A GOAL AND ONTOLOGY BASED APPROACH FOR GENERATING ETL PROCESS SPECIFICATIONS

AZMAN TA'A

DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2012

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
UUMCollege of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Pembangunan sistem gudang data (DW) melibatkanbeberapatugassepertipenentuankeperluan, merekabentukskema DW danmenetapoperasitransformasi data.Sesungguhnya, kejayaansistem DW adalahbergantungkepadakesempurnaanrekabentuk proses penarikan,perubahan, dan pemuatan(ETL).Walaubagaimanapun, masalahbiasa yang berkaitandenganrekabentuk proses ETL sepertipenentuankeperluanpenggunadan spesifikasitransformasi data sukaruntukdiselesaikan.Masalahiniadalahberkaitandengankepelbagaian sumber data, kekaburandalamkeperluanpengguna, dankerumitandalamaktivititransformasi data.Pendekatan semasa mempunyaikekangandalammenyesuaikan semantik keperluan DW kearahrekabentuk proses ETL. Akibatnya, hal ini telah melewatkan proses penjanaan spesifikasi proses ETL. Kerangkakerja semantik sistem DW yang dihasilkandaripadakajian ini digunakan untuk membangunkankaedahanalisiskeperluanbagimerekabentuk proses ETL (RAMEPs) daripada aspek perbezaan perspektif organisasi, pembuat keputusan, danpembangunan sistem denganmengggunakanpendekatanmatlamat danontologi.Ketepatan pendekatan RAMEPs telahditentusahkan denganmenggunakanperisian yang barudibangunkan dandiubahsuai.RAMEPs juga telahdinilaidalamtigakajiankessebenariaituSistem Hal EhwalPelajar, SistemUtiliti Gas, danSistemUsahawanSiswazah.Kajiankesinitelahdigunakanuntukmenunjukkanbagai mana pendekatan RAMEPs bolehdilaksanakandalammerekabentukdanmenjanaspesifikasi prosesETL.Tambahan pula, pendekatan RAMEPs telahdisemakoleh pakar DW untuk mengenalpastikekuatandankelemahannya dan pendekatan baru tersebut telah diterima. Kaedah RAMEPs berjaya membuktikan spesifikasi proses ETL boleh dijana dari fasa awal pembangunan sistem DW dengan menggunakan pendekatan matlamat-ontologi.

Kata Kunci: Analisis keperluan, ProsesETL, Gudang data, Ontologi, Kepintaran Perniagaan

Abstract

Data warehouse (DW) systems development involves several tasks such as defining requirements, designing DW schemas, and specifying data transformation operations. Indeed, the success of DW systems is very much dependent on the proper design of the extracting, transforming, and loading (ETL) processes. However, the common design-related problems in the ETL processes such as defining user requirements and data transformation specifications are far from being resolved. These problems are due to data heterogeneity in data sources, ambiguity of user requirements, and the complexity of data transformation activities. Current approaches have limitations on the reconciliation of DW requirement semantics towards designing the ETL processes. As a result, this has prolonged the process of the ETL processes specifications generation. The semantic framework of DW systems established from this study is used to develop the requirement analysis method for designing the ETL processes (RAMEPs) from the different perspectives of organization, decision-maker, and developer by using goal and ontology approaches. The correctness of RAMEPs approach was validated by using modified and newly developed compliant tools. The RAMEPs was evaluated in three real case studies, i.e., Student Affairs System, Gas Utility System, and Graduate Entrepreneur System. These case studies were used to illustrate how the RAMEPs approach can be implemented for designing and generating the ETL processes specifications. Moreover, the RAMEPs approach was reviewed by the DW experts for assessing the strengths and weaknesses of this method, and the new approach is accepted. The RAMEPs method proves that the ETL processes specifications can be derived from the early phases of DW systems development by using the goal-ontology approach.

Keywords: Requirement analysis, ETL processes, Data warehouse, Ontology, Business Intelligence

Acknowledgement

Alhamdulillah, syukur ke hadrat Allah s.w.t kerana dengan pertolonganNya tesis ini telah dapat disiapkan. Setinggi penghargaan ditujukan kepada penyelia, Dr. Mohd Syazwan Abdullah dan Prof. Madya Dr. Norita Norwawi di atas segala tunjuk ajar, nasihat, dorongan, pengalaman dan ilmu yang dicurahkan semasa menjalankan kajian dan menyiapkan tesis ini. Ucapan terima kasih juga ditujukan kepada pihak sekolah pengajian pengkomputeran dan pihak universiti di atas keperihatinan dan sokongan dan seterusnya pihak Kementerian Pengajian Tinggi kerana membiayai pengajian ini.

Tesis ini juga telah disiapkan dengan bantuan dan sokongan daripada teman seperjuangan, rakan sepengajian, rakan sekerja, rakan penyelidik, Hal Ehwal Pelajar UUM, Perpustakaan dan Pusat KomputerUUM.Terima kasih yang tidak terhingga juga kepada Syarikat Gas Malaysia, Unit Usahawan Kementerian Pengajian Tinggi, dan Institut SAS Malaysia yang telah memberi ruang untuk melaksanakan kajian ini dan membuka peluang yang lebih signifikan kepada pembangunan aplikasi yang berkaitan dengan bidang kajian ini.

