Journal of Theoretical and Applied Information Technology
30th June 2014. Vol. 64 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPLEMENTATION OF DATAWAREHOUSE, DATAMINING AND DASHBOARD FOR HIGHER EDUCATION

[1]RUDY, [2]EKA MIRANDA, [3]ELI SURYANI
[1,2,3]Information Systems Department, Bina Nusantara University
E-mail: [1]rudy2105@yahoo.com , [2]ekamiranda@yahoo.com

ABSTRACT

Today, most of the higher education institutions have been supported by integrated information systems for all activities, such as lecturer recruitment, marketing, student admission, student registration, academic operations, and alumni system. Higher education information system generates many transactional data; and the amount of those data is increased day by day, but the use of data has not been explored optimally. Each data stored at data storage media is issued for separate purposes. On the other hand, a higher education organization needs to gain overall knowledge from all data. Based on these conditions, organization needs an analytical tool to extract information and discover valuable knowledge from huge amount of data. These processes require more processing time and more complex process. The purpose of this research was to design the data warehouse model and dashboard for analytical tool and use data mining technique for higher education institutions. The research method began with designing the knowledge needs; designing and developing transformation model to data warehouse; using data mining techniques and designing dashboard for analytical tool. The obtained result was the model of data warehouse, data mining and analytical tool for higher education institution that will improve institution's performance, and help them to improve analytical and decision making process.

Keywords: *Datawarehouse, Data Mining, Analytic Tools, Higher Education, Dashboard*

1. INTRODUCTION

Nowadays in Indonesia, number of higher education institutions has reached more than 3200 institutions; and in the Jakarta as a Capital City, number of higher education institution are 320 institutions. Higher education institutions challenged with the pressure to improve the quality of education process and management. Higher education is sought to invest more resources in tools that allow them to collect and manage information directly [4]. Higher education stakeholders like parents, students, funding bodies, and government would be interested in quality of education offered by higher education. This quality can be found on various sources of data in higher education institutions [3]. Higher education institutions look for efficient technology to manage and support decision making procedures or assist them to set strategies. The quality of higher education processes and managerial system can be improve by providing new knowledge, that can be extracted from historical and operational data [5].

Every organization including universities, desperately need accurate and relevant information. In fact, the data owned by institutions are way beyond their expectation due to lack of usage of the information systems. Higher Education institutions should take evaluation process continously to improve their quality and carry out the mission of Tridharma Perguruan Tinggi. Evaluation process needs internal and external data. Hence, evaluation process needs database technology, data warehouse and data mining to help them optimizing the data. Many Higher Education that have used information systems could not explore their data to support decision making process [15]. Many organizations had used data warehouse, whereas many Higher Education institutions have not used it. Organizations need information that can be accessed anytime. Furthermore, organizations need accurate and valid data to support decision making process [8]. Competition is one important issue around Higher Education institutions. Not only local but also global Higher Education institution compete to get prospective students. Information technology is one of many answers to win the competition. Information technology and analytical tools are two tools that help Higher Education institutions in decision making process. Hence, decision making system might help Higher Education take ad-hoc decision, so institutions can

Journal of Theoretical and Applied Information Technology
30th June 2014. Vol. 64 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

adjust many changes easily. This system must take information from many sources, internal and external sources [13].

Problem Statements

- Information has not provided added value expected in Higher Education analysis process.
- Integration processes of all information are complex process and time consumed.
- Analysis system has not created useful knowledge in decision making process, strategic planning process, and monitoring process yet.

