REQUIREMENT MODELING FOR DATA WAREHOUSE USING GOAL-UML APPROACH: THE CASE OF HEALTH CARE

SAJARATUL HUDA BT MOHD SHAFIE

MASTER OF SCIENCE (INFORMATION TECHNOLOGY)
UNIVERSITI UTARA MALAYSIA
2017

REQUIREMENT MODELING FOR DATA WAREHOUSE USING GOAL-UML APPROACH: THE CASE OF HEALTH CARE

A Thesis submitted to Dean of Awang Had Salleh Graduate School of Arts and Sciences in Partial Fulfilment of the requirement for degree Master of Science in Information Technology

Universiti Utara Malaysia

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Pembuat keputusan menggunakan Gudang Data (DW) untuk melaksanakan analisis pada maklumat perniagaan. Pembangunan DW adalah satu proses yang panjang dengan risiko kegagalan yang tinggi dan sukar untuk menganggarkan keperluan untuk membuat keputusan pada masa depan. Tambahan, reka bentuk DW semasa tidak mengambil kira analisis keperluan awal dan akhir semasa pembangunannya khususnya dengan menggunakan pendekatan *Unified Modeling Language* (UML). Berdasarkan masalah ini, adalah penting untuk pendekatan pemodelan DW semasa merangkumi kedua-dua analisis keperluan awal dan akhir dalam reka bentuk DW. Satu kajian kes telah dijalankan ke atas Penjagaan Kesihatan Luar Bandar, Malaysia (MRH) untuk mengumpul keperluan reka bentuk DW. Pendekatan berorientasikan matlamat telah digunakan untuk menganalisis keperluan awal dan kemudian dipetakan kepada pendekatan UML untuk menghasilkan model DW baharu yang dipanggil Goal-UML (G-UML). Pendekatan yang disyorkan menekankan proses pemetaan skema konseptual DW kepada gambar rajah kelas untuk menghasilkan reka bentuk MRH-DW yang lengkap. Ketepatan reka bentuk DW itu dinilai melalui ulasan pakar. Kaedah G-UML boleh menyumbang kepada pembangunan DW dan menjadi garis panduan kepada pembangun DW untuk menghasilkan reka bentuk DW yang baik serta memenuhi semua keperluan pengguna.

Kata kunci: gudang data, orientasi-matlamat, skema konseptual, kelas rajah, keperluan

Abstract

Decision makers use Data Warehouse (DW) for performing analysis on business information. DW development is a long term process with high risk of failure and it is difficult to estimate the future requirements for the decision-making. Further, the current DW design does not consider the early and late requirements analysis during its development, especially by using Unified Modeling Language (UML) approach. Due to this problem, it is crucial that current DW modeling approaches covered both early and late requirements analysis in the DW design. A case study was conducted on Malaysia Rural Health Care (MRH) to gather the requirements for DW design. The goal-oriented approach has been used to analyze the early requirements and later was mapped to UML approach to produce a new DW modeling called Goal-UML (G-UML). The proposed approach highlighted the mapping process of DW conceptual schema to a class diagram to produce a complete MRH-DW design. The correctness of the DW design was evaluated using expert reviews. The G-UML method can contribute to the development of DW and be a guideline to the DW developers to produce an improved DW design that meets all the user requirements.

Keywords: data warehouse, goal-oriented, conceptual schema, class diagram, requirement

Acknowledgement

Alhamdulillah. First and foremost, all praise and thanks to Allah for giving me the strength and patience, and providing me the knowledge to accomplish this thesis. And, special dedication to my beloved father, mother and family members.

This thesis would not have been possible without the support of many people. First, I wish to express my gratitude to my supervisors, Dr. Azman Ta'a and Dr. Azizah Ahmad, who abundantly helpful and offered invaluable assistance support and guidance. My sincere thanks must also go to the members of my Viva committee: Dr. Norliza Katuk, Assoc. Prof. Dr. Haslina Mohd, and Dr. Sharhida Zawani Saad, for useful comments and suggestions to improve my thesis. Deepest gratitude to Dr. Norliza Katuk and Mdm Nor Ma and all friends those without their assistance, this study would not have been successful.

