Vertical Information..

by Julyeta Runtuwene

Submission date: 07-Jan-2020 03:50PM (UTC+0700)
Submission ID: 1239733867
File name: FT_Runtuwene_Artikel_2_Jurnal_Vertical_Information...pdf (801.56K)
Word count: 3845
Character count: 20899

Indonesian Journal of Electrical Engineering and Computer Science
Vol. 6, No. 1, April 2017, pp. 42 ~ 49
DOI: 10.11591/ijeecs.v6.i1.pp42-49

Vertical Information System: A Case Study of Civil Servant Teachers Data in Manado City

Julyeta P. A. Runtuwene, Irene R. H. T. Tangkawarow*
Pendidikan Teknologi Informasi dan Komunikasi, Manado State University, Tondano, Indonesia
*Corresponding author, e-mail: irene.tangkawarow@unima.ac.id

Abstract

Information systems have become important factors in a company, where information systems are useful in providing useful information for the company itself. The information system helps companies implement corporate activities at the beginning of the process until the decision making. In the concept Information system, we know the term of vertical information system, which is another strategy for increasing vertical information capacity. Vertical information system includes the periodic report, written information, and computer-based communications distributed to managers. In this article, discusses the approach in the development of vertical information system which will be preceded by comparison of operational systems, management information systems, and business intelligence systems. We will give explanations about the role in the BI system in an organization, the challenges faced in the implementation of vertical information systems, approaches in vertical information systems, approaches that we choose, and how it implement in civil servant teachers data in Manado City.

Keywords: *vertical information system, periodic report, operational system, management information system, business intelligence system.*

1. Introduction

In defining control mechanisms in an organization, one needs to consider the information base of the organization. This information base provides historical trends and analysis of the data over a period of time. As De Leeuw states in his paper, control must know the overall objectives of the organizational system and should have a clear understanding of the model of the system to enforce the right mix of control. However, this approach may face challenges when the information base of the organization grows beyond proportions and a broader view of the organizational information pool is required.

Such requirement can easily be met by introducing vertical information systems (also called Business Intelligence solutions, or Data Warehouses). Such solution typically allows managers to have a clear picture of their organization's performance within the industry. It allows managers to clearly see key performance indicators (KPIs) on different dimensions (time, region, product, etc.).

2. Operational System vs Management Information System vs Business Intelligence System

2.1. Operational System

An operational system is a term used in data warehousing to refer to a system that is used to process the day-to-day transactions of an organization. These systems are designed so processing of day-to-day transactions is performed efficiently and the integrity of the transaction data is preserved. Operational systems are generally designed to support high-volume transaction processing with minimal back-end reporting. It is also generally process-oriented or process-driven, meaning that they are focused on a specific business process or tasks.

The operating system is generally concerned with current data and optimized to perform fast inserts and update of relatively small volumes of data. Data within operating system are generally updated regularly according to need. They are generally application-specific, resulting in a multitude of partial or non-integrated systems and redundant data and require a non-trivial

Received December 20, 2016; Revised March 2, 2017; Accepted March 20, 2017

level of computing skills among the end-user community. Example tasks: billing, registration, etc.

2.2. Management Information System

A management information system (MIS) is a system or process that provides information needed to manage organizations effectively. The term MIS appears to depict a variety of applications is developed for the manager where the application is to provide information about sales, inventories and other data that will assist in the process of managing the company. An 'MIS' is a planned system of the collection, processing, storage and dissemination of data in the form of information needed to carry out the management functions. The terms MIS and information system are often confused. Information systems include systems that are not intended for decision making.

2.3. Business Intelligence System

Business intelligence (BI) refers to specialized system for retrieving decision support information, by analyzing huge amounts of data to help decision makers improve their performance of their company of organization, gain competitive edge, and optimize business process. BI technologies provide historical, current, and predictive views of business operations. Common functions of business intelligence technologies are reporting, online analytical processing, analytics, data mining, business performance management, benchmarking, text mining, and predictive analytics. Business intelligence aims to support better business decision-making. Thus a BI system can be called a decision support system (DSS).