Research Objectives

General objectives of this research are to design data warehouse model, data mining and analytical tool that support analysis and decision making process. This model might be produced valuable knowledge to support decision making and strategic planning process. This research consist of 3 (three) steps, and the specific objectives of each step are :

1. Building the data warehouse model and using data mining technique based on the information model (previous research result)
2. Designing dashboards for analytical tool
3. Evaluating the model

Advantages of Research

Research based on the purpose could give many advantages for Higher Education institutions in Indonesia :

1. Get integrated information from all Higher Education institution activities to create integrated knowledge related to Higher Education institution performance
2. Be a model for supporting analysis process. This model help institution for decision making process and strategic planning process
3. Help Higher Education institution take action responsively based on environmental changes, internal and external environment

Disadvantages of Research

Some problems associated with developing a data warehouse are:

1. Underestimation of resource for data ETL
2. High demand for resources
3. Data ownership
4. Complexity of integration

2. LITERATURE STUDY

2.1 Data Warehouse

A data warehouse is a collection of integrated databases and subject-oriented and designed to support the decision-making function, where each unit of data is relevant to the events at any given time [9].

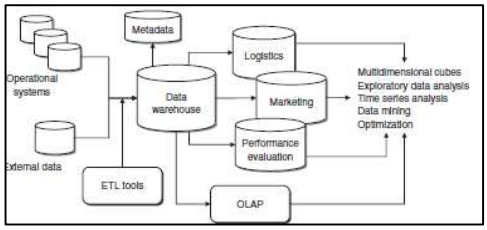

Figure 1: Architecture And Functions Of A Data Warehouse [15]

The ability of a data warehouse that can support the following actions [6]:

1. Running business data: data is produced by enterprise applications
2. Integrate business data: to improve and synchronize two or more enterprise applications, even those that do not designed to work alongside each other
3. Monitor business data: determine the relationship between data, deliver it to end users as reporting tools and support decision making process.

2.2 Data Warehouse Structure

Data warehouse has some level of detail that is older level of detail (usually found in alternative storage), current level of detail, lightly summarized the data (data mart level), and highly summarized data. The flow of data into the data warehouse from the operational environment. The transformation of data usually come from the operational level to the data warehouse level [9]

2.3 Dimensionality Modeling

A dimensional data model that has a fact table in the center, surrounded by denormalized dimension tables [2]. The data warehouse star schemes represent information about the structure and usage of the courses. These consist of two types of tables: Dimension tables which composed of descriptive data and Fact table which contain measurements. Data warehouse model for this Business Intelligence model consist of dimension and fact tables. Fact and dimension tables that exist in data warehouse design model are drawn in graphical form (star scheme form). The star scheme form enables users to analyze data from many dimensions or perspectives.

Journal of Theoretical and Applied Information Technology
30th June 2014. Vol. 64 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

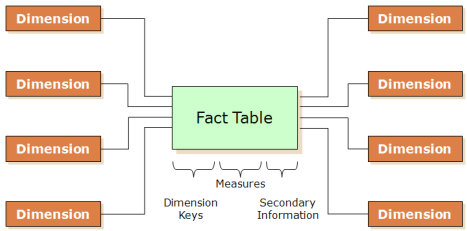

Figure 2:. Star scheme

2.4 Data Mining

Data mining (knowledge discovery from data) is Extraction of interesting (non-trivial, implicit, previously unknown and potentially useful) patterns or knowledge from huge amount of data [7]. Data mining can be defined in several ways, which differ primarily in their focus on different aspects of data mining. Data mining: The use of machine learning algorithms to find faint patterns of relationship between data elements in large, noisy, and messy data sets, which can lead to actions to increase benefit in some form (diagnosis, profit, detection) [11].

Association Rules

The goal of association rules is to detect relationships or associations between specific values of categorical variables in large data sets. This technique allows analysts and researchers to uncover hidden patterns in large data sets [11].

Association Rule, an implication expression of the form $X \rightarrow Y$, where X and Y are itemsets.

