

Prototyping Data Warehouse System for Ministry of Higher Education in Saudi Arabia

Saleh Rehiel A Alenazi¹, Abd Ghani Gholamdin², Ayed Rheal A. Alanzi³ & Ebrahim Mohammed Al-Matari⁴

¹ School of Computing, College of Arts and Sciences (CAS), University Utara Malaysia, Malaysia

² College of Arts and Sciences (CAS), University Utara Malaysia, Malaysia

³ College of Science and Human Studies at Hotat Sudair, Mathematics Department, Majmaah University, Kingdom of Saudi Arabia

⁴ Faculty of Business and Economics, Amman University, Yemen and Othman Yeop Abdullah Graduate School of Business, University Utara Malaysia, Malaysia

Correspondence: School of Computing, College of Arts and Sciences (CAS), University Utara Malaysia, Malaysia. E-mail: sala207@hotmail.com

Received: August 8, 2014 Accepted: August 16, 2014 Online Published: September 3, 2014

doi:10.5539/cis.v7n4p74

URL: <http://dx.doi.org/10.5539/cis.v7n4p74>

Abstract

One of the major issues confronting the ministry of higher education in the kingdom of Saudi Arabia is the procedures for taking managerial decision making on their institutions of higher learning. Thus, there has not been empirical study on the integration of institutions of higher learning's repository to the ministry of higher education as the stakeholder. Since data warehouse is described as a system that capable of storing and exploring large amount of information from various data sources, this paper design a prototyping data warehouse for the ministry of higher institution of learning so as to facilitate searching and decision making on concerned institution. The XAMPP web server was used to design the data warehouse due to its embedded PHP and MySQL functions. Hence, the result presents a prototype that integrates the repositories of some higher learning institutions with the ministry of higher institution of learning in Saudi Arabia Kingdom.

Keywords: data warehouse, prototyping, Saudi Arabia, dimensional model

1. Introduction

Past and present studies have defined data warehouse as an oriented and integrated discipline which helps in collection of data for the purpose of helping management in decision making process (Yi et al., 2013; Mate et al., 2012; Inmon, 1996). Therefore, a data warehouse is premised on multipurpose data model that can provide systematic tools which will aid executives in the organization and usage of their data for strategic decisions making. Importantly therefore, many large corporate organizations have realized that data warehouses are highly important in today's fast growing and highly competitive world (Mate et al., 2012; Han and Kamber, 2006). Hence, the competition and pressures on the industry nowadays call for the intervention of data warehouse as the solution to engage and keep customers giving chances of knowing details about their needs (Han and Kamber, 2006).

In line with that, data warehouses systems have the power to help in exploring and storing of large amount of information from various sources of data (Glorio et al., 2010). Hence, these systems are relevant and meet the current information requirement scenarios. Besides that, the data is arranged by following the multipurpose model which assembles data into facts (measures of the analysis) and scope as the perspective of analysis (Glorio et al., 2010; Han and Kamber, 2006). Moreover, data warehouse have been employed in many domains, such as health, sport, education and many more for solving the problems of storing and exploring the challenges of using the pool of data for managerial decision's making (Nebot and Berlanga, 2010).

Additionally, one of the benefits of data warehouse is to provide repositories integrated information for querying and analysis (Baranovic, et al., 2003; Inmon, 1992). However, the primary purpose of data warehousing is to explore, sort and analyze basic idea behind the data warehousing for the purpose integrating information that are relevant in anticipation of enquiries. This makes a data warehouse to be considered as an active mechanism to information integration when compared to the traditional approaches that process and integrate when query

arrives (Saadatdoost et al., 2011). In relation to that, the use of mining data is necessary as strategic approach for government to make decisions through the use of a storage area of exact data across the value chain (Mohammed et al., 2012).

