

Design Data Warehouse For Centralized Medical Record

Kadek Ary Budi Permana^[1], Gde Brahupadhya Subiksa^[2], and Made Sudarma^[3]

[1][2] Department of Electrical and Computer Engineering, Post Graduate Program, Udayana University

[3] Department of Electrical and Computer Engineering, Udayana University

Email: arybudi1@gmail.com

Abstract- medical record is a note related to service and action which was accepted by the patient on the health facility. Nowadays, data about medical record have been saved electronically both in form of computer file or database file. Today, the medical record can only be accessed by the hospital and the patients themselves. Medical record supposed to be a consideration in taking particular medical action in real situation. The centralized medical record is needed in order to ease the exchange of medical record among hospitals or other institute. The centralized medical record will involve a lot of data from many sources. The design of the data warehouse can be a solution, in which the data warehouse is a place to submit the data or the data which have been submitted systemically so it can be used to help the analysis and report. By designing the centralized medical recorded it is expected that in the future it can give the complex information and analyzing the data.

Keywords - database, centralized medical record, data warehouse.

I. INTRODUCTION

The need of good and fast service and facility in health area is urgently needed in Indonesia mainly from the side patient's medical record. By using the development of information, computer and technology those expectations can be achieved by designing and developing the centralized medical record system. Nowadays, the phenomenon which is happening is the patient who goes to the health center need to do registration related to the data of the patients and later on he/she needs to record the medical data. Those kinds of thing happen repeatedly if the patient goes to different health center. The data of patient's medical record is only used in one medical center which was visited by the patient. The saving system of the patient medical record have been saved electronically in several hospitals, even though there are still some medical center which used the manual saving system. It was done by saving the physic data related to the history of the disease of the patients, actions that was given, and diagnose related to disease that suffered by the patients.

According to the report of strategic plan of Bali in 2016 there were 48 hospitals and 120 health center in Bali [1], while the data related to the registration of new patients was 19.649 in that year [2]. The high number of health center and health facilities will cause the different record in each health facility or health center, besides the use of medical record of the patients cannot be implemented maximally since each of the health center has different medical record. In the other hand, if we see the number patient's which is very huge in a health

center, it needs the big human resources to save and take care of patient's medical record which was organized manually.

Those problems were appeared because the medical record system which has not been centralized at the health center, it makes the use of medical record cannot be used optimally. Another problem is that the medical record which is saved manually costs a lot of money and needs high number of human resources to save and take care of the medical record. If we see from the side of service and facility it can stimulate the negative perceptions from the patients and the users of health facility in which caused by the insufficient service in taking the manual medical record, the speed of medical record sources which is very slow in taking decision and the actions that given to the patients because of the lack of information related to the disease that suffered by the patients or even the malpractice that happens because of the misperception in giving the medicine misdiagnose of the disease.

The centralized medical record could be used by all health facilities by giving input and advices cooperatively in enriching the data of the patients in the centralized medical record. For the contribution of each medical center the data of medical record of the patients will be compiled in form of warehouse. The data warehouse is the storage of information which was submitted from one or more resources which is saved in a place under the integrated scheme by using the centralized storage as the storage of the data will ease the access to the health facility and the user so that it will not depend on the space and time.

II. STUDY LITERATURE

A. Literature Review

Varieties of researches have been conducted related to the use of data warehouse that have been discussed and implemented. The data warehouse can be used for the facility of outpatient and inpatient in Sanglah General Hospital, so that it can give information related to the number of patients from varieties of dimension which have been formed before [3]. In order to support the future service, the planning of patient data warehouse is can also be done in Muhammdiyah Hospital in Palembang. Based on the data analysis the data warehouse can also help in helping the analysis process and in completing the information, because it is fast, accurate for supporting the process of planning, evaluation, and proper decision taking. [4].

B. Medical Record

Medical record is the document related to the patients which is saved in the health center. The medical record consists of notes of patient data which is related to the health service. Those records are very essentials for the health service of the patients, because the complete data can give the information about the decision taking about the medication, handling, and other medical actions [5].