I would also like to thank all my brothers and sister for their support.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

DW	Data Warehouse
MOH	Ministry of Health
NGO	non-government organization
MCH	Mother Child Health Care
OPD	Outpatient Department
ER	Entity Relationship
UML	Unified Modeling Language
NFR	Nonfunctional requirement
G-UML	Goal and UML
RE	Requirement Engineering
BO	Business Object
GOODM	Graph Object Oriented Multidimensional Data
AGDI	Agent-Goal-Decision-Information
OLAP	Online Analytical Processing
CRUD	Create, Read, Update, Delete
ETL	Extraction-Transformation-Loading
GRAnD	Goal-oriented Requirement Analysis for Data Warehouse
SD	Strategic Dependency
SR	Strategic Rationale
GOMA	Goal-oriented Modeling Approach
AGORA	Annotated Goal-Oriented Requirements Analysis
GRL	Goal Requirement Language
GBRAM	Goal-Based Requirements Analysis Method
KAOS	Knowledge Acquisition in automated specification
GORE	Goal-oriented Requirement
MDA	Model Driven Architecture
CIM	Computation Independent Model
PIM	Platform Independent Model
3NF	Third Normal Form
UP	Unified Process
OCL	Object Constraint Language
ETL	Extract, Transform and Load
PSM	Platform Specific Models
KVB	KassenärztlicheVereinigung Bayern, Bavarian Association of Statutory Health Insurance Physicians
HITV	Health Insurance Treatment Voucher
DFM	Dimensional Fact Model
FK	Foreign Key
OME	Organization Modeling Environment
MEAA	Maybank Electronic Application Accommodation
EOS	Enterprise Origination System
MRH	Malaysia Rural Health Care

CHAPTER ONE
INTRODUCTION

1.1 Introduction

This chapter serves as the introductory part of this study. It includes the objectives of the study, the background of study, problem statement, scope and limitation of the research and research significance that need to be gained in this study. In summary, this study is laid for further discussion on how goal-oriented and UML modeling approach contribute to develop DW requirement modeling based on Rural Health Care.

1.2 Background of the Study

Data Warehouse (DW) is decision support systems that are specifically derived for the business environment. It's used mainly by decision maker in organization to improve decision making system and increase organization performance. DW contain multiple databases that stores and organizes enterprise-wide data based on large amount of data integrated from heterogeneous sources (Sharma & Jain, 2013). It helps to enhance data access for analysis and decision making that can be used to deduce useful information in systematic way.

As in other information systems, requirements analysis phase is one of important phase that might influence all the phases in DW development. Requirement analysis phase help to identify accurate end users that represent requirements in different way and to reduce risk of DW failure. The main objective of requirement analysis phase

The contents of the thesis is for internal user only

REFERENCES

Afreen, N., Khatoon, A., & Sadiq, M. (2016). A Taxonomy of Software's Non-functional Requirements. In *Proceedings of the Second International Conference on Computer and Communication Technologies* (pp. 47-53). Springer India.

Ado, A., Aliyu, A., Bello, S. A., & Garba, A. U.(2014). Building a Diabetes Data Warehouse to Support Decision making in healthcare industry. *IOSR Journal of Computer Engineering (IOSR-JCE). 16*(2), Ver. IX (Mar-Apr. 2014), pp 138-143.

Abai, N. H. Z., Yahaya, J. H., & Deraman, A. (2013). User Requirement Analysis In Data Warehouse Design: A Review. *Procedia Technology, 11,* 801-806.

Axel van Lamsweerde (2009). *Requirements Engineering - From System Goals to UML Models to Software Specifications*. England: Wiley.

Axel van Lamsweerde. (2001, August).Goal-Oriented Requirements Engineering: A Guided Tour. *5th IEEE International Symposium on Requirement Engineering (*RE 2001): 49-263

Alencar, F. M., Castro, J., Cysneiros Filho, G. A., & Mylopoulos, J. (2000, July). From Early Requirements Modeled by the i* Technique to Later Requirements Modeled in Precise UML. In *WER* (pp. 92-108).