In the context of the ministry of higher education Jeddah Saudi Arabia which is one of the government's ministries is responsible for providing the needed policies and directions for the higher education of learning in Saudi Arabia (Hamdan, 2005). The functions of ministry of higher education in Saudi Arabia were initially under the ministry of education before their disintegration in 2003, bringing more values to the Saudi Arabia higher learning education (Hamdan, 2005). This provides more departments and agencies to the welfare of higher learning institutions with individual database. However, there has not been empirical study on the creation of data warehouse would allow ministry of higher education to access the repository of all the higher institutions in the kingdom of Saudi Arabia. Accessing the higher learning institutions' data warehouse by the ministry of higher education in the Saudi Arabia Kingdom would be of help to the ministry for the purpose of making decision at managerial level that will eventually lead to further development of Saudi Arabia Universities. Having the primary objective of developing prototype of the data warehouse for the higher institution of learning ministry in Saudi Arabia therefore, this paper would take all necessary steps that will help in achieving a system which is capable exploring, storing and validating data for the ministry.

2. Concept of Data Warehouse

The magnitude of data warehouse is an indication that it has capacity to give support to every enquiry that is being made in order to access high volume of records as well as executing other functions such as calculus, sorting, aggregation of data and so on (Santos and Bernardino, 2008). Therefore, data warehouse has maximum capacity to access predefined and structured data like partition or indexes which may be complex and large in size. This means that the data warehouse update will be very challenging due to useful time and server overloading and because of the frequency and volume of the update. This actually entails large amount of operations to be carried out on the data warehouse structures and which can dramatically reduce the efficiency of OLAP performance (Bruckner and Tjoa, 2002; Santos and Bernardino, 2008).

Data Warehouse systems are instruments of decisions making being used to analyze the position and development of an organization premised on the large quantity of data that has been harnessed from different sources into the multipurpose model which include, facts, dimension and measures (Kimball and Ross, 2002; Soler et al., 2008). As a result of the importance of information that is stored in the data warehouse, protecting the data has become an important requirement that must be given consideration at all levels of development beginning from requirements to execution and maintenance (Devanbu and Stubblebine, 2002; Rizzi et al., 2006).

In addition, one of the significant security requirements of any management information system is to ensure that all data and resources of organization are protected from hacking and unacceptable modification as well as ensure confidentiality of information while ensuring that right users have access to required information. Thus, the classical security model should be replaced with an appropriate model centered on the main concepts of multidimensional and conceptual modeling like facts, dimensions and measures (Rizzi et al., 2006; Soler et al., 2008).

Additionally, data warehouse ensures that necessary information that are required for decision making, critical processing and data mining tools are made available. In line with that, collects data from different operational source systems like On-Line Transaction Processing (OLTP) and On-Line Analytical Processing (OLAP) in a central store used by analytical applications (Rizzi et al., 2006; Soler et al., 2008). The purpose of data area of a data warehouse is to store all the transaction history of an organization premised on the regular process that is often used to obtain information through OLAP tools (Santos and Bernardino, 2008).

3. Data Warehouse Model Design

There is a systematic lifecycle for the development and application of data warehouse which can help to integrate data into the warehouse for the purpose of analysis, implementation design and maintenance in 3 stages which is logical model, physical model and conceptual model (Fasel and Shahzad, 2010). The fact that these stages are completed does not imply that the data warehouse development and application stages have been terminated but the stages mark the transformation into another higher stage. On the other hand, the development and application process of data warehouse represent normal 3 stages and it has been transformed into the design of the model (Fasel and Shahzad, 2010). Hence, original data warehouse is recommended for the users.