Akhirnya, kejayaan tesis ini dihadiahkan kepada seluruh anggota keluarga, khususnya kepada ibu, isteri dan anak-anak yang turut sama mengharungi cabaran serta memahami perjuangan pengajian ini seadanya. Kepada ayah, abah, dan emak mertua yang telah kembali kerahmatullah, semangat kalian akan terus menjadi inspirasi kepada kejayaan ini dan yang seterusnya.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Acronym	Meaning
AAD	Academic Affairs Department
DAAD	Director Academic Affairs Department
AI	Artificial Intelligence
AKEM	Application Knowledge Engineering Methodology
ASIS	Academic Student Information System
BI	Business Intelligence
BISE	Business Intelligence for Student Entrepreneur
CCS	Call Center System
CDLNR	Conceptual Data Language for N Repositories
CEDI	Co-operative and Entrepreneurship Development Institute
CIF	Corporate Information Factory
COG	Corporate Ontology Grid
CS	Computer Sciences
CWM	Common Warehouse Meta-model
DAML	DARPA Agent Mark-up Language
DDL	Data Definition Language
DFM	Dimensional-Fact Model
DL	Description Logic
DM	Dimensional Model
DMS	Document Management System
DOME	Domain Ontology Management Environment
DSA	Design Science Approach
DSS	Decision Support System
DW	Data Warehouse
DW-Tool	Data Warehouse Tool
EAI	Enterprise Application Integration
EII	Enterprise Information Integration
EIS	Enterprise Information System
EKP	Enterprise Knowledge Portal
EPC	Event-Driven Process Chain
ER	Entity Relationship
ETL	Extract-Transform-Loading
EVE	Evolvable View Environment
GAIS	Graduate Academic Information System
GAV	Global-as-View
GRL	Goal-oriented Requirements Language
HR	Human Resources

IaR	Information as Required
IHE	Institute of Higher Education
IRS	Internet Reasoning System
IS	Information System
JDBC	Java Database Connectivity
JDE	J.D. Edwards System
KAON	Kalrsruhe Semantic Web and Ontology Infrastructure
KM	Knowledge Management
KST	Knowledge Sharing Technology
LAV	Local-as-View
MAKMUM	Majlis Keusahawanan Universiti-Universiti Malaysia
MAS	Multi-agent system
MDC	Metadata Coalition
MDM	Multidimensional Modeling
MER	Multi-Entity Relationship
MOF	Meta-Object Facility
MoHE	Ministry of Higher Education
MOMIS	Mediator Environment for Multiple Information Sources
NGDS	Natural Gas Distribution System
NIAM	Natural Language Information Method
NLP	Natural Language Processing
OBSERVER	Vocabulary heterogeneity resolution
ODBC	Object Database Connectivity
ODS	Operational Data Store
OIL	Ontology Inference Layer
OIM	Open Information Model
OLAP	On-Line Analytical Processing
OMG	Object Management Group
ONION	Ontology Composition.
OO	Object-Oriented
OODB	Object-Oriented Databases
ORDB	Object-Relational Databases
ORM	Object Role Modeling
OWL	Web Ontology Language
P2P	Peer-to-Peer
PROMPT	Formalism-independent algorithm for ontology merging and alignment
RDF	Resource Description Framework
RDF-S	Resource Description Framework Schema
RFI	Request for Information
RO	Response Obtained
SA	Staging Area

SCD	Slowly Changing Dimension
SEAL	Semantic Portal
SIM	Semantic Information Management
SPARSQL	Protocol and RDF Query Language
SQL	Structured Query Language
STS	Socio-Technical System
TA-Tool	Transformation Analysis tool
UBIS	Utility Billing Information System
UCM	Use Case Maps
UML	Unified Modeling Language
UMIS	University Management Information System
UniMAP	Universiti Malaysia Perlis
URN	User Requirements Notation
UUM	Universiti Utara Malaysia
WFMS	Workflow Management Systems
WWW	World Wide Web
XMI	XML Metadata Interchange
XML	Extensible Markup Language
XOL	eXtended Ontology Language

CHAPTER ONE–INTRODUCTION

This chapter presents the background and motivation of this research. The chapter defines the research problems and the research gaps, as well as the research questions and research objectives. Then, the research strategy is discussed in three phases, followed by the scope and the research contributions. This chapter ends withan overview of the thesis organization and summary of the thesis.

1.1 Background

The trend of Business Intelligence (BI) system utilization for decision-making and monitoringperformance (e.g., Key Performance Indicator - KPI) has increased tremendously. The BI Verdict (formerly known as the OLAP Report) (2006)[1] reported that the On-Line Analytical Processing (OLAP) market grew from one billion US dollars in the year 1996 to 5.7 billion US dollars in the year 2006. The industry analyst firm, IDC, predicted that the business analytics software will grow by 10.3 percent annually through the year 2011[2]. This prediction is in line with the market survey conducted by BetterManagement[3], which showed that 84 percent of various organizations were using BI systems. Indeed, many studies conducted by researchers and practitioners have shown increasing use of the BI system by small, medium and larger organizations.

[1] http://www.bi-verdict.com/index.php?id=122 (Previously known as olapreport.com)
[2] http://www.oracle.com/corporate/analyst/reports/infrastructure/bi_dw/208699e.pdf
[3] http://www.bettermanagement.com/default.aspx

The contents of the thesis is for internal user only

REFERENCES

Abdullah, M. S. (2006). *A UML Profile for Conceptual Modelling of Knowledge-Based Systems.*Unpublished PhD, University of York.

Abello, A., Samos, J., & Saltor, F. (2002).*YAM2 (Yet Another Multidimensional Model): An extension of UML*. Paper presented at the IDEAS'02.

Abiteboul, S., Cluet, S., Milo, T., Mogilevsky, P., Simeon, J., & Zohar, S. (1999). Tools for Data Translation and Integration.*Bulletin IEEE Computer Society Technical Committee on Data Engineering*.