Rule Evaluation Metrics

- Support (s)
  Fraction of transactions that contain both X and Y
- Confidence (c)
  Measures how often items in Y appear in transactions that contain X

Support $(A \rightarrow B) = P(A \cup B)$

Confidence $(A \rightarrow B) = P(B|A) =$

$$\frac{\text{support}(A \cup B)}{\text{support}(A)} = \frac{\text{support\_count}(A \cup B)}{\text{support\_count}(A)}$$

2.5 Balanced Score Card

Balanced Scorecard was first built by Kaplan and Norton [10], which help decision-makers within the organization to drive the organization to be success. This technique enables organizations to translate mission and strategy into a comprehensive performance measurement package that provides the framework for a strategic measurement. Fundamental concept of the Balanced Scorecard is to reduce the size of the objectives scale and size that come from the vision and company strategies. Balance scorecard basic concept is 4 (four) prespectives as a balance frame work for monitoring process to achieve organization goals.

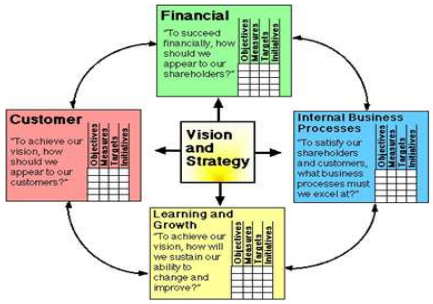

Figure 3: Balanced Scorecard: Mapping Strategy Into Operational Area. Source: [10]

Balanced Scorecard is a performance measurement strategy for organization. This strategy set specific target and size that link a strategy and mesurenet to strategy and important measurement in business [10]. All strategy extend to all unit and all level [12]. Balanceed scorecard manage financial measurement and others 3 (three) prespectives : Customer, Internal process and Learning and growth. So it act as "balanced meaasurement", qualitative and quantitative factor.

2.6 Dashboard

"Scorecards and dashboards are common components of most performance management system, performance measure systems, BPM suites, and BI platforms. Dashboard and scorecards both provide visual displays of important information that is consolidated and arranged on a single screen so that information can be digested a single glance and easily explored [14].

3. METHODS

This research based explorative descriptive research, without making comparisons or linking this research to any preceding research. This research using several techniques:

1. Interview top-level management as a source of information to obtain information about the organization's strategy, managerial plans, analysis activities and critical decision making process
2. Case studies, analysis, evaluation and study of the literature of the organization documents, also other information

Journal of Theoretical and Applied Information Technology
30th June 2014. Vol. 64 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

resources about implementation of analytical system in Higher Education institution.

The purposes of this study were:

1. Analyze the literature study result, internal and external data, as well as identify the finding
2. Analyze the information model on previous research's result
3. Designing the data warehouse model using SQL Server. The result of this research were: fact table model, dimension table model, and relationship between tables
4. Using data mining technique
5. Planning the technology used; and design dashboard as the user interface model of analytical system.
6. Evaluate the model using interview technique

4. DISCUSSION

4.1 Higher Education Information System Model

Currently, Many colleges have used information system in their academic operations. In 1990 Higher Education Directorate had designed The National Higher Education Information System (SINAS-DIKTI), see figure 4. This system consist of 10 (ten) subsystems.

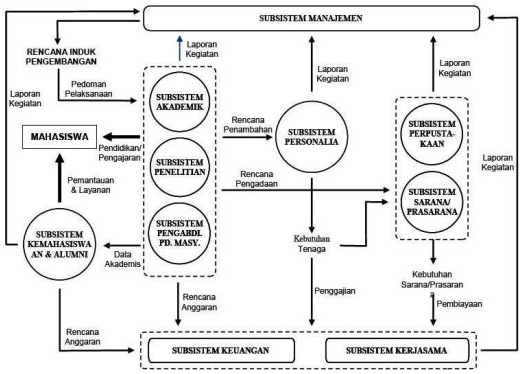

Figure 4: The National Higher Education Information System (SINAS-DIKTI) [1]

Higher Education subsystems

The subsystems consist of:

a. Academic subsystem
b. Research subsystem
c. Comunity Service subsystem
d. Personel subsystemI
e. Library subsystem
f. Infrastructure subsystem
g. Financial Subsystem
h. Coorporation subsystem
i. Student and Alumni subsystem
j. Management subsystem

4.2 Higher Education Star Schemes

Based on Higher Education Information System (HEIS), Standard Accreditation for HE and score card subsequently we identify all fact tables, dimension tables and design star scheme for developing the analytical tool. Here are fact tables, dimension tables and star scheme that are used to develop the analytical tool:

Star scheme for Curriculum Development

The main views of this scheme are: defining the number of courses for each Faculty and Study program by Year (time of curriculum developed).