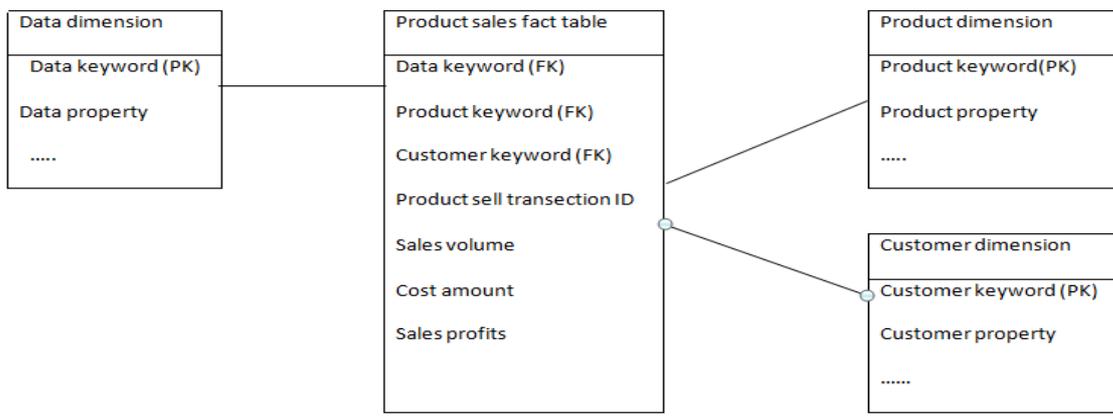


Figure 1. Data warehouse model

Furthermore, the present data warehouse system can be improved through continuous and systematic approach (Wang, 2009). The figure 1 explicitly depicts a sample of the warehouse model design. As shown, in this situation, the system should complete the product sales trend analysis through the development and application. The model as shown will bring about efficiency, convenience, and will also help in making enquiries about the historical performance of sales of products in different department for the purpose of aiding managers’ decision making. The design of data warehouse is done through systematic renovation steps of conceptual model, logical model and physical model (Wang, 2009; Fasel and Shahzad, 2010). This has eventually helped to transform the traditional data base into a refined one. Going by this, the application of traditional database has improved into data warehouse in an organized way. Thus, the prime task of conceptual model design is to define the boundaries of the system and determine the subject domain and its contents.

4. Data Warehouse Architecture

The direct integration of external data into the warehouse undoubtedly bring about external process of operational data base resources which also influences the efficiency of the data warehousing loading data capacity (Yebai and Li, 2009). In line with this view, have revealed that data warehouse design often permits repository area for data pre-processing with the purpose of improving total performance of the data warehouse (Yebai and Li, 2009; Wang, 2009; Fasel and Shahzad, 2010). Meanwhile, the data warehouse design as shown in Figure 2 stands for the analysis of inventory and decision support system architecture.

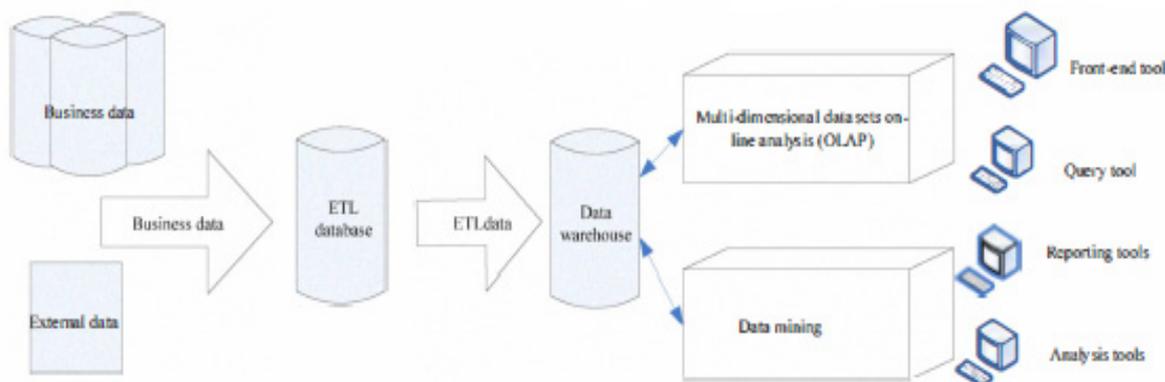


Figure 2. Data warehouse architecture

While pursuing this further, all the data sources that are stored in organization's ERP system database are mined to the extract, transform and load (ETL) database has shown in Figure 2 as this represents the data groundwork area based on the theme and ETL tool. Certain factors such as conversion and cleaning of the ETL data base, loading into the warehouse through ETL tool will determine the completion of the data. Moreover, through the

use of Microsoft analysis such as OLAP server, all necessary data meant to be used can be stored in the warehouse by the dimensions and fact tables. The analysis Services can easily define dimensions and build OLAP cube, front-end query tools, reporting tools, analysis tools, access decision support systems and analyze through data browsing.