Agosta, L. (2002). Market Overview Update: ETL. Retrieved September 20, 2007, from http://www.gigagroup.com/

Ahmad, M. N., & Colomb, R. M. (2007).Overview of Ontology Servers Research, Webology 4(2), Article 43.*Webology* Retrieved January 15, 2008, from http://www.webology.ir/2007/v4n2/a43.html

Akkaoui, Z. E., Mazón, J.-N., Vaisman, A., & Zimányi, E. (2012).*BPMN-Based Conceptual Modeling of ETL Processes.* Paper presented at the 14th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2012), Vienna University of Technology.

Aleksovski, Z. (2008). *Using Background Knowledge in Ontology Matching.*Vrije University.

Alexiev, V., Breu, M., Bruijn, J. d., Fensel, D., Lara, R., & Lausen, H. (2005).*Information Integration with Ontologies: Experiences from an Industrial Showcase*: John Wiley & Son Ltd.

Ali, R., Dalpiaz, F., & Giorgini, P. (2010).A Goal-based Framework for Contextual Requirements Modeling and Analysis.*Springer - International Journal of Requirements Engineering, 15*(4), 439–458.

Allemang, D., & Hendler, J. (2008).*Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL*: Morgan Kaufmann.

An, Y. (2007).*Discovering and Using Semantics for Database Schemas.*Unpublished PhD, University of Toronto.

Antonio, A. d., Ramırez, J., Imbert, R., & Mendez, G. (2005). Intelligent Virtual Environments for Training: An Agent-Based Approach. *LNAI*(3690), 82-91.

Antoniou, G., & Harmelen, F. V. (2003). *Web Ontology Language: OWL*: Springer-Verlag.

Aparício, A. S., Farias, O. L. M., & Santos, N. d. (2005).*Applying Ontologies in the Integration of Heterogeneous Relational Databases.* Paper presented at the Conferences in Research and Practice in Information Technology (CRPIT), Sydney.

Archer, M. S., & Tritter, J. Q. (2000).*Rational choice theory: resisting colonization*: Routledge.

Arpírez, J. C., Corcho, O., Fernández-López, M., & Gómez-Pérez, A. (2003).WebODE in a Nutshell.*AI Magazine, 24*(3), 37-48.

Baader, F., Horrocks, I., & Sattler, U. (2005).Description Logics as Ontology Languages for the Semantic Web.*Lecture Notes in Artificial Intelligence, 26*(5), 228-248.

Barzdins, G., Barzdins, J., & Cerans, K. (2009).From Databases to Ontologies.*IGI Global*, 242-266.

Bekke, J. H. t. (1992).*Semantic Data Modeling*: Prentice Hall.

Beneventano, D., Bergamaschi, S., Guerra, F., & Vincini, M. (2003).*Building an integrated Ontology within SEWASIE system.* Paper presented at the 1st InternationalWorkshop on Semantic Web and Databases (SWDB), Berlin, Germany.

Berenbach, B., Paulish, D. J., Kazmeier, J., & Rudorfer, A. (2009).*Software & Systems Requirements Engineering: In Practice*: McGraw-Hill.

Berners-Lee, T., Hall, W., Hendler, J., Shadbolt, N., & Weitzner, D. J. (2006).Creating a Science of the Web.*Science, 313*(5788), 769-771.

Booch, G., Rumbaugh, J., & Jacobson, I. (1999).*The Unified Modeling Language User Guide*: Addison Wesley.

Bostrom, R. P., & Heinen, J. S. (1977). MIS Problems and Failures: A Socio-Technical Perspective. *MIS Quarterly, 1*(3), 17-32.

Bouzeghoub, M., Fabret, F., & Matulovic-Broqué, M. (1999).*Modeling Data Warehouse Refreshment Process as a Workflow Application.* Paper presented at the International Workshop on Design and Management of Data Warehouse (DMDW'99), Heidelberg, Germany.

Bresciani, P., Perini, A., Giorgini, P., Giunchiglia, F., & Mylopoulos, J. (2004). Tropos: An Agent-Oriented Software Development Methodology. *Autonomous Agents and Multi-Agent Systems, 8*(3), 203-236.

Brisaboa, N. R., Penabad, M. R., Places, A. S., & Rodriguez, F. J. (2002).Ontologies for Database Federation.*UPGRADE, Vol. III*(3), 52-61.

Bruckner, R. M., List, B., & Schiefer, J. (2001).*Developing Requirements For Data Warehouse Systems With Use Cases.* Paper presented at the 7th Americas Conference on Information Systems.

Bruckner, R. M., List, B., & Schiefer, J. (2002).*A Holistic Approach for Managing Requirements of Data Warehouse Systems.* Paper presented at the 8th Americas on Information Systems.

Buccella, A., Cechich, A., & Brisaboa, N. R. (2003).*An Ontology Approach to Data Integration.* Paper presented at the JCS&T.

Calvanese, D., Giacomo, G. D., Lenzerini, M., Nardi, D., & Rosati, R. (1998, 20-22 August).*Information Integration: Conceptual Modeling and Reasoning Support.* Paper presented at the 3rd IFCIS International Conference on Cooperative Information Systems, New York, NY, USA.

Calvanese, D., Giacomo, G. D., Lenzerini, M., Nardi, D., & Rosati, R. (2001). Data Integration In Data Warehousing. *International Journal of Cooperative Information Systems, 10*(3), 237-271.

Cao, L., Zhang, C., & Liu, J. (2005). Ontology-based Integration of Business Intelligence. *International Journal of Web Intelligence and Agent System*.

Chaudhri, V. K., Farquhar, A., Fikes, R., Karp, P. D., & Rice, J. P. (1998, July 26-30). *OKBC: A Programmatic Foundation for Knowledge Base Interoperability.* Paper presented at the AAAI-98, Madison, WI.