According to the rule of minister of health number 269/Menkes/per/III/2008, in the clause number 1, stated that medical record is document which consists of notes and documents related to the patient identity, checkup, medication, action and other service which was given to the patients. Moreover, the types of medical record stated in written, and completely and clearly in form of electronic system. In determining the content or the data which is saved in medical record, it should minimally contain:

1. Identity of the patient
2. Date and time
3. Result of the anamnesis, at least the complaint and the history of the disease
4. The checkup result and the medical supporting
5. Diagnose
6. Planning of the implementation
7. Treatment or action
8. Other services which was given to the patients
9. For patient which has problems with their teeth should be completed with odontogram.
10. The approval of treatment if it is needed

The purpose of medical record is to reach the well-organized administration in order to increase the health service at the hospitals. Without supported by appropriate and correct management system, it will be impossible that the will organized administration to be achieved as what expected. In the other side, the well-organized administration is one of the factors that takes role in increasing the health service in the hospital [6].

C. Data Warehouse

Data warehouse is storage where the data is organized systematically to be meaningful which is later on will be beneficial in analyzing and reporting. The sources in data warehouse is taken from Online Transaction Processing (OLTP). OLTP is a process which directly oriented to the transaction process. The data warehouse usually save historical data in order to help the decision making [7]. The data warehouse is developed to overcome the technique and business in cases which are related to the use of the data and information to help the decision making of business and management [8].

D. ETL

ETL (extract, transform, and load) system is a system which has ability to connect to the data source, and later on will be able to read the data, change the data, and load the data to a target system. In loading the data to the data warehouse, ETL can is able to integrate, change, and load the data to the Dimensional Data Storage (DSS) [9]. ETL system is not only used to load the data into data warehouse but also ETL system largely and can be used to every kind of movement of the data.

The transformation can be used to change the data based on the format and criteria which is suggested by the target system, to decrease the new values which will be loaded to the target system, and to validate the data from the data sources.

E. Star Schema

Star Schema is a system of data warehouse which is simple enough where this scheme is also called as star schema because the diagram was look like a star. Star schema is a design of table where there is a fact table which is located in the middle of the dimension, this table's function as the connector of other dimension tables. Star Scheme has function as the model of the business scheme accurately by using several fact tables [9]. The strength of the use of star scheme is from the system performance system, where the star schema is more reliable compared to other schemes. It is caused by the use of join level among the fact tables with the less dimension table, as the result it can help the system in doing aggregation operation.

III. RESEARCH METHODOLOGY

A. Data Source

The data sources that was used in designing the data warehouse for the centralized medical record comes from the varieties data of the patients from the hospital. The data from each of the hospital were gained in one centralized place which we called as Centralized Medical Record.

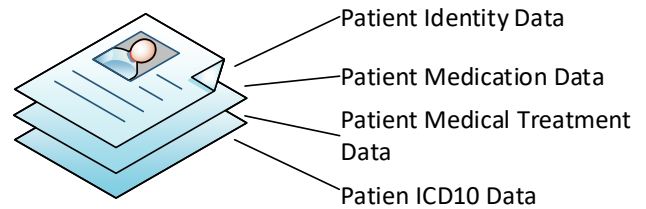


Figure 1. Data Source

The figure 1 shows data which were needed in designing the ware house. The data which is referred here is the data related to the used of the medicine when she/he is in the hospital, the data about the medical treatment which was received by the patients when they were hospitalized, and the ICD10 from every person who had registered. The main concept of the data warehouse is to design the centralized is to design the data of medical record in which later on can be accessed by other hospitals.

The data about the patient's identity is going to be main data from other data. The data related to the identity of the patients from several hospitals can be different. This case could be happened because of the mistype or the error in inputting the data of the patients which was done by the hospital. This case could become a problem that inhibit the development of centralized medical record. In this research we suggest that every data related to the patient's identity should be completed with the identity which has unique character, single, and sit should stick to certain person for life long.

B. The Kimball Lifecycle

Bottom-up approach is recommended to build one data mart per business process. The data bus is the architecture which ensures the capability of various data mart. Data bus requires all the data metrics modeled in a consistent standard data.

Kimball recommends the process of development which have 4 steps for each data mart, where the model of data dimension play important role. The process of model development data dimension includes the facts tables in metric, data and table dimension that modified those data. The model can be used actively by the last user. It helps to ensure the last user will be included actively in developing the data warehouse. Then, the usage and response query time which is make sense is the another purpose of the guideline model of data dimension [10].