5. Data Warehouse in Education

Currently, data warehousing has become aware about the profits for public administration. It has also boosted the interaction of data analysts and reporters (Inmon, 1996). Moreover, it also provides organized structure for fixed and variable data. It helps in the automatic conversion of any increase in fixed and variable data for control purpose. Data structure also has the capacity to modify any pertinent data base. It is essential that the structure has multipurpose pattern of processing for all the stakeholders. This will help all the universities in the world to be able to work independently.

The automation of the data process can be achieved and enhanced through the help of e-governance application. Necessary information that is required for approval can be obtained from data warehouse. In addition, through e-Government and through the usage of centralized data base, avalanche of opportunities are available for the universities. It will also be possible for all the universities' departments to provide good platform for companies and best performing government that are seeking interaction between employees and researchers. It therefore becomes easy for all the departments in the universities to get relevant information about the fields which every university specializes using data warehouse. (Bhanti et al., 2011).

6. Data Warehouse Model in Higher Learning

The higher institution of learning data warehouse possesses attributes of accessibility, transparency, efficiency and impact as shown in Figure 3.

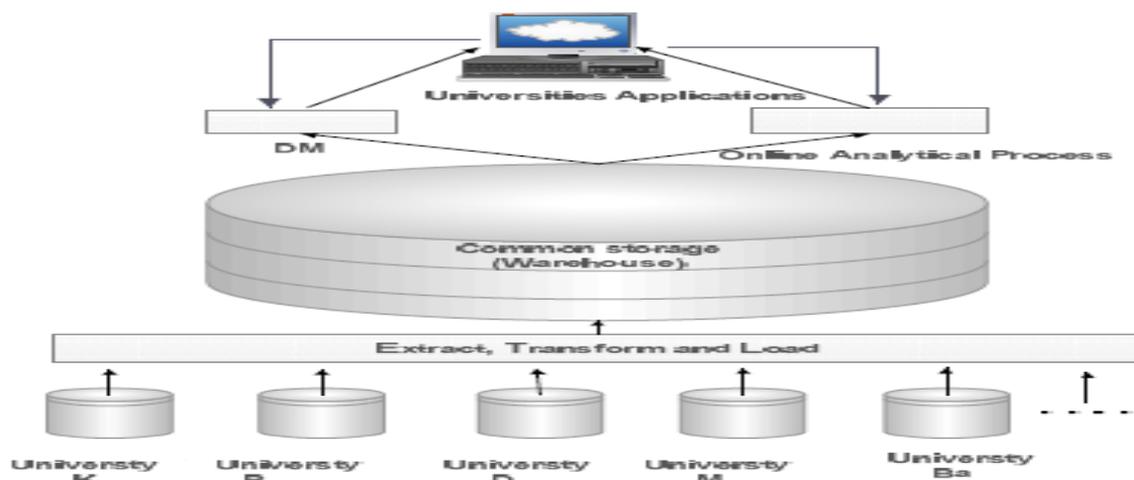


Figure 3. Data warehouse model of higher institution monitored by the ministry

Based on the figure3; the explanation of Institution Higher learning data ware house proses follow.

6.1 Databases Stage

The function of the data base is to ensure that all necessary documents, reports and all other information are gathered from data base of the universities such as K, B, d, M and Ba.

6.2 The ETL Stage

This stage includes tool called ETL (Extract, Transform and Load). The function of ETL is to mine information from the data bases of universities and transform such information with the purpose of meeting the needs of the users especially after it has loaded the mined information into the warehouse

6.3 The Warehousing Stage

The storage system called warehouse that is used to Store University's information depend on type and time factors (current information and historical information). Therefore, the normal warehouse has the ability that can

help to improve the process of sharing information among the departments of the universities. It also helps to improve interaction among the stakeholders and other agencies. It further helps to easily respond to the enquiries of staff from different universities since the store is carrying large amount of information (Demigha, 2009).

6.4 The Tools Stage

Online analytical processing (OLAP) it makes analytical data. The function of the analytical data is to make available necessary and valuable information and reports to the all the departments in the university. Moreover, it also gives assistance by viewing the dimension from different views with the purpose of analyzing the data. In other words, the essence of data mining is to extract necessary data from the warehouse and make such available to the universities (Demigha, 2009; Bhanti et al., 2011).