Chaudhuri, S., & Dayal, U. (1997). An Overview of Data Warehousing and OLAP Technology.*ACM SIGMOD Record, 26*(1).

Chen, P. P.-S. (1976). The Entity-Relationship Model-Toward a Unified View of Data. *ACM Transactions on Database Systems, 1*(1), 9-36.

Cheng, B. H. C., & Atlee, J. M. (2007).*Research Directions in Requirements Engineering*. Paper presented at the 2007 Future of Software Engineering.

Codd, E. F. (1979). Extending the Database Relational Model to Capture More Meaning.*Communications of the ACM, 4*(4), 397-434.

Connolly, T., & Begg, C. (2005).*Database System A Practical Approach to Design, Implementation, and Management* (4th ed.): Pearson Education Limited.

Corcho, O., Fernandez-Lopez, M., Gomez-Perez, A., & Lopez-Cima, A. (2005).*Building Legal Ontologies with METHONTOLOGY and WebODE.* Paper presented at the Law and Semantic Web Conference.

Cui, Z., & O'Brien, P. (2000).*Domain Ontology Management Environment.* Paper presented at the 33rd International Conference on System Sciences, Hawaii.

Cure, O., & Jablonski, S. (2007). *Ontology-Based Data Integration in Data Logistics Workflows.* Paper presented at the ER Workshops CMLSA, Auckland, New Zealand.

Cysneiros, L. M., Werneck, V., & Yu, E. (2004).*Evaluating Methodologies: A Requirements Engineering Approach Through the Use of an Exemplar.* Paper presented at the 7th Workshop on Autonomous Agents.

Daconta, M. C., Obrst, L. J., & Smith, K. T. (2003).*The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*: Wiley Publishing Inc., Indianapolis, Indiana.

Daft, R. L. (2008). *Organization Theory and Design* (10 ed.). Mason, USA: South-Western Cengage Learning.

Denny, M. (2004).Ontology Tools Survey, Revisited. Retrieved September 20, 2007, from http://www.xml.com/pub/a/2004/07/14/onto.html

Doan, A., & Halevy, A. Y. (2005). Semantic-integration research in the database community: A Brief Survey. *AI Magazine, 26*(1).

Dou, D., & LePendu, P. (2006).*Ontology-based Integration for Relational Databases.* Paper presented at the SAC'06, Dijon, France.

Drucker, P. F. (1974). *Management: tasks, responsibilities, practices*: Butterworth-Heinemann.

Farhan, M. S., Marie, M. E., El-Fangary, L. M., & Helmy, Y. K. (2012).Transforming Conceptual Model into Logical Model for Temporal Data Warehouse Security: A Case Study.*International Journal of Advanced Computer Science and Applications (IJACSA), 3*(3), 115-122.

Fellbaum, C. (1998). A Semantic Network of English: The Mother of All WordNets. *Computers and the Humanities, 32*(2), 209-220.

Fensel, D. (2004). *Ontologies: A Silver Bullets for Knowledge Management and Electronic Commerce* (2nd ed.): Springer-Verlag.

Firat, A., Madnick, S., & Grosof, B. N. (2002). Knowledge Integration to Overcome Ontological Heterogeneity: Challenges from Financial Information Systems. *Twenty-Third International Conference on Information Systems*.

Fonseca, F. T., & Martin, J. (2007). Learning The Differences Between Ontologies and Conceptual Schemas Through Ontology-Driven Information Systems. *Association for Information Systems, 8*(2), 129-142.

Franconi, E., & Kamble, A. (2004).*A DataWarehouse Conceptual Data Model.* Paper presented at the SSDBM.

Friedman, T., & Gassman, B. (2005).Magic Quadrant for Extract, Transformation and Loading. Retrieved September 20, 2007, from http://www.gartner.com/

Galhardas, H., Florescu, D., Shasha, D., & Simon, E. (2000).*Declaratively Cleaning Your Data Using AJAX*. Paper presented at the Journ?es Bases de Donn?es Avanc?es (BDA), Portugal.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., & Schneider, L. (2002). Sweetening Ontologies with DOLCE. *LECTURE NOTES IN COMPUTER SCIENCE*(2473), 166-181.

Garcia-Molina, H., Labio, W., & Yang, J. (1998).*Expiring Data in a Warehouse.* Paper presented at the 24th VLDB'98, San Mateo, CA.

Gardner, S. P. (2005). Ontologies and Semantic Data Integration.*Drug Discovery Today, 10*(14), 1001-1007.

Gennari, J. H., Musen, M. A., Fergerson, R. W., Grosso, W. E., Crubézy, M., Eriksson, H., et al. (2003). The Evolution of Protégé: An Environment for Knowledge-Based Systems Development. *International Journal of Human-Computer Studies, 58*(1), 89-123.

Geroimenko, V. (2004).*Dictionary of XML technologies and the semantic Web*: Springer.

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

Giunchiglia, F., & Shvaiko, P. (2004).Semantic Matching.*The Knowledge Engineering Review, Cambridge Univ Press*.

Gogolla, M., Bohling, J., & Richters, M. (2005).Validating UML and OCL models in USE by automatic snapshot generation.*Software and Systems Modeling, 4*(4), 386-398.

Goh, C. H. (1997). *Representing and Reasoning About Semantic Conflicts in Heterogenous Information Systems.* Unpublished PhD, MIT.

Golfarelli, M. (2010).From User Requirements to Conceptual Design in Data Warehouse Design - a Survey.*Data Warehouse Design and Advanced Engineering Applications: Methods for Complex Construction*, 1-16.

Golfarelli, M., Maio, D., & Rizzi, S. (1998). THE DIMENSIONAL FACT MODEL: A CONCEPTUAL MODEL FOR DATA WAREHOUSES. *International Journal of Cooperative Information Systems, 7*(2-3), 215-247.