Annie I. Antón: Goal-Based Requirements Analysis. ICRE 1996: 136-144.

Bennet S., McRobb S. & Farmer R., (2010). *Object-Oriented Systems Analysis and Design Using UML* (4th Edition). UK: McGraw-Hill Higher Education.

Bresciani, P., Perini, A., Giorgini, P., Giunchiglia, F., & Mylopoulos, J. (2002). Tropos: An agent-oriented software development methodology. *Autonomous Agents and Multi-Agent Systems*, *8*(3), 203-236.

Ballad C., Herreman D., Schau D., Bel R., K. Eunsang, Valencic A. (1998). *Data Modeling Techniques for Data Warehouse.* Retrieved from http://www.redbooks.ibm.com

Berson A., Stephen J. Smith (1997). *Data Warehousing, Data Mining and OLAP* (Mcgraw-Hill)

Chhabra, R., Kumar, P., & Pahwa, P. (2016). An approach to Design Object Oriented Data Warehouse. *International Journal of Research and Engineering*, 3(3), 54-56.

Cravero A. & S. Sepúlveda (2014). Multi Dimensional Design Paradigms for Data Warehouse: A Systematic Mapping Study. *Journal of Software Engineering and Application*, *7*, pp. 53-61.

Chung, L., & do Prado Leite, J. C. S. (2009). On non-functional requirements in software engineering. In *Conceptual modeling: Foundations and applications*(pp. 363-379). Springer Berlin Heidelberg.

Cysneiros, L. M., Werneck, V. M. B., Amaral, J., & Yu, E. (2005). Agent/Goal Orientation versus Object Orientation for Requirements Engineering: A Practical Evaluation Using an Exemplar. In *Proc. of VIII Workshop in Requirements Engineering* (pp. 123-134).

Duy Cu Nguyen, Anna Perini, Paolo Tonella (2007). A Goal-oriented Software Testing Methodology. In *8th International Workshop on Agent-Oriented Software Engineering*, AAMAS

Eridaputra, H., Hendradjaya, B., & Sunindyo, W. D. (2014). Modeling the requirements for big data application using goal oriented approach. In Data and Software Engineering (ICODSE), 2014 International Conference on (pp. 1-6). IEEE.

Ellis-Braithwaite, R., Lock, R., Dawson, R., & Haque, B. (2013). Towards an approach for analysing the strategic alignment of software requirements using quantified goal graphs. *arXiv preprint arXiv:1307.2580*.

Farhan, M. S., Marie, M. E., El-Fangary, L. M., & Helmy, Y. K. (2011). An Integrated Conceptual Model for Temporal Data Warehouse Security. *Computer and Information Science*, *4*(4), 46.

Gupta, V., Chauhan, A., Kumar, A., & Taneja, S. (2011). UREM-A UML-Based Requirement Engineering Model for a Data Warehouse. In *Proceedings of the 5th National Conference*.

Guerra, J., McGinnis, J., & Andrews, D. (2011). Why you need a data warehouse. *Andrews Consulting Group, www. rapiddecision. net/pdfs/Why-You-Need-a-Data-Warehouse. pdf*.

Golfarelli, M. (2010). From User Requirements to Conceptual Design in Data Warehouse Design.

Golfarelli, M., & Rizzi, S. (2009). A comprehensive approach to data warehouse testing. In *Proceedings of the ACM twelfth international*

workshop on Data warehousing and OLAP (pp. 17-24). ACM.

Goldsby, H., & Cheng, B. H. (2006). Goal-oriented modeling of requirements engineering for dynamically adaptive system. In *14th IEEE International Requirements Engineering Conference (RE'06)* (pp. 345-346). IEEE.

Giorgini, P., Kolp, M., Mylopoulos, J., & Pistore, M. (2003). The tropos methodology: An overview. *Methodologies And Software Engineering For Agent Systems, Kluwer Academic Publishing (New York)*.

Garzetti, M., Giorgini, P., Mylopoulos, J., & Sannicolo, F. (2002). Applying Tropos Methodology to a real case study: Complexity and Criticality Analysis.