Star scheme for Lecturer

The main view of this scheme is: defining the number of lecturers for each semester by Major, Lecturer status (part time or full time lecturer), Education level, Lecturer academic area, Lecturer certification status and Year (time of academic calendar).

Star scheme for Student Intake

The main views of this scheme are: formulate marketing activities for the program (activities and target results), evaluate and scoring/grading to determine which prospective student can continue to the next phase and re-registration for student that passing the test with particular grade to finish administration and financial obligation

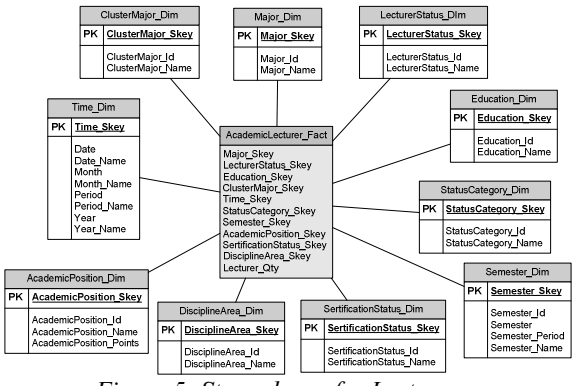

Figure 5: Star scheme for Lecturer

Journal of Theoretical and Applied Information Technology
30th June 2014. Vol. 64 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

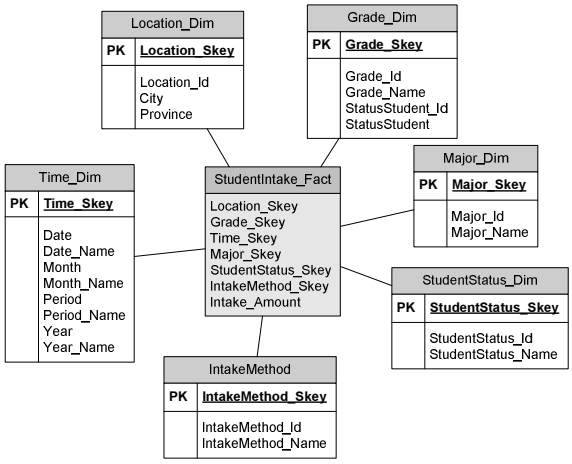

Figure 6: Star scheme for Student Intake

Star scheme for Registration and Payment

The main view of this scheme is: collect the number of registered active students by Semester, Major, Registration status and Time.

Star scheme for Teaching Learning process

The main views of this scheme are: the number of attendance for each student and lecturer by Semester, Study Program and Course.

Star scheme for Evaluation process

The main views of this scheme are: the number of students that pass the evaluation, and can be categorize by Course, Department, Major, Grade, Exam type and Semester.

Star Scheme for Thesis Guidance

The main view of this scheme is: the number of meetings for thesis guidance activities by Study Program, Semester and Lecturer.

Star Scheme for Student Coaching Activity

The main view of this scheme is: the number of students that are assisted by lecturer. This scheme is shown student couching activity by Study Program, Type of Coaching and Lecturer.

Star Scheme for Student's Achievement

The main view of this scheme is: the number of student's achievements by Study Program, Type of Activity and Scope of Activity.

Star scheme for Graduation (Alumni)

The main view of this scheme is: the number of graduation student by Major, and GPAType.
Furthermore this scheme view the number of graduation students who got a job before or just after graduation day. This number can be view by Work industry, Work position and Work location.

