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

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 is to identify the business goals and elaborate the requirement to enhance organization performance (El Mohajir & Jellouli, 2014). There are three approaches involve in requirement analysis to develop DW. One approach is data driven, second is requirements driven approach and lastly is mixed driven (Kumar & Singh, 2010).

Health informatics is currently one of the top focuses of computer science researchers. This study focused on Rural Health Care based in Malaysia. Health care in Malaysia is mainly under the responsibility of the government's Ministry of Health (MOH). Health care in Malaysia is divided into private, public and non-governmental organization (NGO) sector. Based on observations and interview which was done with medical practitioners such as medical officers and nurses at Klinik Kesihatan Mempaga, Klinik Kesihatan Bandar Alor Setar, Klinik Kesihatan Tunjang and Klinik Kesihatan Gua Musang for rural health services, for example patient registrations, Mother Child Health Care (MCH), and Outpatient Department (OPD) Services, in several clinics, there are no DW system implemented in this area to integrate outpatient records to hospital system. These means, all the patient information's still record in paper with particular id number as reference to clinic. But as in medical field, there are huge clinical data needs to systematically record and extracting of these records is complex time consuming. Therefore, DW system is a crucial asset to integrate hospital and clinic records. Availability of timely and accurate data is essential for medical decision making.

However, DW system is a long-terms project, so it is very difficult to estimate future requirements for the decision-making process and high risk to failure. There are a few reasons that contribute to DW system failure; some of them are user requirement analysis (Abai, Yahaya & Deraman, 2013). Most of failed DW system is because developers are tending to focus on development phase without considering the user requirements. User requirements for DW systems are not easy to identify at the early phase because decision processes are structured and shared across the multiple sectors of the organization. Hence, requirements analysis must start with early requirements analysis and further on detailed requirements analysis. Therefore, the requirement analysis in DW system development has been defined into an early and late requirement analysis (Kumar & Singh, 2010).

Analyzing early requirements will decrease the possibility of misunderstanding the user's requirements and reduce the risk of failure for the DW project. Horkoff, (2012) state that the early requirement analysis is typically concerned with the identification and analysis of the stakeholder goals, while late requirement analysis usually focuses on completeness, consistency and automates verification of requirements. All the requirements analysis are useful to meet the goals in order to complete the business operation, decision making, and job performance.

Therefore, in view of attending to these observed gaps, this study aims to develop a DW requirement modeling based on Rural Health Care. Goal-oriented modeling and Unified Modeling Language (UML) approach is used as to produce new requirement modeling approach that comprises early and late requirements. The deliverable of this study is capable to improve DW modeling design in development process.

1.3 Problem Statement

Requirement analysis in DW development process has been divided into two phases which is early requirements and late requirements (Yu & E.S.K., 1997). During early requirement phase, goal-oriented approach assists to identify the information relevant to decision making process. Meanwhile, late requirement phase focuses on completeness, consistency and automated verification of functional demands for a DW (Alencar, Castro, Cysneiros, & Mylopoulos, 2000). However, there are a few problems and limitations identified in early and late requirements phases.

The problems and limitations are identified such as the lack of a current approach to combine both aspects in the requirements analysis to the conceptual design causes many DW projects fail to meet their goals in satisfying users' needs and expectations (El Mohajir & Jellouli, 2014). The failure of DW projects effect the organizations to bear considerable losses to cover development costs DW. Current approach such as ER and UML modeling is designed to capture late requirements in requirements analysis and it facilitates to produce requirements document as required by decision-maker (Gupta, Chauhan, Kumar & Taneja, 2011).

The other problems and limitations identified are lack of standards for requirement analysis (El Mohajir & Jellouli, 2014) to analyse early and late requirements and lack of mechanisms to ensure that the users' requirements are well mapped into the conceptual schema and then validate the obtained schema in the conceptual phase (Golfarelli, 2010). Requirement analysis in DW development is typically overlooked in many ways. DW developer and analyst used the data driven approach ignore early requirement phase and they only focus on analysing the operational data sources to elaborate the target DW conceptual schema (Saroop & Kumar, 2011). Therefore, this cause the requirements are continually changed over time and need to be managed effectively to meet business objective. But to make any changes and modify the DW system often, it is essential to understand the organization environment, the users themselves, goals, constraints and risks to DW system itself. The failure to change the required information totally will affect the decision-making process in organization. Therefore, it is important to start with early requirement phase in requirement analysis to allow the decision-maker to identify the user requirements and make changes accordingly to meet the business goals.

Another problem in DW development process is decision-maker difficult to identify and define the requirements from stakeholders and from the goal of the organization since the current DW modeling does not consider the goal-oriented concept in the DW development (Luján-Mora, 2005). Current DW modeling such as Entity-Relationship (ER) modeling and UML modeling used does not provide the required information and techniques to map high level users' goal to design a DW modeling for understanding the end users due to the lack of communication between the DW developers and the business analysis (Giorgini et al., 2005). This approach has their own graphical notations which is the designer of DW required to learn a new specific model and notations together to design DW modeling in development process.