Golfarelli, M., Maio, D., & Rizzi, S. (1998). The dimensional fact model: a conceptual model for data warehouses. *International Journal of Cooperative Information Systems*, *7*(02n03), 215-247.

Horkoff J. & E. Yu. (2012). Comparison and evaluation of goal-oriented satisfaction analysis techniques. doi: 10.1007/s00766-011-0143-y

Horkoff, J. M. (2012). *Iterative, interactive analysis of agent-goal models for early requirements engineering.* (Doctoral dissertation) University of Toronto.

Hugh J.W, Dale L.G, & Barbara H.W (2002). The Benefit Of Data Warehousing: Why Some Organizations Realize Exceptional Payoffs. *Information & Management, 39* (2002), 491-502.

Hüsemann, B., Lechtenbörger, J., & Vossen, G. (2000). *Conceptual data warehouse design* (pp. 6-1). Univ.

Inuwa, I., Oye. D., N., (2015) Design of a Data Warehouse Model for a University

Decision Support System. *Information and Knowledge Management*, Vol.5, No.12, 2015

Inmon, W. H. (2002). *Building the Data Warehouse*. John Wiley & Sons (3rd ed.). Retrieved from http://inmoncif.com/

Jiang, L. (2010). *Data Quality By Design: Agoal-Oriented Approach* (Doctoral dissertation, University of Toronto).

Kumar, M., & Singh, Y. (2010). Stakeholders driven requirements engineering approach for data warehouse development. *Journal of information processing systems*, *6*(3), 385-402.

Kamal Alaskar and Akhtar Shaikh, (2009). Object Oriented Data Modelling For Data Warehousing: And Extension Of UML Approach To Study Hajj Pilgrim's Private Tour As Case Study. *International Arab Journal of e-Technology*, *1*(2).

Kimball (2002). *Data Warehouse Toolkit*. *The Complete Guide to Dimensional Modeling.* Canada; John Wiley & Sons (2nd ed.).

Kimball R., & Ross, M. (2011). *Data Warehouse Toolkit*. *The Complete Guide to Dimensional Modeling.* Canada; John Wiley & Sons (2nd ed.).

Luján-Mora, S., Trujillo, J., & Song, I. Y. (2006). A UML profile for multidimensional modeling in data warehouses. *Data & Knowledge Engineering*,*59*(3), 725-769.

Lapouchnian, A. (2005). Goal-oriented requirements engineering: An overview of the current research. *University of Toronto*.

Liu, L., & Yu, E. (2004). Designing information systems in social context: a goal and scenario modelling approach. *Information systems*, *29*(2), 187-203.

El Mohajir, M., & Jellouli, I. (2014). Towards A Framework In Corporation Functional And Nonfunctional Requirements For Data Warehouse Conceptual Design. *IADIS International Journal on Computer Science & Information Systems*, *9*(1).

M. Gollfarelli, (2010). From User Requirement to Conceptual Design in Data warehouse Design. *Data Warehousing Design and Advanced Engineering Applications:Method fromComplex Construction,* p.1.

Martínez, A., Pastor, O., Mylopoulos, J., & Giorgini, P. (2006, June). From Early Requirements to Late Requirements: A goal-based approach. In *Proceedings of Eight International Bi-Conference Workshop on Agent-Oriented Information System (AOIS-2006)* (pp. 5-12).

Mai, Y., Li, J., & Viktor, H. L. (2004). UML for data warehouse dimensional modeling. *Management Information Systems 2004: GIS and Remote Sensing,8,* 201.

Maxwell, J. A. (1998). Designing a qualitative study. *Handbook of applied social research methods*, 69-100.

Mazón, J. N., Pardillo, J., & Trujillo, J. (2007). A model-driven goal-oriented requirement engineering approach for data warehouses. In *International Conference on Conceptual Modeling* (pp. 255-264). Springer Berlin Heidelberg.

Nasiri, A., Zimányi, E., & Wrembel, R. (2015, April). Requirements Engineering for

Data Warehouses. In *EDA* (pp. 49-64)

Nguyen, D. C., Perini, A., & Tonella, P. (2007). A goal-oriented software testing methodology. In *International Workshop on Agent-Oriented Software Engineering* (pp. 58-72). Springer Berlin Heidelberg.

