. Identification of Needs Data, data that has been collected further identification to be able to filter the data to be processed.
3. System Requirement Analysis, analysis conducted in the form of system design to be built.
4. Process Data Warehouse, perform data processing from various sources so that it can be used as a source in the form of queries.
5. Results Data Warehouse, data warehouse results in the form of a system that can provide information based on data that has been processed.

V. ANALYSIS AND RESULT

A. Identification of Data Source

Indetification of data sources to be find out the data source to be obtained for the preparation of the standard 3. There are several data sources obtained from various divisions including: Student and Graduate Profile, Student Service Profile, Graduation Evaluation Profile and Alumni Association. Information on the Student Profile and Graduates are available from SINAK (Academic Information System), Excel form of Student Division and New Student Admission System (CAMA). Information on the Student Service Profile there are several sources ie from the division of Counseling in the form of Excel and from the marketing division of Excel. Information on Graduation Evaluation Profile there is a source ie Tracer System of Study and Information on Alumni Association in Alumni system. Identification of data source in detail can be seen in the table 1

TABLE I. IDENTIFICATION OF DATA SOURCE

Data Category		Information	Data Source
a.	Profile of Student and Graduation	Profile of Reguler Student and Graduation in the last 5 years	SINAK
		Profile of Non-Regular Student in the last five years	SINAK
		Profile of student achievement in the last three years in academic and non academic	PK III (Excel)
		Profile data of the summary of regular students in the last seven years	CAMA
b.	Profile of Student Services (Forms of activities, execution and results)	Guidance and counseling	Counseling (Excel)
		Interests and talents (extracurricular)	Not Available
		Coaching Soft Skill	Not Available
		Scholarship	Marketing (Excel)
c.	Graduate Evaluation Profile	Performance Evaluation Graduates by graduates	Not Available
		The average waiting time for graduates gets the first job	Tracer Studi
		Percentage of graduates who work in midwives who are suitably skilled	Tracer Studi
d.	The Alumni Assocation	The Alumni Assocation Profile	Alumni

B. Identification of Data Requirment

The data requirement identification is based on the report format from the accreditation study form, so the next step can be mapping the source reference of the report format with the data source obtained from the application such as SINAK, New Student Admission System (CAMA), Tracer Study and Alumni and some excel from various divisions. The results of data needs analysis and system information developed from the data source are arranged in accordance with the stages of the data warehouse, as in table 2.

TABLE II. IDENTIFICATION OF DATA REQUIRMENT

Data Category	Information	Data Source
New Student	New student data based on last 5 years	CAMA
	New student data based on provinces	CAMA
	New student data based on student from	CAMA
	New student data based on join the selection test	CAMA
	New student data based on not join the selection test	CAMA
	New student data based on summary student	CAMA
Student	Student data based on academic years	SINAK
	Reguler student data	SINAK
	New student data based on transfer	SINAK
Graduates	Graduates data based on GPA	SINAK
	Graduates data based on long study	SINAK
	Minimum, maximun and average of GPA	SINAK
	Graduate have GPA < 2.75	SINAK
	Graduate have GPA 2.75 until 3.5	SINAK
	Graduate have GPA > 3.5	SINAK
Student Achievement	Student achievement data	PK III (Excel)
Percentage of Graduates	Percentage of graduates	Tracer Studi
Alumni	The association of alumni data	Alumni

C. Design of Data Warehouse

Based on the identification that has been done, can be seen that the data obtained lies from various sources such as from Apliakasi Cama, SINAK, Tracer Studies and Alumni and some are from Excel. The data are then made into one that can be called as a data source, then brought to the data warehouse technology. After the data warehouse has been completed then it can be made an application so that it can be used by the executive user in this case is the prodi as depicted in the figure 4.

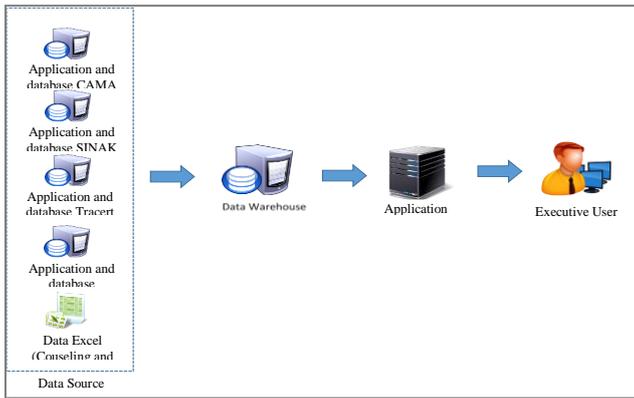


Fig. 4 Design of Data Warehouse

D. Design of Schema Star

Data warehouse is a fundamental approach to data abstraction accumulated in various data sources. Translation of business dimension into logical data model is done by using dimensional modeling. Implementation of dimensional modeling is translated in the form of a star schema which is a schema consisting of fact tables that relate to one or more dimension tables so as to establish a data mart.