BI uses technologies, processes, and applications to analyze mostly internally, structured data and business processes while competitive intelligence gathers, analyzes and disseminates information with a topical focus on company competitors. Business intelligence understood broadly can include the subset of competitive intelligence. Business Intelligence System (BIS) has become the essential tools in helping management make decision in most organizations. With the vast amount of data collected through the transactions made through the years, organizations can stay ahead of competition. Business Intelligence is also supported in strategic planning and processes.

Discussion about Business Intelligence solution wouldn't be complete without introducing the three heavyweights in the BI community. Ralph Kimball, Bill Inmon and Dan Lindsted are leading specialists who have their own different theories on how to best go towards a business intelligence solution. Kimball believes that organizational information should be put in vertical information systems at the atomic level, but in a de-normalized manner, whereas Inmon argues that such endeavors should be attempted in a normalized way as done in operational systems. For both, Lindsted argues that a hybrid approach is the best path since Kimball's denormalized approach is pertinent only for small data warehouses whereas Inmon's approach is pertinent to operational systems. The BI community seems to fall in the path of Lindsted's approach recently. Here are highlights from approaches proposed by these BI gurus:

Table 1. The Highlights From Approaches Proposed By Bi Gurus

Bill Inmon	Ralph Kimball	Dan Lindsted
Firstly, process the design with carefully, accommodate the user need from that time to next time and then built the data warehouse.	Built the data mart in departments which have needs and initiatives, then process the integration data mart if needed.	A collection of detail-oriented, history-tracing and uniquely linked collection of normalized tables.
Data marts sourced from data warehouse after the data warehouse builds, data marts which in departments take the information source from one (and only one) data warehouse.	The data warehouse is a collection of data marts. It means, first built the data marts and then built the data warehouse.	A hybrid approach combining the best of 3NF and dimensional modelling.
Architecture name is Corporate Information Factory (CIF)	Name architecture is bus data warehouse.	Dan argues that his Data Vault is not architecture; it is more like a standard.
Information stored in the 3NF relational form.	Information stored in multidimensional database.	Flexible, scalable, consistent and adjustable to business need.

3. Challenges of implementing Vertical Information Systems

Business intelligence has changed the fate of many big organizations, efficient business intelligence systems in place have made the manger's life little easier by providing accurate and concise information which helps to make a better decision. Vertical Information systems can provide unprecedented information to the managers by the aid of which managers are able to make thoughtful decisions that is best for the company.

Though it looks all good from the outside and with amazing success stories to back up the fact that implementing an effective business intelligence solution is beneficial for an organization it's not sans challenges to implement one functional and useful solution. Basically the challenges that come into significance are related to the business organization and how they function. If the business processes the business organization do not comply with any standard and are haphazard then it will not be possible to build a vertical information system to serve any meaningful purpose.

Due to the size and the enormous scope of the vertical information system in the form of a business intelligence solution, it requires a very high political will to implement the solution. It also involves a high degree of time, effort which equates to soaring costs to build these kinds of systems. This is a prominently evident challenge that as the costs and resource investment goes high on a system that is not yielding anything fast the management has to rethink about investing more funds and resources to the system. So the management should be committed to get a vertical information system that will eventually help to make better decisions resulting in higher turnover of the company.

Another challenge in implementing a business intelligence solution that commands high priority is that the system should be focused on the business and management aspect rather than just the IT aspect. As the system is generally owned and maintained by the IT department, but its main users are the management of the company. So making it usable and focused on the information needs of the user/business executive makes the system sustainable and easily accepted by the end user.

Data availability and data quality are two big questions to ask if a vertical information system is to be put in practice. In some Online transaction processing (OLTP) system the vendor might prohibit other systems from accessing the database. Chances are there might not be any proper system in place to get the data from. Even if there is sample data, guaranteeing data quality to get the accurate and desirable results needed by the management is a big challenge in itself.