Park, G., Chung, L., Hong, J. E., Garrido, J. L., & Noguera, M. (2016). Problem-Aware Traceability in Goal-Oriented Requirements Engineering. DOI reference number: 10.18293/SEKE2016-210.

P. Giorgini, S. Rizzi and M. Garzetti, (2008). GRAnD: A Goal- Oriented Approach to Requirement Analysis in Data Warehouses. *Decision Support Systems, 45* (1), pp. 4-21.

P. Giorgini, S. Rizzi and M. Garzetti, (2005). Goal Oriented Re- quirement Analysis for Data Warehouse Design. *DO-LAP*'05, *Proceedings of the 8th ACM International Work-shop on Data Warehousing and OLAP*, Bremen, *45*, pp 47-56.

P.Giorgini, Mylopoulos, J., & Sebastiani, R. (2005). Goal-oriented requirements analysis and reasoning in the tropos methodology. *Engineering Applications of Artificial Intelligence, 18*(2), 159-171.

P. Vassiliadis, A. Simitsis, P. Georgantas,M. Terrovitis, and S. Skiadopoulos. 2005). A Generic and Customizable Framework for the Design of ETL Scenarios. *Information Systems, Elsevier Science Ltd, vol. 30,* no. 7, pp. 492-525, Nov 2005.

Priebe, T., & Pernul, G. (2000). Towards OLAP security design: Survey and research issues. In *Proceedings of the ACM International Workshop on Data Warehousing and OLAP* (pp. 33-40). Washington, DC.

Pedersen, T. B., & Jensen, C. S. (1998). Multidimensional data modeling for complex data. In *Data Engineering, 1999. Proceedings., 15th International Conference on* (pp. 336-345). IEEE. Sydney, Australia

Raghupathi, W., & Umar, A. (2008). Exploring a model-driven architecture (MDA) approach to health care information systems development. *International journal of medical informatics*, *77*(5), 305-314.

Rizzi, S. (2007). Conceptual modeling solutions for the data warehouse. *Data Warehouses and OLAP: Concepts, Architectures and Solutions*, 1-26.

Ramudhin, A., Chan, E., Benziane, R., & Mokadem, A. (2006). Modeling and optimization of health care processes. *IIE Annual Conference.Proceedings*, 1-6.

Raisinghani, Mahesh S (2000). Adapting Data Modeling techniques for data warehouse design. *The Journal of Computer Information System*; Spring 2000; 40,3.

Supakkul, S., Zhao, L., & Chung, L. (2016, June). GOMA: Supporting Big Data Analytics with a Goal-Oriented Approach. In *Big Data (BigData Congress), 2016 IEEE International Congress* on (pp. 149-156). IEEE.

Sheta, D., Osama, E., & Eldeen, A. N. (2013). The technology of using a data warehouse to support decision-making in health care. *arXiv preprint arXiv:1307.3061*.

Sharma, S., & Jain, R. (2013). Enhancing business intelligence using data warehousing: A Multi Case Analysis. *International Journal*, *1*(7).

Saroop, S., & Kumar, M. (2011). Comparative Analysis of Data warehouse Design Approaches from Security Perspectives. *International journal of computer trends and technology*. ISSN: 2231-2803. Retrieved from http://www.internationaljournalssrg.org

Saeki, M. (2010). Semantic requirements engineering. In *Intentional Perspectives on Information Systems Engineering* (pp. 67-82). Springer Berlin Heidelberg.

Siena, A, Bonetti, A, & Giorgini, P. (2008). Balanced Goalcards: Combining Balanced Scorecards and Goal Analysis. *In Third Int. Conf. on Evaluation of Novel Approaches to Software Engineering (ENASE 2008)* (pp. 107–114). Funchal, Portugal: INSTICC Press.

Sarkar, A. (2012). Data Warehouse Requirements Analysis Framework: Business-Object Based Approach. International Journal, 3.