Kimball recommends a development methodology which is very unique for data warehouse. It includes the bottom-up which means developing a data mart in every business process which is choosing process. The four steps of the methodology are:

1. Selection of Business Process
Doing the selection of business which is needed by the data warehouse for centralized medical record. this process determined the dimensions and facts which are need to be found.
2. Grain Declaration
The content of the facts table can be identified through this declaration. Grain states the details which are related to the measurement of facts table. By grain declaration it can also be determined the level of detail which is expected in dimension model. The more details included, the level of granularity becomes lower. If the detail was not really included, the level of granularity becomes higher.
3. Determining Dimension Table
The dimension table is needed to answer all questions that are exist in facts table. The dimension table takes role in making the data warehouse be easier to be understood and used., the dimension need to be interpreted and described as clear as possible.
4. Identifying Facts Table
The fact table is a foundation of the data warehouse. The facts table consist of data which can be measured and it is the main target of query in data warehouse. Facts table showed the result of business which have been declared at the beginning.

IV. ANALYSIS AND DESIGN OF DATA WAREHOUSE

A. Selection of Business Process

The first step in planning data warehouse based on Kimball approach is defining the business process and the need which is expected by the users. The expectation of centralized medical record in this research is to know how the information which is received by the patient from one hospital can be used by the hospitals. The information can be a

reference in taking medical action which is going to be taken afterwards. The data exchange such as the use of medicine, medical treatment when the patient was hospitalized become the main focus in this research.

The process of business that was chosen in data warehouse planning can be seen as following:

1. Registration of the patient in hospital
2. The use of the medicine when the patient was hospitalized.
3. The treatment which received by the patient when she/he was hospitalized.
4. Patient's ICD10 data

Through those business process, the data from various hospitals can be submitted. By the submission of those data we do an analysis and planning which has purpose on designing the centralized medical record in form:

1. The record of the use of medicine
2. Record of patient's medical treatment
3. Record of patient's ICD10
4. Data of patient's

B. Architecture of Data Warehouse

Data warehouse which is developed was including several sources of data (hospital data). The sources of the data later on it will be preceded through the ETL which is written in data warehouse. The data will be the congeries of information from various information of data sources. The data warehouse will produce data mart which submit the information in form of particular scheme, data mart will be accessed by the user of centralized medical record to gain the information which is expected.

The architecture of data warehouse referred to two kinds of users, the first one is the hospitals which exchange the information each other and government institutes which is related with it. The hospital users will exchange information each other as well as with the government. the hospital users will get the information related to the patient data which has not been known before, in this case patient got the medical treatment from other hospital. Besides the hospital users, we also refer to another important part, that is the government. Government or the related agency could get information about how the disease spread in each hospital, and it can also be a reference for the government to take the decision. The figure 2 is the further explanation related to the architecture that we have designed before.

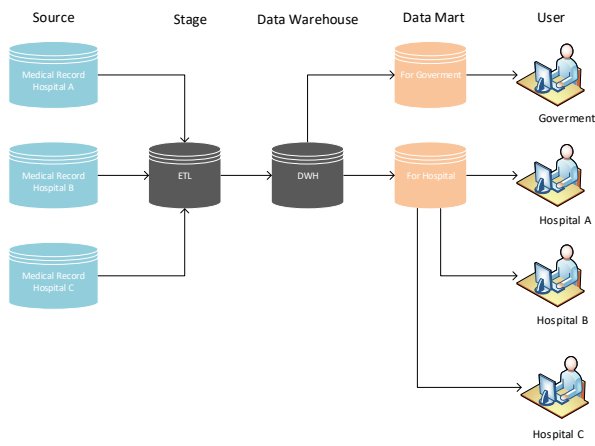


Figure 2. Architecture of data warehouse for centralized medical record

C. Data Warehouse Bus

The business process produces the metric which is translated to the fact table. Most of the facts focused on the result of single business. Choosing the process is an essential thing since deciding the specific target design allows grains, dimension, and facts to be declared. Every business process matches to the row in metrics which can be interpreted in warehouse bus data.

Data bus is the part of Kimball architecture which allow the number of data mart to be an overall data warehouse. Bus architecture is another way to say that every data mart need to use dimension which is suitable to the standard. The basic requirement of dimension which is suitable is that the key, , the name column, definition of attribute, and the attribute value is consistent in the whole business. In other word, two dimensions will look a like if they are similar, or one of them become the perfect subset for other.