Another practical challenge is to get over the silver bullet syndrome. Implementing business intelligence will solve all the problems of the business, it will give answer to any data and information related question one can ask is the general assumption of the user. The over expectations should be trimmed down on the realistic level, the realistic picture and the things that can be realized should be communicated. Managing this mountain of expectation of the user that business intelligence solution is the silver bullet for all business problems is a hand task to execute.

4. Approaches to the Vertical Information System

There exists approaches to build and implement vertical information system in the form of a Business Intelligence solution, the approaches basically comprise of:

1. Top down Approach: the vertical information system is made because the management layer feels the need of the system. It is derived from the mission, objective and vision of the organization. This approach is very positive for the system because as management is committed and wants to implement the system, there is easy to implement this system and the financial factors are also covered as it's the management who wants the system.
2. Data Driven Approach: is based on the assumption that the data already exists in the OLTP system or some external database which can be accessed by the organization. In this approach the organization wants to analyze the existing data and derive a vertical information system to assist the business executives to make meaningful and logical decisions. In this approach the data history and forecasting should be kept in mind to give out the results demanded by the management.
3. Process driven approach: is based on the business processes. Processes that run on the organization define the organization. The success of a company can simply be measured by

determining how efficient its processes are. Success largely depends on the processes being carried out, so it this approach to implement a vertical information system helps to do the things on a horizontal process level.

4. Off the shelf: seems to be the easiest approach where the business can buy a rebuilt solution and implement it. It is a tool as well as a solution. The strict financial regulation on how the financial reports should be prepared decides how the reports should be. A business intelligence solution in this approach is benchmarking devices in itself as the same solution can work out for company X and your competition, but it might not be the right solution for your business.

5. Suggested Approach for Vertical Information System Development

In his widely acclaimed book entitled 'The Data Warehouse Toolkit', Ralph Kimball [1] strongly argues that the job of a data warehouse manager is basically similar to the job of a publishing editor-in-chief of a magazine. As the editor-in-chief's job ranges from basic lay outing to customer demographic study and maintaining reader's trust, the IT manager's job involved in BI developments follows a similar path. BI developer teams need to understand the users of business area and their business processes, publish the right data at the right time, and maintain the trust of executives.

In light of these facts, we believe that the process-oriented way of BI development should be the leading approach used. In fact, the very first task in developing a vertical information system should be selecting the business process to model. We need to understand the business process of organizations. In following this approach, however, we need to understand that a business process DOES NOT mean a business function or a business department.

To highlight this point further, let's take a case of a company that is trying to implement a VI system for its order's management process. The first question we should be asking in addressing such a requirement should be what is the business process and with what source system is it supported. A wrong question to ask would be 'which department needs this vertical information system'. The reason behind this is that the order management can be used by both Marketing department and Sales department. These are two different organizational units with the same interest in the final published data. If we treat them separately then there will be duplication of resources and data and inconsistencies may occur. In BI world, the best way to ensure consistency of data is to publish the right data once. A single publishing run not only reduces our technical efforts in getting the data to the user, but it also delivers the same information to all users (irrespective of their departments).

In his book, Kimball [1 pg.32] interestingly argues against the data-driven approach as follows:

"We strongly encourage you to resist the temptation to model the data by looking at the source data files alone. We realize that it may be much less intimidating to dive into the file layouts and copy-books rather than interview a business person; however they are no substitute for user input. Unfortunately, many organizations have attempted this path-of-least-resistance data-driven approach, but without much success."

We also argue that the process-drive approach cannot succeed in its own without concerted effort by company management. Unlike operational systems that capture the nitty-gritty details of the operational business world, vertical information systems are targeted towards the business people in their posh offices in 10th floor. These people may need to exercise their political muscles to get the BI solution live. The information dealt is most sensitive, even within the organization so successful with such endeavor needs management commitment and vision. To address this we need to couple a top-down approach and convince the business executives and get their commitment in this regard.

6. Implementation of Business Intelligence System in ETL Process using Pentaho Data Integration

BI system changes the enterprise data, e.g. operational data, transactional data or others (On Line Transactional Processing-OLTP) into the dashboard or graphical view. This application analyzes data in the past, then using dashboard or graphic to support the decision

and support the enterprise plan. BI is equal to briefing books, report and query tools, and executive information system.