Salim, N., & Ibrahim, R. (2011). Towards data warehouse quality through integrated requirements analysis. In *Advanced Computer Science and Information System (ICACSIS), 2011 International Conference on* (pp. 259-264). IEEE.

Singh, Y., Gosain, A., & Kumar, M. (2009). From early requirements to late requirements modeling for a data warehouse. In *INC, IMS and IDC, 2009. NCM'09. Fifth International Joint Conference on* (pp. 798-804). IEEE.

S. Luján-Mora (2005). *Data Warehouse Design with UML* (PhD Thesis). Universiti d'Alacant.

S. Luján-Mora and J. Trujillo. (2004). Physical Modeling of Data Warehouse using

UML. (*DOLAP*'04), *Washington, DC*, USA.

S. Luján-Mora and J. Trujillo. (2004). *A Data Warehouse Engineering Process.* In Proceeding if the 3rd Biennal International onference on dvance in Information Systems (ADVIS'04),of *Lecturer Notes in Computer Science*, Izmir, turkey, Springer-Verlag. 3261, pp. 14-23.

Stefanov V, & List B. (2002), *A UML Profile for Modeling Data Warehouse Usage* *[PDF]. Advances in Conceptual Modeling – Foundations and Applications Lecture Notes in Computer Science Volume 4802, 2007, pp 137-147.

S. Luján-Mora, J. Trujillo. (2002, September). *Extending UML for Multidimensional Modeling*. In Proceedings of the 5th International Conference on Advances in Information Systems (ADVIS'04), of *Lecturer Note in Computer Science,* Dresden, Germany, Springer-Verlag.2460, 290-304.

Silva, C. T., & Castro, J. (2002). Modeling Organizational Architectural Styles in UML: The Tropos Case. In *WER* (pp. 162-176).

Sharma, U., & Gosain, A. (2010). Dimensional Modeling for Data Warehouse. *Indraprastha University, Delhi, India*.

Van Lamsweerde, A., Darimont, R., & Massonet, P. (1995, March). Goal-directed elaboration of requirements for a meeting scheduler: Problems and lessons learnt. In *Requirements Engineering, 1995., Proceedings of the Second IEEE International Symposium on* (pp. 194-203). IEEE.

Wrembel R. & Koncilia C. (2007). Data Warehouse Refreshment. In Idea Group Inc (IGI), *Data Warehouse and OLAP: Concepts, Architectures, and Solutions* (pp. 111-134). Retrived from http://books.google.com.my/books

Wirtz, K., Tauscher, M., Zwerenz, M., Munte, A., und Strategie, V., & Bayern, K. V. (2006). Data Warehousing for Bavarian Out-Patient Public Health Care. In*ECEH* (pp. 263-274).

Winter Corporation, (2005). Introduction to Data Warehousing. In Connolly T. & Begg C.(Eds), *Database System: A practical approach to design, implementation, and management* (1196 -1221). Pearson Education, Inc Boston, Ma.

Winter, R., & Strauch, B. (2003, January). A method for demand-driven information requirements analysis in data warehousing projects. In *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on* (pp. 9-pp). IEEE.

Wang, J., Chen, T. J., & Chiu, S. H. (2005), *Literature Review On Data Warehouse Development.*IACIS Pacific 2005 Conference Program.

Yu, E., & Cysneiros, L. M. (2002). Agent-Oriented Methodologies-Towards a Challenge Exemplar. In *Proc of the 4 Intl. Bi-Conference Workshop on AOIS, Toronto* (Vol. 151).

Yu, E, & Mylopoulos, J. (1998). Why Goal-Oriented Requirements Engineering. In *Fourth Intl. Workshop on Req. Eng.:Foundation for Software Quality (REFSQ'98), Pisa, Italy.*

Yu, & E. S.K (1997). Towards modeling and reasoning support for early-phase requirements engineering. In *Requirements Engineering, 1997., Proceedings of the Third IEEE International Symposium on* (pp. 226-235). IEEE.

E. Yu, (1995) *Modelling Strategic Relationships for Process Reengineering*, Ph.D. Thesis, also Tech. Report DKBS-TR-94-6, Dept. of Computer Science, University of Toronto