The basic schema drafting is the matrix design shown in table 3. Each star schema forms a data mart composed by a fact table which is an event or business process contained in each row and dimension tables that are attributes, rules or perspectives to the columns of the matrix.

TABLE III. MATRIX OF BUSINESS PROCESS AND DIMENTIONS

Business Process	Candidate of student class	Academic years	selection	Registration	High School	Study program	Alumni	Curriculum	Student Value	Transcript	Guidance	Achievement
Capacity of Student Candidate	x	x	x									
Selection of Candidate Student Regular	x	x	x									
New Student Regular and Transfer			x	x	x	x						
Graduate of Regular and Transfer Student		x			x	x		x				
Student Grade		x			x		x	x	x			
Percentage of GPA graduate		x			x		x	x	x	x		
Student Achievement		x						x				x
Guidance					x			x			x	
Alumni Association						x		x				

Each star scema will form a fact table, namely:

1. Fact_capacity_of_sudent_candidate, contains data relating to the information of the candidates new student candidates, including prospective students, classes and academic year.
2. Fact_selection_of_candidate_student_reguler, contains data relating to information about candidate candidates, including prospective students, academic year and selection.
3. Fact_new_student_reguler_and_transfer, contains data relating to the information of new students who are regular and non regular, covering the school year, registration, home sma and majors
4. Fact_graduate_of_student_reguler_and_transfer, contains data relating to the information of regular graduate students and transfer, covering the school year, majors and alumni
5. Fact_student_grade, contains data relating to the value information of each student, covering the school year, majors, courses, students and the value
6. Fact_percentage_of_GPA_graduate, contains data relating to information GPA of graduate students, covering the school year, majors, courses, students, grades and transcripts.
7. Fact_student_achievement. contains data relating to information achievements held by students, covering the school year, students and achievements.
8. Fact_guidance. contains data relating to guidance information conducted by students, including majors, students and guidance.
9. Fact_alumni_association, contains data related to information alumni from STMIK STIKOM Bali, covering alumni and students.

E. ETL (Extraction, Transform, Load) Process

ETL process is a process of data integration from multiple data sources to produce a single point of view of all the data that is integrated. The ETL process requires mapping from a data source into a schema star design. The result of this mapping will be used to specify the fields of dimension tables and fact tables. The ETL process that occurs in the data warehouse can be seen in the figure 5

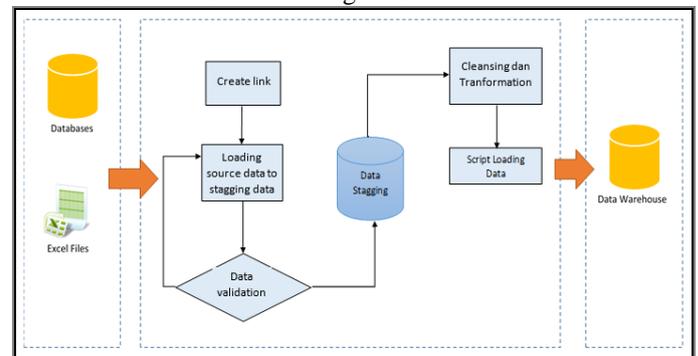


Fig. 5 Process ETL

In the ETL process there are several processes performed, among others:

1. Extranction is sorting data from one source or more data source. In this study there are two data sources namely from the database (DBMS) and Excel Files. Sorting is done based on data needs.
2. Cleansing is data cleaning to ensure the validity, quality and consistency between data and eliminate duplicate data.
3. Transformastion is the process of adjusting the data when there is integration of data from multiple sources to fit the target data warehouse.
4. Loading is loading data into the target data warehouse.

Based on the scheme it can be made an ETL process using a third application called Apatar.

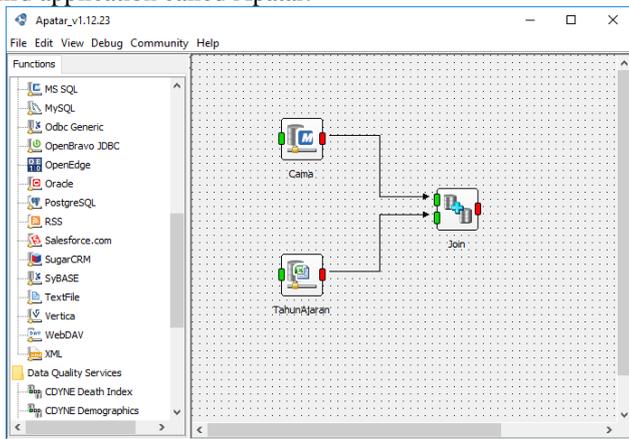


Fig. 6 Apatar to ETL Process

F. Testing of Data Warehouse

The tests are performed on the side of the readiness to support information system that will be built. This test is realized by forming the required reporting by the executive with a needs analysis which has been recorded in table 2. One of the results of the test report is the new student candidates based on the incoming school year can be seen in the figure 7.