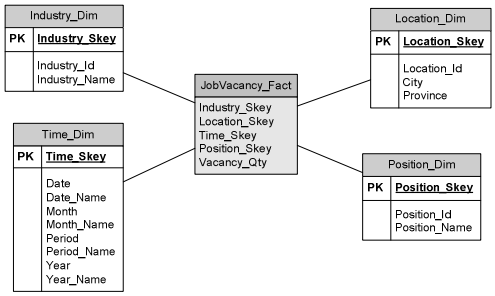

Figure 7: Star scheme for Job vacancy

Star scheme for Library Collection

The main view of this scheme is: the number of Library Collection by Type of Collection, Study Program, and Language.

Star Scheme for Publication

The main view of this scheme is: the number of Publication by Study Program, Lecturer and Type of Publication.

Star Scheme for Research

The main view of this scheme is: the number of Research activities by Research category, Major, Academic position and Time.

Stars Scheme for Grant

The main view of this scheme is: the number of Grant by Grant category, Major, Academic position and Time.

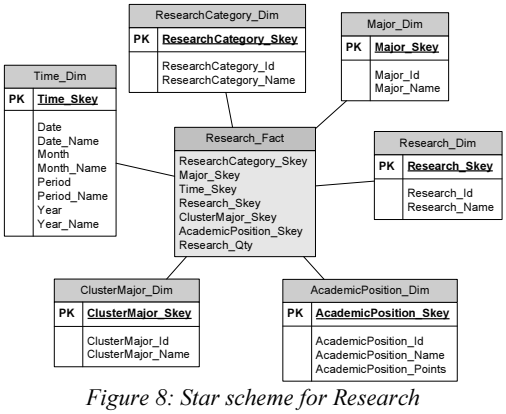

Figure 8: Star scheme for Research

Star Scheme for Community Services

The main view of this scheme is: the number of Community Service activities by Major, Community service category, Academic position and Time.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Star scheme for Infrastructures
The main views of this scheme are: the number of Facilities and Infrastructure by Study Program and Period.

Star Scheme for Industry and Government relationship
The main view of this scheme is: the number of relationship by Study Program, Period, Type of Industry and Scope of Industry.

Stars schemes for Information and Communication Technology
The main view of this scheme is: the number of IT devices by Study Progran, Period and Type of Device.

4.3 Higher Education Data Mining

For evaluation and data analysis purposes we use Association mining technique and data for testing. Here are some analysis results using Association mining technique:

Evaluation process
The rules for this process are:
1. TestType=mid exam 114290 ==> StudentStatus=pass the exam 114245 conf:(1)
2. Department=Language TestType=mid exam 40894 ==> StudentStatus=pass the exam 40872 conf:(1)
3. Department=Math 35551 ==> StudentStatus=pass the exam 35522 conf:(1)
4. Department=Language 90054 ==> StudentStatus=pass the exam 89973 conf:(1)
5. Department=Language ExamType-final exam 49160 ==> StudentStatus=pass the exam 49101 conf:(1)
6. ExamType=final exam 137459 ==> StudentStatus=pass the exam 137285 conf:(1)

Student's evaluation produce 6 (six) rules with Minimum Support : 0.1, Minimum Metric (confidence) : 0.9, Number of cycles performed : 18, NumRules : 1000.

Registration process
The rules for this process are:
1. Province=West Java 10418 ==> StudentStatus=pass the exam 10407 conf:(1)
2. Province=East Java 9683 ==> StudentStatus=pass the exam 9672 conf:(1)
3. Year=2009 23043 ==> RegistrationMode=SNMPTN 23043 conf:(1)
4. Year=2008 12579 ==> RegistrationMode =SNMPTN 12575 conf:(1)
5. Year=2009 RegistrationMode=SNMPTN 23043 ==> Grade=a 23035 conf:(1)
6. Year=2009 Grade=a 23035 ==> RegistrationMode=SNMPTN 23035 conf:(1)
7. Year=2012 15940 ==> RegistrationMode=UMB 15509 conf:(0.97)
8. Year=2012 15940 ==> RegistrationMode=UMB Grade=b 15506 conf:(0.97)

Student's registration produce 8 (eight) main rules with Minimum Support : 0.1, Minimum Metric (confidence) : 0.9, Number of cycles performed : 18, NumRules : 1000.