6.5 The Interface Stage

The interface is prepared for the government agency, like ministry of higher education to perform analyses and mining information. The agencies in universities can as well accesses warehouse the data warehouse with ease and obtain the needed information in the data warehouse (Mohammed et al., 2012).

7. Methodology

Previous studies have shown that methodology is the process that can help in articulating and executing all activities that are necessary for the completion of the research (Hoffer et al., 2005). Academic scholars have therefore proved that research methodology is not just about gathering methods for the research but it also involve looking for better ways of resolving existing problem. Therefore, this research shall adopt Vaishnavi and Kuechler (2004) approach in order to achieve the objectives of this study. This approach has five major stages which include awareness of problem, suggestion, development, evaluation and conclusion.

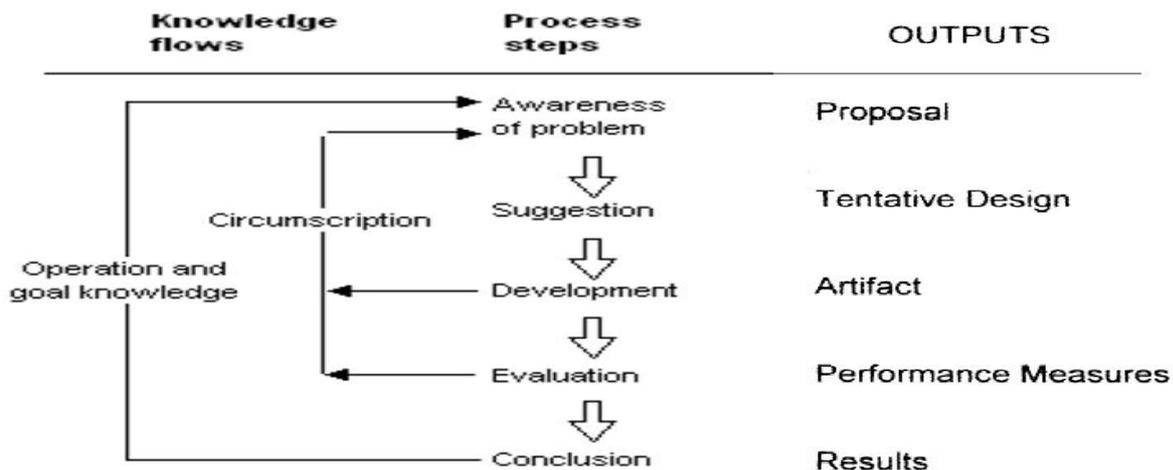


Figure 4. Research design methodology

Source: Vaishnavi and Kuechler (2004).

8. Awareness of the Problem

Previous researchers have emphasized that a good research starts from the understanding of the research problem which digs dip into the awareness of the problems to be addressed (Bezzi et al., 2010; Uljens, 2004). In other words, when there is awareness about the problems at hand, understanding about research idea can be enhanced as it will also help in finding long term solution to the problem Therefore, the problem facing the ministry of higher education in terms of searching for specific information of individual higher learning institution in Saudi Arabia was digested extensively. The dimensions of the higher learning institutions were understood from the angels of ministry of higher education to the higher learning institutions in Saudi Arabia for proper justification of the this research.

9. Suggestion

The suggestion to solve the existing problem faced by the ministry of higher education in Saudi Arabia through the prototyping of data warehouse would allow transparency of higher learning institution in Saudi Arabia. The study of Demarest (2013) suggested the prototyping of data warehouse in order to avoid risk of tampering with

data since ministry of higher education deals with public data. In this respect, the prototype must therefore be designed in such a way that it will be user friendly and will be able to assist the intending users to achieve their objectives. When the interface is interactive, it will also help the users to have more knowledge about the usage of data warehouse systems (Sridara and Iyakutti, 2009).

10. Development

The prototyping development stage for the data warehouse for the higher Institution of learning ministry in Saudi Arabia was made possible through critical analysis of requirements, star schema, and multipurpose model.