The data that we use for the implementation is data of civil servant teacher in Manado city. This data we use for show all the implementation of Business Intelligence System and Pentaho has chosen as the open source BI application.

Extraction, Transformation, Loading (ETL) process is the important process to built data warehouse. A data warehouse is needed to store all transaction data (OLTP) before it analyzes into On Line Analytical Processing (OLAP). In General, the benefits of ETL are as follows:

1. Transform data from OLTP to OLAP (data warehouse)
2. Integration between applications
3. Clean data
4. Data migration
5. Export data

The ETL processes are inextricably linked to the data source and the database, which consists of:

1. Extract the data from an external source. An outside source can be a database, csv, excel, xml, web services, database, etc.
2. Transform (Edit) the data so as to suit the needs (can also incorporate elements of data quality).
3. Load (entering) the data into the target end. The final target could either be csv, excel, databases, XML, web services, database, etc.

Pentaho Data Integration (PDI) or Kettle ETL utility is open source under Pentaho Corp Amerika. This was originally an initiative of Matt Casters, a programmer and consultant Business Intelligence (BI) from Belgium who has managed projects for enterprise BI.

Currently Kettle ETL is a utility that is very popular and one of the best on the market. Some advantages are as follows:

1. Have a collection of data processing modules that quite a lot. More than 100 modules or step.
2. Have a module that facilitates the design of the data warehouse model as Slowly Changing Dimensions Dimension and Junk.
3. Performance and scalability are well-proven.
4. Can be developed with a variety of additional plugins.
5. Utility Kettle to be used in the integration of this data using the Spoon

In the multidimensional model, a database consists of several facts table and dimension table. A fact table contains the value of the aggregation that is the basis of measurement (measure) as well as several keys that are related to the tables of dimensions that would be the point of view of the measure.

In the process, the arrangement of the fact table and the dimension table has a standard design or schema because it is proven to improve performance and ease of translation to the OLAP system.

The schema which served as the basis for performing data warehousing. Two of the most common schema used by the various OLAP engine is the star schema (Star Schema) and snow grain scheme (Snowflake Schema).

We use star schema to show the multidimensional model, for civil servant teachers' data. Star schema based on a single fact table surrounded by one or more dimension tables as 'branches' so that look like stars. Each branching off at one of the tables of dimensions. Or in other words the table dimensions with this schema are all in the form of leaves and has no other branches.

Pentaho Data Integration which called Kettle used for an integrated diversity civil servant teacher's data available in the database. In this case, database platform we use MySQL. Below show all the dimension tables of data warehouse civil servant teachers in Manado city:

Dimension table Guru (dim_guru)

Data source of dim_guru table is derived from a list of names of teachers who registered as civil servent teachers in Manado city until 2016. Where the total registered is 4208 teachers. Here is an overview transformation design to use PDI-Kettle.