Data warehouse bus which included in table 1, interpret how the need which was already explained before connected with the suitable dimension.

Table 1 Data Warehouse Bus Matrix

Business Process	Dimension					
	Patient	Medicine	Medical Treatment	ICD10	Hospital	Time
Patient Medicine	X	X			X	X
Patient Medical Treatment	X		X		X	X
Patient ICD10	X			X	X	X
Patient Distribution				X	X	X

Each rows in metrics bus is the submission of the process business data which is suitable to the unit work to develop ETL system. Every submission of business process data need a special ETL guideline to extract the transaction data and connect it to the suitable dimension.

D. ETL

There are some steps that are need to be done in ETL process, they are:

1. Extraction is filtering the data from one source or more. In this article the source comes from the transactional data which can be find in the hospital. The data can be in form of database (DTMS) and excel files. The filtration is conducted based on the data need which have been explained before.
2. Cleaning is the process of cleaning the data to ensure the validity and consistency among the data and reducing the double transaction data which was already exist. The purpose is to make sure the data which was provided in the hospital is already matched with the need of data warehouse design.
3. Transformation is the process of adapting the data from various sources in order to make it suitable to the data warehouse. This process makes the data which was inputted from the similar format.
4. Loading is the process to load the data into data warehouse.

E. Star Schema

The guideline model will be translated in form of star scheme which is consists of the facts table which is connected to one or more dimension which will form a data mart. Every scheme for, a data mart which is organized by a fact table that become an event or business process which exist in each dimension table and dimension which becomes the attribute, rule, or perspective.

The star scheme which is form such as:

1. Fact_Medicine_Patient which contains data about patient medicine use history. This table is connected to the patient dimension table, time dimension table, and medicine dimension table.
2. Fact_Medical_Treatment_Patient which is consist of the history related to the medical treatment which was received by the patients. This fact table is related to the patient dimension table, time dimension table, and medical treatment dimension table.
3. Fact_ICD_10_Patient which contains the data related to the historical disease of patient ICD 10. This fact table is related to the patient dimension table, time dimension table, and ICD10 dimension.
4. Fact_Distribution_Patient which contains data from each disease in every hospital in form of ICD 10. This data can interpret how the spread of the disease in every hospital. The fact table is related to the hospital dimension, time dimension table and dimension of ICD10.

The star scheme in figure 3 interpret three table of facts, they are Fact_Medicine_Patients, Fact_Medical_Treatment_Patient, and Fact_ICD_10_Patient. This scheme consists of information and form a data mart which can be accessed by hospital to exchange the information each other.

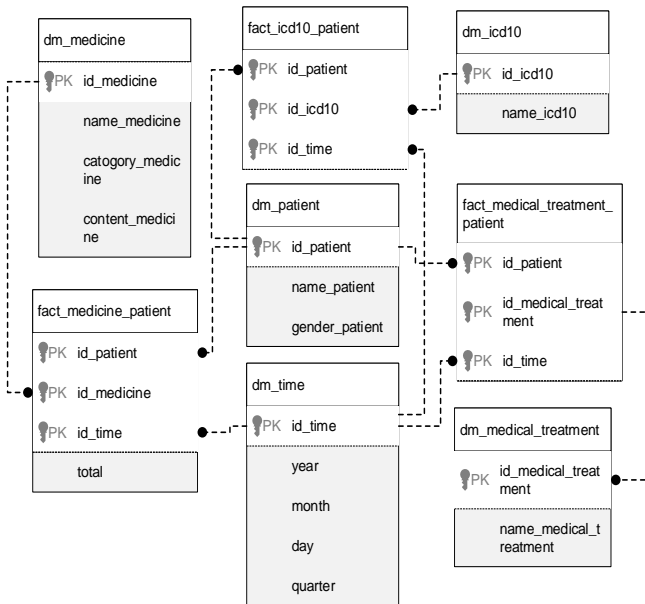


Figure 3. Star Scheme (data mart for the hospital)

In the other hands, star scheme in figure 4 interpret the fact table Fact_Spread_Patient. The data mart which is produced from the scheme in figure 4 is referred to be accessed for government agencies which are related to it.