11. Dimensional Model

Data warehouses are built using dimensional data models which consist of fact and dimension tables (Santos and Bernardino, 2008). Therefore, IBM (2006) argued that to overcome performance issues for large queries in the data warehouse there should be use of dimensional models. The dimensional modeling approach provides a way to improve query performance for summary reports without affecting data integrity. The study of Sing and Kumar (2012) emphasised that a dimensional model is sometimes called a star schema which is popular in data warehousing as a result of its ability to provide much better query performance, especially on very large queries, than an E/R model. However, it also has the major benefit of being easier to understand. It consists, typically, of a large table of facts.

12. Prototype Design

The design of data warehouse prototype was made possible for higher Institution of learning in Saudi Arabia with the help of XAMPP web server application package. Previous studies have expressed that XAMP provides avenue to create the dimensional database for all the tables that would represent the actors in the data warehouse. Therefore, the services provided by MySQL function of XAMPP help to smooth contact between the system and system users who may be making enquiries. Importantly, the developers are attracted to the environment because it is language-neutral oriented and it can also help the users to easily achieve their daily objectives through effective content delivery without necessarily consider the type of platform that is involved. in use (Road, 2002). Hence, the data warehouse prototype for ministry of education would allow different stakeholders in the ministry of education in searching for the details of a respective university.

13. Evaluation

System evaluation is a significant step in the development in the development of information system which helps in the improvement and assessment of the quality of prototype (Ahmad et al., 2013). However, among the limitations of this system is that waste time and it is very expensive to maintain because of the accessibility of the system which may vary depending on the method being adopted by individual system evaluator (Saad and Salim, 2008). In addition, when system is evaluated, a feedback is given and this helps to guaranty the successful implementation of the system especially when it comes to searching for related issues and information in Saudi Arabia Universities. Hence, the research was evaluated by testing the ability of prototype to display the repositories of institutions of higher learning in the kingdom of Saudi Arabia.

14. Conclusion

Prototyping of data warehouse for the ministry of higher education in Saudi Arabia would provide a lasting solution to the lack of integration of information about the universities and ministry of higher education. Moreover, the practical searching of detailed information for managerial decision making at the ministry of higher education was achieved with ease by presenting prototype in the subsequent chapter. Hence, the ease of use and relevant of the data warehouse system were confirmed in chapter four as well.

15. Recommendation and Future Researchers

The importance of the developing prototype data warehouse for taking managerial decision making in the kingdom of Saudi Arabia on the higher learning institutions cannot be overemphasized in achieving effective educational management. Therefore, this calls for immediate recommendation of this research in the ministry of higher education in Saudi Arabia. Hence, this research helps approach of managerial decision making on the institutions of higher learning in Saudi Arabia and helps in forecasting the future status of the higher learning institutions.

16. Conclusion of the Study

As the main objective of this research is to design and validate a data warehouse prototype which allows ministry of higher education in Saudi Arabia to access repository of higher learning institution, the research has revealed that the only panacea to have a formidable institution of higher learning is to provide access to the

ministry. This chapter has established the achievement of objectives as stated in chapter one of the study. Lastly, the limitations of the study, directions of future research and recommendations of the research have been discussed.