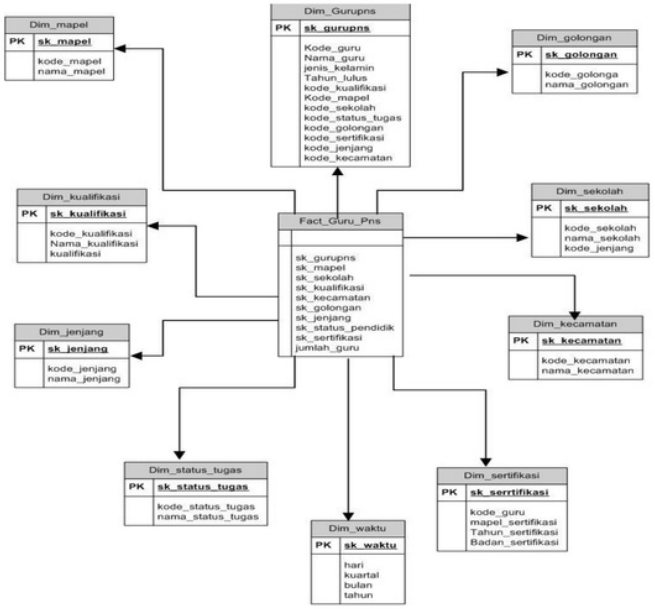

Figure 1. Star schema of data warehouse civil servant teachers in Manado city

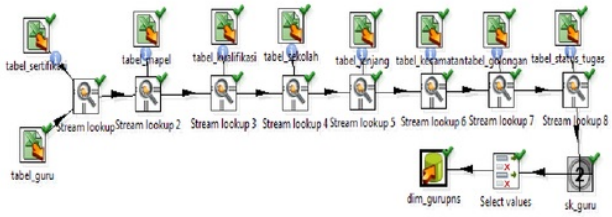

Figure 2. The transformation to a table dim_guru

Once the transformation is resulted student dimension tables as follows:

Table 1. Teachers Dimension Table

First rows | Last rows | Off

#	sk_guru	kode_guru	nama_guru	jenis_kelamin	kode_kualifikasi	thn_lulus	kode_mapel	kode_sekolah
1	1	123456789111	NELTJEJOHANIS	P	12	1991	200000000	6111
2	2	123456789112	MAGDALENA MATHIAS	P	12	1987	200000000	6150
3	3	123456789113	ROSMIATI S. SALALE	P	12	1988	200000000	6150
4	4	123456789114	PATRISIA ELLA	P	18	2014	200000000	6151
5	5	123456789115	THERESIA JULIEN AGUSTINA SAMURA	P	12	1988	200000000	6151
6	6	123456789116	SURIATY DONIO	P	18	2013	200000000	6182
7	7	123456789117	TIRSA PUDE	P	12	1991	200000000	6200
8	8	123456789118	DETTY SALINDEHO	P	12	1981	200000000	6201
9	9	123456789119	ADRIANA TAWARIS	P	18	2014	200000000	6201
10	10	123456789120	Emmy Annie Ropa	P	18	2012	200000000	6127
11	11	123456789121	Deiji C.S Lolowang	P	18	2006	200000000	6127
12	12	123456789122	Meyke Rampengan	P	18	2010	200000000	6128
13	13	123456789123	SULASTRI LAHIFE	P	13	1977	200000000	6134
14	14	123456789124	MARTHINA MAJUSIP	P	12	1984	200000000	6142
15	15	123456789125	Santje Sinjal	P	12	1984	200000000	6142
16	16	123456789126	Altje Amye Panderot	P	18	2006	200000000	6143
17	17	123456789127	Kening Aweeh	P	18	2005	200000000	6143
18	18	123456789128	Sry Lenny Semuruk	P	18	2011	200000000	6143
19	19	123456789129	SANGIAN MARLINA KETTY	P	12	1983	200000000	6143
20	20	123456789130	Meifi Telap	P	18	2010	200000000	6143
21	21	123456789131	INEKE KARAUWAN	P	18	2010	200000000	6152
22	22	123456789132	LAKSMY EVELIN LINELEJAN	P	12	1981	1000110170	6152
23	23	123456789133	ELSJE ESMI LUMENDEK	P	13	1984	200000000	6153

Cube is the logical data modeling from a dimension to multidimensional. Before create dashboard, we need to pass the MDX Query to Crete cube with schema workbench application. Cube guru_pns created shown in Table 2, shown the total teachers in each district. In Manado city consists of 9 districts (*kecamatan*). Every district has more than a hundred teachers.

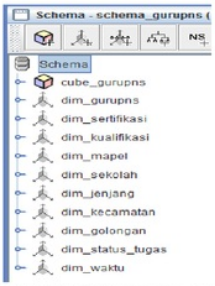

Figure 3. Cube guru_pns

Table 2. Total teacher in each district

nama_kecamatan	jumlah_guru
Bunaken	137
Malalayang	466
Mapanget	448
Sario	352
Singkil	165
Tikala	824
Tuminting	475
Wanea	663
Wenang	601

The result of this implementation is a dashboard which represent all the dimension created before. To show the dashboard we use BI server Plugin Saiku. This plugin is more user friendly and easy to show the chart, diagram and table. Figure 4 show the pie graph of the total teacher in each district. Figure 5 shown the graph of total teacher each subject and each their last degree