Graciela, B., Ma. Laura, C., & Omar, C. (2006).*A process for building a domain ontology: an experience in developing a government budgetary ontology*. Paper presented at the Proceedings of the second Australasian workshop on Advances in ontologies - Volume 72.

Gruber, T. R. (1993). A translation approach to portable ontology specifications.*Knowledge Acquisition, 5*(2), 199–220.

Gruber, T. R. (1994). Toward Principles for the Design of Ontologies Used for Knowledge Sharing.*IJHCS, 43*(5/6), 907-928.

Guarino, N. (1998). *Formal Ontology and Information Systems.* Paper presented at the FOIS'98, Trento, Italy.

Guo, M., Li, S., Dong, J., Fu, X., Hu, Y., & Yin, Q. (2003). *Ontology-based Product Data Integration.* Paper presented at the 17th International Conference on Advanced Information Networking and Applications (AINA'03), Xidian University, Xi'an, China.

Haas, L. M., Miller, R. J., Niswonger, B., Roth, M. T., Schwarz, P. M., & Wimmers, E. L. (1999).Transforming Heterogeneous Data with DatabaseMiddleware: Beyond Integration.*Bulletin Data Engineering, 22*(1), 31-36.

Halevy, A. Y. (2005). Why Your Data Won't Mix: Semantic Heterogeneity. *ACM Queue, 3*(8).

Halpin, T. (2001).*Information Modeling and Relational Databases - From Conceptual Analysis to Logical Design*: Morgan Kaufman.

Hammond, M. (2004).*The Fact Gap: The Disconnect Between Data and Decisions*: Business Onjects.

Hansson, S. O. (1994). *Decision Theory - A Brief Introduction*. Royal Institute of Technology (KTH), Stockholm: Uppsala University.

Hatch, M. J., & Cunliffe, A. L. (2006).*Organization theory: modern, symbolic, and postmodern perspectives*: Oxford University Press.

Hellerstein, J. M., Stonebraker, M., & Caccia, R. (1999).Independent, Open Enterprise Data Integration.*Bulletin IEEE Computer Society Technical Committee on Data Engineering*.

Horkoff, J. M. (2012). *Iterative, Interactive Analysis of Agent-Goal Models for Early Requirements Engineering.* Toronto.

Husemann, B., Lechtenborger, J., & Vossen, G. (2000).*Conceptual Data Warehouse Design.* Paper presented at the DMDW, Stockholm, Sweden.

Hutter, D., Stephan, W., Baader, F., Horrocks, I., & Sattler, U. (2005).Description Logics as Ontology Languages for the Semantic Web.In *Mechanizing Mathematical Reasoning* (Vol. 2605, pp. 228-248): Springer Berlin / Heidelberg.

Hwang, M. I., & Xu, H. (2007). The Effect of Implementation Factors on Data Warehouseing Success: An Exploratory Study. *Journal of Information, Information Technology and Organizations, 2*(1).

IEEE.(2004). *SWEBOK - Guide to the Software Engineering Body of Knowledge*. Los Alamitos, CA.

Inmon, W. H. (2002). *Building the Data Warehouse - Third Edition*: John Wiley & Sons, Inc.

Jacky, A., Isabelle, C.-W., & Nicolas, P. (2001).*Dimension hierarchies design from UML generalizations and aggregations.* Paper presented at the Conceptual modeling - ER 2001, Yokohama.

Jarrar, M. (2005).*Towards Methodological Principles for Ontology Engineering.* Unpublished Computer Science, Vrije Universiteit Brussel.

Jasper, R., & Uschold, M. (1999).*A Framework for Understanding and Classifying Ontology Applications.* Paper presented at the IJCAI-99 ontology workshop.

John, T., Lin, P., & James, H. (2002). Representation and reasoning for goals in BDI agents. *Aust. Comput. Sci. Commun., 24*(1), 259-265.

Jureta, I. J., Faulkner, S., & Schobbens, P.-Y. (2007). Achieving, Satisficing, and Excelling. *LECTURE NOTES IN COMPUTER SCIENCE, 4802*, 286-295.

Kaiya, H., & Saeki, M. (2006).*Using Domain Ontology as Domain Knowledge for Requirements Elicitation.* Paper presented at the 14th IEEE Requirements Engineering, Minneapolis/St. Paul, MN.

Karp, P. D., Chaudhri, V. K., & Thomere, J. (2000).*XOL: An XML-based ontology exchange language, Version 0.5, February 17, 2000*.

Kelly, S., & Pohjonen, R. (2009).Worst Practices for Domain-Specific Modeling.*Software, IEEE, 26*(4), 22-29.

Kerremans, K., Temmerman, R., & Tummers, J. (2003).*Representing Multilingual and Culture-Specific Knowledge in a VAT Regulatory Ontology: Support from the Termontography Method.* Paper presented at the OTM 2003 Workshops.

Kim, W., Hong, B.-J., Hong, E.-K., Kim, S.-K., & Lee, D. (2003).*A Taxonomy of Dirty Data.* Paper presented at the Data Mining and Knowledge Discovery.

Kimball, R. (1996). *The Data Warehouse Toolkit - Practical Techniques for Building Dimensional Data Warehouses*: John Wiley & Son.

Kimball, R. (2006). Kimball University: Integration for Real People. Retrieved June 15, 2007, from http://www.intelligententerprise.com/showArticle.jhtml?articleID=190500064

Kimball, R., & Caserta, J. (2004).*The Data Warehouse ETL Toolkit. Practical Technique for Extracting, Cleaning, Conforming and Delivering Data*: Wiley Publishing, Inc., Indianapolis.