Hence all evaluation and analysis result present information and insight of data that can be used for Higher Education management help them to improve analytical and decision making process

4.4 Higher Education Dashboard

After determining each star scheme for the needs of information, the next process is to create a user interface in the form of the dashboard. This user interface is used by the Higher Education management. Information is presented in graphical form to facilitate the monitoring process. Here are some of the user interfaces that can be used by the Higher Education management.

Dashboard for Lecturer
Lecturer academic dashboard presenting information about the number of lecturer for each Major and Discipline area. For more detail this dashboard displays the number of lecturer for each Academic position, Education degree, Status and National certification status.

Journal of Theoretical and Applied Information Technology
30th June 2014. Vol. 64 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

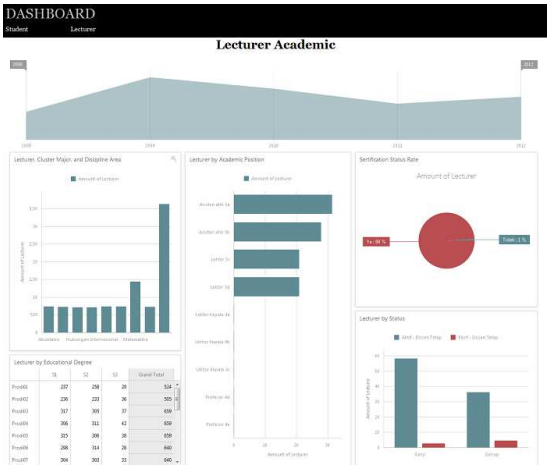

Figure 9: Dashboard for Lecturer

Dashboard for Student Intake and Payment

Student Intake dashboard presents information on the number of students who enroll at the university (the applicant) for each province and the number of students who passed the entrance examination (Applicant) for each department (faculty).

Payments Dashboard presents information on the number of students who have completed the registration and tuition fees.

Dashboard for Registration

Registration dashboard presenting information on the number of prospective students and their academic status (Active, Drop out, On-thesis process, Alumni, Leave, Resign and Without explanation).

Dashboard for Evaluation process

Evaluation dashboard displays information about student learning outcomes. Information is displayed in the comparison form of the number of passed and failed students. This information than can be viewed in detail by a comparison between the distribution of courses and grades.

Dashboard for Graduation (Alumni)

Graduation dashboard displays information about the number of graduates with the distribution of types of graduates. This dashboard can compare the number of graduate students from one study program to another study program; in addition, this dashboard can also view the status of graduates by employed /unemployed graduate students after graduation date. All information above can assist Higher education preparing new programs that are needed in the future.

Job vacancy dashboard displays information about the number of students who have worked before or after graduation date. For more detail the dashboard displays the information for each work industry, position and location.

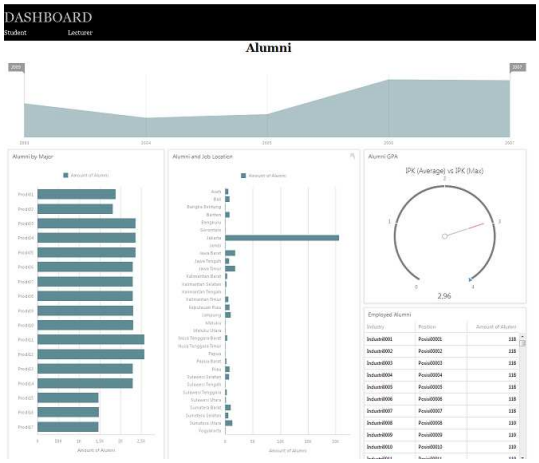

Figure 10: Dashboard for Alumni

Dashboard for Research

Research dashboard displays information about research activities which involving lecturers and students. For more detail this dashboard display the number of this activities by department (faculty), research category and lecturer grade (academic position).