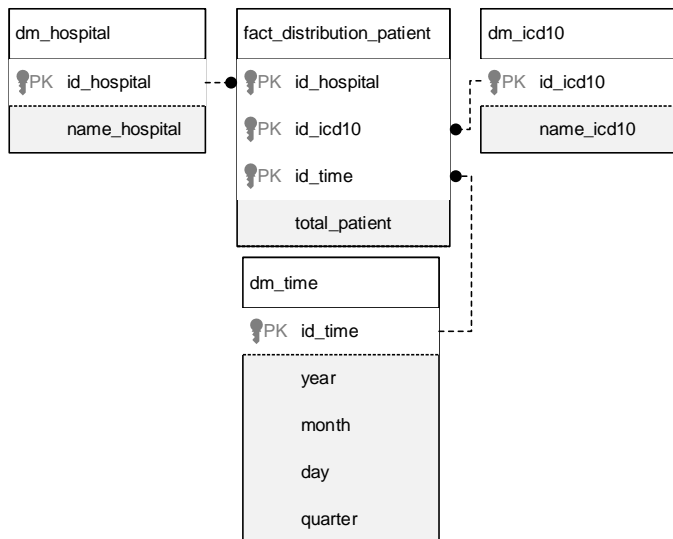


Figure 4 Star Schema (data mart for the government)

V. CONCLUSION AND FUTURE WORK

In designing the centralized data warehouse exchange information from one hospital to other is expected to be executed. The data exchange of the patient’s medicine, and the patient’s medical treatment, and ICD10 data can help the process of medical in real situation. Exchanging the information is expected to be a reference in taking further decision. The centralized medical record can also be beneficial for the government in supervising the spread of disease which can be as a reference in taking decision.

In the further development the design can be developed from the side of the user. The data exchange can be a solution which the centralized medical record can be accessed by the patient (the user). The data like the history of medicine use and history of medical treatment can be supervised by the patient in form of mobile application, which can increase the awareness that health is more important.

REFERENCES

- [1] Dinas Kesehatan Pemerintah Provinsi Bali, “Perubahan Renstra Dinas Kesehatan Provinsi Bali 2013 - 2017,” Pemerintah Provinsi Bali, Juli 2017.
- [2] Rumah Sakit Wangaya Kota Denpasar, “Laporan Tahunan Rumah Sakit Umum Daerah Wangaya Kota Denpasar Tahun 2016”, 2016.
- [3] Putu Veda Andreyana, Putu Angelina Widya, Made Suartika, Designing Data Warehouse At the Inpatient and Outpatient (Case Study: SanglahHospital), International Journal of Engineering and Emerging Technology, Vol. 2, No. 1, January—June 2017
- [4] Rahman, “Perancangan Data Warehouse Pasien pada RS. Muhammadiyah Palembang,” 11 April, 2013
- [5] Endang Lestari, Sistem Informasi Rekam Medik Pada Rumah Sakit Bersalin Graha Rap Tanjung Balai Karimun, JSI Universitas Sriwijaya, 2011
- [6] Novita Yuliani, Jurnal Ilmiah Rekam Medis dan Informatika Kesehatan Analisis Keakuratan Kode Diagnosis Penyakit Commotio Cerebri Pasien Rawat Inap Berdasarkan Icd-10 Rekam Medik Di Rumah Sakit Islam Klaten, 2010
- [7] Suparto Darudiato, Perancangan Data Warehouse Penjualan Untuk Mendukung Kebutuhan Informasi Eksekutif Cemerlang Skin Care. Seminar Nasional Informatika 2010 (semnasIF 2010) ISSN: 1979-2328 UPN “Veteran” Yogyakarta, 22 Mei 2010 E-350, 2010
- [8] Heribertus Himawan, "Implementasi Sistem Business Intelligence Untuk Melakukan Analisis Data Guna Mendukung Pembuatan Keputusan Manajer," Techno.COM, Mei 2008; Vol. 7, No. 1.
- [9] Rainardi, V., Building a Data Warehouse: With Examples in SQL Server. New York: Springer-Verlag New York, Inc., 2008
- [10] Mary Breslin. Data Warehouse Battle of The Giants: Comparing The Basic of The Kimball and Inmon Models. Minneapolis: Business Intelligence Journal, 2004