Kimball, R., & Ross, M. (2002).*The Data Warehouse Toolkit - The Complete Guide to Dimensional Modeling* (Second ed.): John Wiley and Sons.

Lamsweerde, A. v. (2009). *Requirements Engineering - From System Goals to UML Models to Software Specifications*: John Wiley & Sons Ltd.

Leffingwell, D., & Widrig, D. (2003).*Managing software requirements: a use case approach*: Pearson Education, Inc.

Lenat, D. B. (1995). Cyc: A Large-Scale Investment in Knowledge Infrastructure. *Communications of the ACM, 38*(11).

Lenzerini, M. (2002).*Data Integration: A Theoretical Perspective.* Paper presented at the 21st ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, Madison, Wisconsin.

Leuf, B. (2006). *The Semantic Web - Crafting Infrastructure for Agency*: John Wiley & Son, Ltd.

Levy, A. Y. (1999).*Logic-Based Techniques In Data Integration.* Paper presented at the Workshop on Logic-Based Artificial Intelligence, Washington, DC.

Lujan-Mora, S. (2005).*Data Warehouse Design With UML.*Unpublished PhD, University of Alicante.

Lujan-Mora, S., Trujillo, J., & Song, I.-Y.(2006). A UML Profile for Multidimensional Modeling in Data Warehouse.*Data & Knowledge Engineering, 59*(3), 725 - 769.

Maedche, A., Staab, S., Studer, R., Sure, Y., & Volz, R. (2002).SEAL - Tying Up Information Integration and Web Site Management by Ontologies.*Buletin of The IEEE*.

Mazon, J.-N., Pardillo, J., & Trujillo, J. (2007).A Model-Driven Goal-Oriented Requirement Engineering Approach for Data Warehouses.*LECTURE NOTES IN COMPUTER SCIENCE*(4802), 255–264.

Mazon, J.-N., Trujillo, J., Serrano, M., & Piattini, M. (2005).*Designing Data Warehouse: From Business Requirement Analysis to Multidimensional Modeling.* Paper presented at the REBNITA.

Meersman, R. (2001). *Ontologies and Databases: More than a Fleeting Resemblance.* Paper presented at the OES/SEO Workshop, Rome.

MoHE.(2010). *Dasar Pembangunan Keusahawanan Institusi Pengajian Tinggi*.Retrieved.from http://www.mohe.gov.my/portal/pelajar/program-keusahawanan.html.

Moss, L. (2005).Ten Mistakes to Avoid for Data Warehouse Project Managers.*TDWI's Best of Business Intelligence, 3,* 16-23.

Mouratidis, H., Giorgini, P., Barley, M., Mouratidis, H., Unruh, A., Spears, D., et al. (2009). Enhancing Secure Tropos to Effectively Deal with Security Requirements in the Development of Multiagent Systems

Safety and Security in Multiagent Systems. In (Vol. 4324, pp. 8-26): Springer Berlin / Heidelberg.

Mumford, E. (2000). A Socio-Technical Approach to Systems Design.*Requirement Engineering, 2000*(5), 125-133.

Mumford, E. (2003). *Redesigning Human Systems*: IRM Press.

Niedrite, L., Solodovnikova, D., Treimanis, M., & Niedritis, A. (2007, February 16-19). *Goal-Driven Design of a Data Warehouse-Based Business Process Analysis System* Paper presented at the 6th WSEAS Int. Conf. on Artificial Intelligence, Knowledge Engineering and Data Bases, Corfu Island, Greece.

Nielsen, J. (1997). The use and misuse of focus groups.*Software, IEEE, 14*(1), 94-95.

Nimmagadda, S. L., Dreher, H., & Rudra, A. (2005).*Ontology of Western Australian Petroleum Data for Effective Data Warehouse Design and Data Mining.* Paper presented at the 3rd IEEE International Conference on Industrial Informatics (INDIN).

Noy, N. F., & McGuinness, D. L. (2001).*Ontology development 101: A guide to creating your first ontology* (No. SMI-2001-0880): Stanford Medical Informatics.

Noy, N. F., & Musen, M. A. (2000).*PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment.* Paper presented at the AAAI'00.

Noy, N. F., Sintek, M., Decker, S., Crubézy, M., Fergerson, R. W., & Musen, M. A. (2001).Creating Semantic Web Contents with Protege-2000.*IEEE INTELLIGENT SYSTEMS AND THEIR APPLICATIONS, 16*(2), 60-71.

Nuseibeh, B., & Easterbrook, S. (2000). *Requirements Engineering: A Roadmap.* Paper presented at the The Future of Software Engineering, Limerick, Ireland.

Nwana, H. S., & Ndumu, D. T. (1999).A Perspective on Software Agents Research.*KNOWLEDGE ENGINEERING REVIEW, 14*(2), 125-142.

Ogawa, H., Kumeno, F., & Honiden, S. (2008).*Model Checking Process with Goal Oriented Requirements Analysis.* Paper presented at the 15th Asia-Pacific Software Engineering Conference.

Olivé, A. (2007). *Conceptual Modeling of Information System*: Springer-Verlag Berlin Heidelberg.

OMG.(2003). *Common Warehouse Metamodel (CWM) Specification*.

OMG.(2007). *Ontology Definition Metamodel* (No. ptc/2007-09-09).

Papastefanatos, G., Vassiliadis, P., Simitsis, A., & Vassiliou, Y. (2009). Policy-regulated Management of ETL Evolution. *Springer Journal on Data Semantics (JoDS XIII)*(5530), 146-176.