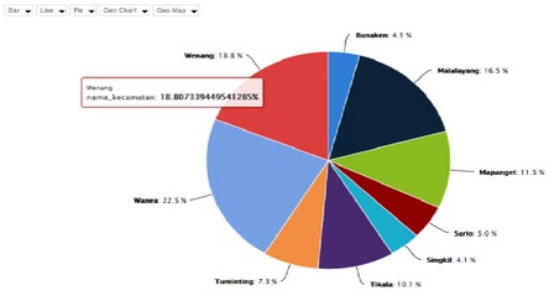

Figure 4. Pie graph of the total teachers every district in Manado city

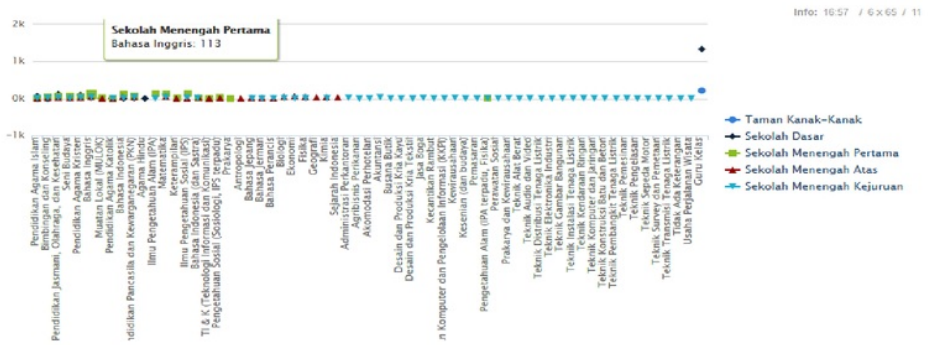

Figure 5. Graph of the teacher in each subject by their last degree

7. Conclusion

Vertical information systems can help an organization address its ever-increasing uncertainties within the organization and with the outside environment. They become even more useful when the information base of an organization grows.

Several approaches can be used in designing BI systems. However, we feel the best approach is the process-oriented approach coupled with management commitment. Since management is the "real" user of such systems and since the information at such aggregated, 'higher', levels are sensitive, full commitment of management should be harnessed to fully realize a vertical information system.

Going into the intricacies of the data oriented approach at an early stage of BI development will end up in a maze and will not address the key requirements of business people. In addition, such approach may suffer from data inconsistencies as data may be replicated by several processes.

Considering the above facts, the business oriented process provides a platform for consistent data publishing by the BI system. In addition, the business oriented approach ensures that data belonging to one business process will only be published once, irrespective of the number of organizational departments who will be using it.

Using a BI system like Pentaho for doing the ETL process to solve the problem of the diversity of data sources. After that, the data will be stored in a database called the datawarehouse. This an ongoing project for creating mdx-query, cube and dashboard of data warehouse of civil servant teachers in Manado city.

References

[1] Kimball Ralph, Ross Margy. The Data Warehouse Toolkit. Second Edition. Canada: John Wiley & Sons, Inc. Canada. 2002.
[2] Kimball & Caserta, The Data Warehouse ETL Toolkit, 2004
[3] Thia Feris. Pentaho Knowledge Based (http://pentaho-en.phi-integration.com)
[4] https://online.han.nl/sites/5-ICA-ALG/MasterMISD/Business%20Intelligence%20with%20OLAP/Presentations/business_intelligence.pdf
[5] http://www.rpi.edu/datawarehouse/dw-about-cmp.html
[6] Business Intelligence with OLAP study material-Data Vault 08-09. Pdf

Vertical Information..

ORIGINALITY REPORT

23%	14%	13%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

2%

★ J P A Runtuwene, I R H T Tangkawarow, C T M Manoppo, R J Salaki. "A Comparative Analysis of Extract, Transformation and Loading (ETL) Process", IOP Conference Series: Materials Science and Engineering, 2018

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		