References

- Ahmad, O., Hyder, I., Iqbal, R., Murad, M. A. A., Mustapha, A., Sharef, N. M., & Mansoor, M. (2013). A Survey of Searching and Information Extraction on a Classical Text Using Ontology-based semantics modeling: *A Case of Quran. Life Science Journal*, 10(4).
- Baranovic, M., Madunic, M., & mekterovic, I. (2003). Data Warehouse as a Part of the Higher Education Information System in Croatia. 25th Int. Conf. Informafion Technology Interfaces /TI, cavtat Croatia, pp. 121-126.
- Bezzi, M., Duquenoy, P., Hubner, S. F., Hansen, M., & Zhang, G. (2010). Privacy and Identity Management for Life. IFIP International Federation for Information Processing, Germany. <http://dx.doi.org/10.1007/978-3-642-14282-6>
- Bhanti, et al. (2011). E-Governance in Higher Education: Concept and Role of Data Warehousing Techniques. *International Journal of Computer Applications*, 18, 23-41.
- Bruckner, R. M., & Tjoa, A. M. (2002). Capturing Delays and Valid Times in Data Warehouses – Towards Timely Consistent Analyses. *Journal of Intelligent Inf. Systems (JIIS)*, 19(2), 169-190. <http://dx.doi.org/10.1023/A:1016555410197>
- Demarest, M. (2013). *Data warehouse prototyping: Reducing risk, securing commitment and improving project governance* (pp. 1-14). Wherescape whitepaper.
- Demigha, S. (2009). A Data Warehouse System to Help Assist Breast Cancer Screening in Diagnosis, Education and Research (pp. 1-6). IEEE Explore.
- Devanbu, P., & Stubblebine S. (2002). Software engineering for security: A roadmap. In A. Finkelstein, (Ed.), *The Future of Software Engineering* (pp. 227-239), ACM Press, New York.
- Fasel, D., & Shahzad, K. (2010). *A DataWarehouse Model for Integrating Fuzzy Concepts in Meta Table Structures* (pp. 100-109). IEEE computer society.
- Glorio, O., Mazon, J. N., Garrigoz, I., & Trujillo, J. (2010). *Using Web-based Personalization on Spatial Data Warehouses* (pp. 1-8). EDBT-ACM.
- Hamdan, A. (2005). Women and Education in Saudi Arabia: Challenges and Achievements. *International Education Journal*, 6(1), 42-64.
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques* (2nd ed.). Morgan Kaufmann Publishers.
- Hijazi, S. (2013). *Using Portable Technology to Teach Web Programming and Database Classes* (pp. 34-50). Proceedings of ASCUE.
- Inmon, W. H. (1992). *Building the Data Bridge: The Ten Critical Success Factors of Building a Data Warehouse*. Database Programming & Design.
- Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit*. Wiley & Sons.
- Mate, A., Trujillo, J., Gregorio, E. D., & Song, Y. (2012). *Improving the Maintainability of Data Warehouse Designs: Modeling Relationships between Sources and User Concepts* (pp. 25-32). DOLAP-ACM.
- Mohammed, M. A., Hasson, A. R., Shawkat, A. R., & Al-Khafaji, N. J. (2012). *E-Government Architecture Uses Data Warehouse Techniques to Increase Information Sharing in Iraqi Universities* (pp. 1-5). IEEE Explore.
- Nebot, V., & Berlanga, R. (2010). *Building Data Warehouses with Semantic Data* (pp. 1-8). EDBT-ACM.
- Rizzi, S., Abell'o, A., Lechtenb'orger, J., & Trujillo, J. (2006). Research in data warehouse modeling and design: dead or alive? (pp. 3–10). In: DOLAP.
- Road, B. (2002). *Authentication and Security Mechanisms in ASP*. Retrieved December 22, 2010, from http://documents.iss.net/whitepapers/asp_net_whitepaper.pdf
- Saad, S., & Salim, N. (2008). Methodology of Ontology Extraction for Islamic Knowledge Text. Paper presented at the Postgraduate Annual Research Seminar.
- Saadatdoost, R., Sim, A. T. H., & Jafarkarimi, H. (2011). Application of Self Organizing Map for Knowledge Discovery Based in Higher Education Data (pp. 1-6). IEEE Explore.

- Santos, R. J., & Bernardino, J. (2008). Real-Time Data Warehouse Loading Methodology (pp. 49-58). IDEAS-ACM.
- Soler, E., Trujillo, J., & Medina, F. (2008). Towards comprehensive requirement analysis for DW: Considering security requirement Published in IEEE Conference in Web Applications. *Journal of Object Technology*, 8(2).
- Sridaran R., Padmavathi, G., & Iyakutti, R. (2009). A Survey of Design Pattern Based Technologies. Idea Publishing Group.
- Vaishnavi, V., & Kuechler, B. (2004). Design Research in Information Technology.
- Wang, F. (2009). Application Research of Data Warehouse and Its Model Design (pp. 798-801). IEEE computer society.
- Yebai, L., & Li, Z. (2009). *Interactive Modeling of Data Warehouse on E-business System* (pp. 80-83). IEEE computer society.
- Yi, X., Paulet, R., Bertino, E., & Xu, G. (2013). *Private Data Warehouse Queries* (pp. 25-35). SACMAT- ACM.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).