Parviainen, P., Tihinen, M., Lormans, M., & Solingen, R. V. (2005). Requirement Engineering: Dealing with the Complexity of Sociotechnical Systems Development *Requirement Engineering for Sociotechnical Systems*, 1-20.

Patel-Schneider, P. F., & Fensel, D. (2002).*Layering the Semantic Web: Problems and Directions.* Paper presented at the First International Semantic Web Conference (ISWC2002), Sardinia, Italy.

Patil, P. S., Rao, S., & Patil, S. B. (2011). Data Extraction, Transformation and Loading.*International Journal of Computer Science and Application*, 5.

Ponniah, P. (2007). *Data Modeling Fundamentals - A Practical Guide for IT Professionals*: John Wiley & Sons.

Prakash, N., & Gosain, A. (2008).An approach to engineering the requirements of data warehouses.*Requirements Engineering, 13*(1), 49-72.

Priebe, T., & Pernul, G. (2003).*Ontology-based Integration of OLAP and Information Retrieval.* Paper presented at the 14th International Workshop on Database and Expert System Applications (DEXA'03).

Rahm, E., & Bernstein, P. A. (2001).A Survey of Approaches to Automatic Schema Matching.*VLDB Journal, 10*, 334-350.

Rahm, E., & Do, H. H. (2000).*Data Cleaning: Problems and Current Approaches*.

Raman, V., & Hellerstein, J. M. (2001).*Potter's Wheel: An Interactive Data Cleaning System.* Paper presented at the 27th VLDB, Roma, Italy.

Rizzi, S. (2007).Conceptual Modeling Solutions for the Data Warehouse.*Idea Group Inc.*, 1-26.

Rizzi, S., Abello, A., Lechtenborger, J., & Trujillo, J. (2006).*Research in Data Warehouse Modeling and Design: Dead or Alive?* Paper presented at the DOLAP'06, Arlington, Virginia, USA.

Romero, O., & Abelló, A. (2007).*Automating Multidimensional Design from Ontologies.* Paper presented at the DOLAP'07, Lisboa, Portugal.

Ropohl, G. (1999). Philosophy Of Socio-Technical Systems. *Society for Philosophy and Technology, 4*(3).

Rudin, K., & Cressy, D. (2003). Will the Real Analytic Application Please Stand Up? Retrieved January 20, 2008, from http://www.dmreview.com/issues/20030301/6427-1.html

Rundensteiner, E. A., Koeller, A., & Zhang, X. (2000). Maintaining Data Warehouses Over Changing Information Sources. *Communication of ACM, 43*(6), 57-62.

Sane, S. S., & Shirke, A. (2009).*Generating OWL Ontologies from a Relational Databases for the Semantic Web.* Paper presented at the ICAC3 '09, Mumbai, India.

Sapia, C., Blaschka, M., Hofling, G., & Dinter, B. (1998).*Extending the E/R model for the multidimensional paradigm.* Paper presented at the ER Workshop on Data Warehouse and Data Mining.

Schreiber, Z. (2003). Semantic Information Architecture: Creating Value by Understanding Data. *DM Review*.

Schreiber, Z. (2004). *Semantics: Delivering One Language to The Enterprise.* Paper presented at the 2nd Semantic Technologies for eGov.

Schreiber, Z., & Gonchar, I. (2004).Industry Models for Semantic Information Management. Retrieved September 20, 2007, from http://www.dmreview.com/

Sell, D., Cabral, L., Motta, E., Domingue, J., & Pacheco, R. (2005, August 22). *Adding Semantics to Business Intelligence.* Paper presented at the Database & Expert Systems Application 2005 - 16th International Workshop.

Sen, A., & Sinha, A. P. (2007).Toward Developing Data Warehousing Process Standards: An Ontology-Based Review of Existing Methodology.*Transactions on Systems, Man, and Cybernetics, 37*(1), 17-31.

Shen, G., Huang, Z., Zhu, X., & Zhao, X. (2006).*Research on the Rules of Mapping from Relational Model to OWL.* Paper presented at the OWLED'06, Athens, Georgia (USA).

Shibaoka, M., Kaiya, H., & Saeki, M. (2007). GOORE: Goal-Oriented and Ontology Driven Requirements Elicitation Method. *ER Workshops (LNCS), 4802,* 225-234.

Simitsis, A. (2004). *Modeling and Optimization of Extraction-Transformation-Loading (ETL) Processes in Data Warehouse Environments.*Unpublished PhD, National Technical University of Athens, Athens.

Simon, H. A. (1996). *The sciences of the artificial*: MIT Press.

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A Practical OWL-DL Reasoner *Web Semantic, 5*(2), 51-53.

Skoutas, D., & Simitsis, A. (2006).*Designing ETL Processes Using Semantic Web Technologies.* Paper presented at the DOLAP'06, Arlington, Virginia, USA.

Skoutas, D., & Simitsis, A. (2007). Ontology-Based Conceptual Design of ETL Processes for Both Structured and Semi-Structured Data. *Semantic Web & Information Systems, 3*(4), 1-24.

Smith, B. (2003). Ontology.*Blackwell Guide to the Philosophy of Computing and Information*, 155-166.

Sobreperez, P. (2008). Using plenary focus groups in information systems research : more than a collection of interviews. *Electronic Journal of Business Research Methods, 6*(2), 181-188.

Sommerville, I. (2007). *Software Engineering* (Eighth ed.): Addison-Wesley.

Spyns, P., Meersman, R., & Jarrar, M. (2002).Data Modeling Versus Ontology Engineering.*ACM SIGMOD.*

Stefanov, V., & List, B. (2005).*Bridging the Gap between Data Warehouses and Business Processes: A Business Intelligence Perspective for Event-Driven Process Chains.* Paper presented at the 9th IEEE International EDOC Enterprise Computing (EDOC'05), Enschede, The Netherlands.