Dashboard for Grant

Grant dashboard displays information about the number of grant achievement by department (faculty), grant category and lecturer grade (academic position).

Dashboard for Community Services

Community Services dashboard displays information about community services activities which involving lecturers and students. For more detail this dashboard display the number of this activities by department (faculty), community service category and lecturer grade (academic position).

5. CONCLUSION

Data warehouse models are built with the presentation of information in the form of graphs / dashboards to help Higher Education management get valuable information. Hence data warehouse and data mining extract information and discover valuable knowledge from huge amount of data. This information and knowledge can be used for operational and strategic interests, where today there are many Higher Education institutions have not been able to utilize and to take advantage from

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

operational data which generated by the current information system.

For future plan research can involve many public and private universities in Indonesia. Moreover, various mining techniques can be used for evaluation and data analysis purposes. Finally detail evaluation need to be done to make sure the implementation of data warehouse, data mining and dashboard meet Higher Education management needs.

REFRENCES:

[1] BAN –PT, "Self Evaluation Guidance", Jakarta, BAN-PT, 2010

[2] Connolly, Thomas M., Begg, Carolyn E., "Database systems : a practical approach to design, implementation, and management", Pearson Education, 2010

[3] Bhanti, P., Kaushal, U., Pandey, A. "E-Governance in Higher Education: Concept and Role of Data Warehousing Techniques", Intenational Journal of Computer Applications, Vol. 18 No.1, March 2011

[4] Bresfelean, V.P., Ghisoiu, N., "Higher Education Decision Making and Decision Support Systems", WSEAS Transaction on Advances in Engineering Education, Vol. 7, Issue 2, Feb 2010

[5] Chauhan, A.S., Singh, Y., Soam, A., "Effective Decision Making in Higher Educational Institutions Using Data Warehousing And Data Mining", Journal of Computer Science Engineering and Information Technology Research (JCSEITR), Vol. 2, Issue 1, Sep 2012

[6] Hammergren, T.C and Simon A,R., "Data Warehousing For Dummies®", 2nd Edition, Wiley Publishing, Inc., Indianapolis, Indiana, 2009

[7] Han, J., Kamber, M., & Pei, J., "Data Mining: Concepts and Techniques", Third Edition. Morgan Kaufmann, 2011

[8] Heise, David L. "Data Warehousing and Decision Making in Higher Education in United States", Andrew University, March 2005

[9] Inmon, W.H, "Building the Data warehouse . Fourth Edition".Wiley Publishing Inc.,Indianapolis. 2005

[10] Kaplan, Robert S. & David P. Norton, "Balanced Scorecard – Translating Strategy into Action", Harvard Business School Press, 1996

[11] Nisbet, Robert., Elder, John., Miner, Gary, "Handbook of statistical analysis and data mining applications", Elsevier Inc, 2009

[12] Niven, Paul R."Balanced Scorecard Step-by-step: Maximizing Performance and Maintaining Result" . John Wiley & Sons, Inc., New York. .2002

[13] T.N., Manjunath, S. Hegadi., Ravindra, "Design and Analysis of DWH and BI in Education Domain", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011

[14] Turban, E., Sharda, R., Delen, D., & King, D, "Business Intelligence a Managerial Approach ", Second Edition. New Jersey: Pearson, 2011

[15] Wilarso, Iik "The use of Data warehouse in Indonesian Higher Education" Information System Journal MTI-UI, Volume 4, No. 1, ISBNM 1412-8896, 2008

[16] Vercellis, C, "Business Intelligence: Data Mining and Optimization for Decision Making. Wiley, 2009