No.	angkatan	str_id_om	str_nm_om	bel_transfer	str_id_sekolah
1	2007	2007000096	SUSTI AYU KETUT YULIANINGTENE	0.0	0000
2	2007	2007000108	I MADE DODIE SETIAWAN	0.0	1098
3	2007	2007000113	NANANG DWI KUNICORO	0.0	1487
4	2007	2007000114	DEPIL	0.0	0000
5	2007	2007000122	MUCHAMMAD YUNUS	0.0	0000
6	2007	2007000127	MELANI DARMAWAN	0.0	1107
7	2007	2007000128	I PUTU ERIK DARHANTO	0.0	1492
8	2007	2007000136	LILIH GEDE HENCAHAWATI	0.0	1346
9	2007	2007000140	I PUTU ADI PRATAMA	0.0	2027
10	2007	2007000142	BAWU BUDI PRACOSO	0.0	1005
11	2007	2007000127	I PUTU ASTRY WANGI	0.0	2062
12	2007	200700018	LAURA	0.0	1339
13	2007	200700020	NI MADE DEVI SHAWITRI	0.0	1040
14	2007	200700021	DEVINIA GABE	0.0	1175
15	2007	200700022	I MADE JUNIAWAN	0.0	1010
16	2007	200700023	I GSTI NGR SUEHA PRATAMA	0.0	1497
17	2007	200700029	I MADE SARASINDA	0.0	1059
18	2007	200700030	I GO HENKRA PRASACINATHA	0.0	1059
19	2007	200700031	EKA SUPAKSA	0.0	1010
20	2007	200700032	CATIYARA DASIA	0.0	1034
21	2007	200700033	NI LILIH GEDE WIYU SUSANTI	0.0	1034
22	2007	200700034	I GSTI PT PRAMUDITYA KARISNATA	0.0	1007
23	2007	200700035	ARYA KUSUMADIANA	0.0	1530
24	2007	200700038	I PUTU EKA ADITYA	0.0	1005
25	2007	2007000126	I PUTU NGDEN PRATIWI	0.0	1005
26	2007	200700041	I GUSTI AGLING ADI SANJAYA	0.0	2044

Fig. 7 Example Reporting of new student candidates

This reporting can be made in the form of a desktop application and loaded into the portal in the form of web base so that more easily accessible by the executive. In addition this report can be displayed in the form of pdf, excel, html or

csv file so that it is easy for executives perform analysis process independently using the format of the report is available.

VI. CONCLUSION

Based on the discussion that has been presented, it can be concluded that :

1. The process of the development of the Data Warehouse is done with identifying the needs of the information required by the executives with the approach of the strategy troubleshooting information which can be reported by the executive of study program.
2. The results from the data warehouse berupan report on each of the scheme based on reference reporting format standard accreditation borang 3 that can be displayed in the form of pdf, excel, html or csv file
3. With the existence of the data warehouse on Academic STIMIK STIKOM Bali can memeudahkan executives to get the information to support the standard accreditation standart 3 and can be used as a reference in decision making.

REFERENCES

- [1] Winarto, "Pemanfaatan Data Warehouse Sebagai Sarana Penunjang Penyusunan Borang Akreditasi Standar 3 Pada Fakultas Teknologi Informasi Universitas Budi Luhur", Jurnal Telematika Mkom, Vol.3 No.2, September 2011.
- [2] Deppi Linda, "Pengisian Borang Akreditasi Standar 3 Menggunakan Model Data Warehouse (Studi Kasus Ibi Darmajaya)", Jurnal Informatika, Vol. 15, No. 2, Desember 2015.
- [3] Khusnul Khotimah, Sriyanto, "Perancangan Dan Implementasi Data Warehouse Untuk Mendukung Sistem Akademik (Studi Kasus Pada Stkip Muhammadiyah Kotabumi)", Jurnal TIM Darmajaya Vol. 02 No. 01 Mei 2016.
- [4] Manjunath T. N., Ravindra S. Hegadi, Umesh I. M., and Ravikumar G. K., "Realistic Analysis of Data Warehousing and Data Mining Application in Education Domain", International Journal of Machine Learning and Computing, Vol. 2, No. 4, August 2012.
- [5] Youssef Bassil, "A Data Warehouse Design for A Typical University Information System", Journal of Computer Science & Research (JCSCR), Vol. 1, No. 6, Pages. 12-17, December 2012.
- [6] Khalid Ibrahim Mohammed, "Data Warehouse Design And Implementation Based On Quality Requirements", International Journal of Advances in Engineering & Technology, July, 2014.
- [7] Sonal Sharma, Rajni Jain, "Enhancing Business Intelligence using Data Warehousing: A Multi Case Analysis", International Journal of Advance Research in Computer Science and Management Studies, Vol. 1, Issue 7, December 2013.