Stefanov, V., & List, B. (2007).*A UML Profile for Modeling Data Warehouse Usage.* Paper presented at the ER 2007 Workshops CMLSA, Auckland, New Zealand.

Storey, V. C. (1993).Understanding Semantic Relationships.*VLDB Journal, 2*, 455-488.

Stylianou, A. C., & Kuman, R. L. (2000). An Integrative Framework for IS Quality Management. *Communications of the ACM, 43*(9), 99-104.

Sumathi, S., & Esakkirajan, S. (2007).*Fundamentals of Relational Database Management Systems*: Springer-Verlag Berlin Heidelberg.

Sung, S., & McLeod, D. (2006). *Ontology-Driven Semantic Matches Between Database Schemas.* Paper presented at the 22nd Int'l Conference on Data Engineering.

Sure, Y., Angele, J., & Staab, S. (2002). Guiding Ontology Development by Methodology.*Inferencing, 31*(4), 18-23.

Ta'a, A., Abdullah, M. S., & Norwawi, N. M. (2008).*Ontology-Based Extraction-Transformation-Loading (ETL) Processes Model in Data Warehouse Environments.* Paper presented at the CAMP'08, UPNM, Kuala Lumpur.

Ta'a, A., Abdullah, M. S., & Norwawi, N. M. (2010). RAMEPs: A Goal-Ontology Approach To Analyse The Requirements For Data Warehouse Systems. *WSEAS TRANSACTIONS on INFORMATION SCIENCE and APPLICATIONS, 7*(2), 295-309.

Ta'a, A., Bakar, M. S. A., & Saleh, A. R. (2008).*ACADEMIC BUSINESS INTELLIGENCE SYSTEM DEVELOPMENT USING SAS TOOLS.* Paper presented at the SAS Global Forum, San Antonio, Texas, USA.

Tang, Y., & Meersman, R. (2005).*Judicial Support Systems: Ideas for a privacy Ontology-Based Case Analyzer.* Paper presented at the OTM Workshops 2005, LNCS 3762.

Thayer, R. H., & Dorfman, M. (1990).*System and Software Requirements Engineering*. Los Alamitos, CA: IEEE Computer Society Press.

Tieniu, W., Jianhua, H., Haihe, Z., Yinglin, W., & Tianrui, L. (2011).Design and Implementation of an ETL Approach in Business Intelligence Project. In *Practical Applications of Intelligent Systems* (Vol. 124, pp. 281-286): Springer Berlin / Heidelberg.

Toivonen, S., & Niemi, T. (2004).*Describing Data Sources Semantically for Facilitating Efficient Creation of OLAP Cubes.* Paper presented at the 3rd International Semantic Web, Hiroshima, Japan.

Ullman, J. D. (2000). Information Integration Using Logical Views.*Theoretical Computer Science, 239*(2), 189-210.

Uschold, M., King, M., Moralee, S., & Zorgios, Y. (1998). The Enterprise Ontology.*Knowledge Engineering Review, 13*, 71–88.

Vassiliadis, P. (2000). *Data Warehouse Modeling and Quality Issues.*Unpublished PhD, National Technical University of Athens.

Vassiliadis, P., Simitsis, A., Georgantas, P., & Terrovitis, M. (2003).A Framework for the Design of ETL Scenarios.*LECTURE NOTES IN COMPUTER SCIENCE*(2681), 520-535.

Wache, H., Vogele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., et al. (2001).*Ontology-Based Integration of Information - A Survey of Existing Approaches.* Paper presented at the IJACAI-01.

Walker, G. H., Stanton, N. A., Salmon, P. M., & Jenkins, D. P. (2008). A review of sociotechnical systems theory: a classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science, 9*(6), 479-499.

Wand, Y., & Wang, R. Y. (1996).Anchoring Data Quality Dimensions in Ontological Foundations.*Communications of the ACM, 39*(11), 86-95.

Wang, T.-W., & Murphy, K. E. (2006). Semantic Integration in Multidatabase Systems: How Much Can We Integrate? *Advanced Topics in Database Research, 3*(xxi), 420-439.

White, C. (2006). *The Next Generation of Business Intelligence: Operational BI* (Sponsored by Sybase): BI Research.

Winter, R., & Strauch, B. (2004).*Information Requirements Engineering for Data Warehouse Systems.* Paper presented at the ACM Symposium on Applied Computing.

Yu, E. (1995). *Modeling Strategic Relationships for Process Reengineering.*Unpublished Ph. D thesis, University of Toronto.

Yu, E., & Cysneiros, L. M. (2002).Agent-Oriented Methodologies - Towards a Challenge Exemplar.

Yu, E., Giorgini, P., Maiden, N., & Mylopoulos, J. (2011).*Social Modeling for Requirements Engineering*: The MIT Press.

Zeleznikow, J., & Stranieri, A. (2001).*An Ontology for the Construction of Legal Decision Support Systems.* Paper presented at the 2nd International Workshop on Legal Ontologies.

Zhao, G., Gao, Y., & Meersman, R. (2004). *An Ontology Based Approach To Business Modelling.* Paper presented at the International Conference of Knowledge Engineering & Decision Support (ICKEDS'04).

Zhuolun, Z., & Sufen, W. (2008, 12-14 Oct. 2008). *A Framework Model Study for Ontology-Driven ETL Processes.* Paper presented at the Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08.4th International Conference on.

Ziegler, P., & Dittrich, K. R. (2004).*Three Decades of Data Integration - All Problems Solved.* Paper presented at the 1st Int'l IFIP on Semantics of a Network World.

Appendix A– Case Study forStudent Affairin University

(i) Organizational Modeling