This is because the ER and UML modeling is designed to capture late requirements in requirements analysis and it facilitates to produce requirements document as required by decision-maker (Gupta, Chauhan, Kumar & Taneja, 2011). Most existing requirements languages and frameworks are intended more to the late requirement phase and this modeling focuses on completeness, consistency and automated verification of functional demands for a DW (Alencar et al., 2000). It shows that ER and UML modeling is less equipped to capture early requirements and business goals (Alencar et al., 2000) in requirement analysis and not an appropriate modeling and design for large DW. Hence, the ER and UML modeling becomes difficult to build if the demands change quickly.

Therefore, the requirement analysis phase for DW development process should start by discovering the goal of decision makers and stakeholders in early requirements phase by using goal-oriented approach. Goal-oriented approach is focused on the early requirements phases (Ellis-Braithwaite, Lock, Dawson, & Haque, 2013). This is because goals are often unclear when first elicited from decision-makers and stakeholders. Based on goal-oriented approach, the requirements and intentions from stakeholder and decision-maker is details analyzed using four analysis techniques. However, goal-oriented approach is suffers from a number of pragmatic concerns such as availability of efficient tools (Cysneiroset al., 2005) to design the model and learning curve due to the complex syntax and semantics, together with inadequate documentation (Cysneiros et al., 2005). Moreover, most of DW requirements engineering approach isolated the early requirements phase from the late requirements phase (Kumar & Singh, 2010) to design a stable DW model and give less risks of DW failure to the organization.

From the issues of DW development processes show that several gaps have been identified in those phases such as the lack of standards for requirement analysis (El Mohajir & Jellouli, 2014), the lack of mechanisms to ensure that the users' requirements are well mapped into the conceptual schema and then validate the obtained schema in the conceptual phase (Golfarelli, 2010). DW systems are often changed due to the user requirement in organization. The current DW modeling doesn't use the goal concept in their modeling. These problems and limitations give high risks to failure and damage the whole current DW and it make the cost become high.

The other problem arise in DW development processes is the current modeling such as ER and UML modeling is both approaches are not suitable for design large DW systems. It is because ER and UML modeling approach does not capture the early requirements has their own notation and required another extension to support the modeling to design large DW. Furthermore, none of these modeling approaches has been widely accepted as a standard modeling to design DW systems.

Therefore, this research is focused on early and late requirements analysis in requirement modeling for designing DW in order to resolve the problems. Early requirements analysis consists of identifying and analyzing the stakeholder's intention, while late requirement analysis focuses on the details of functional and non-functional requirements of the DW (Bresciani, Perini, Giorgini, Giunchiglia & M, J. 2002). All the pre-requisites must be isolated, identified and analyzed within early and late requirement analysis by using suitable modeling.

This study propose goal-oriented modeling which is mapped to UML modeling approaches is to identify and analyze the early and late requirements for design stable DW model that comprised all the needs of stakeholders and decision makers. A comprehensive UML modeling for goal-oriented modeling represent a mechanism that enables the integration of goals with the UML modeling such as Class Diagram. It also establishes as a guideline for developers to assist them in resolving modeling issues in DW domain especially in early and late requirements. The mapping of Goal and UML modeling approach give rationale requirements for analysis and decision making to identify the stable requirements and information that required by stakeholders and help analyst guide the requirements elaboration in requirements engineering to design a stable DW modeling.

1.4 Research Questions

This main question of this research focused on how the DW requirement is modeled by using Goal and UML (G-UML) modeling approaches. The following are the related research questions:-

i. How to identify the early and late requirements for designing DW systems by using G-UML modeling approach?

ii. How to develop DW requirement model by using G-UML modeling approach?

iii. How to verify the correctness of DW requirement model developed?

1.5 Research Objective

The objectives of this research are as follows.

i. To identify the early and late requirements for designing DW systems by using G-UML modeling approach.

ii. To develop DW requirement model by using G-UML modeling approach.

iii. To verify the correctness of DW requirement model developed.

1.6 Scope and Limitation

This study focuses on the DW requirement modeling by using a G-UML modeling approach that have a high possibility for design stable DW model that comprised all the needs of stakeholders and decision makers. Requirement analysis stage is required in order to identify and model the early and late requirements for decision makers and derive a suitable model. A G-UML modeling approach is explored deeper. The requirements of DW were analyzed by using goal-oriented approach and later mapped to the UML approach to realize the design into the implementation. The Malaysian Health Care DW is used as case study, and it focused on the Rural. The G-UML modeling approach is verified by expert review to produces the exactly requirement needs based on the goals of stakeholders clarified in the early and late phase of requirement analysis. This research did not cover the development and implementation of the DW system.

1.7 Research Significance

There are two approaches employed to lead this contribution of G-UML modeling approach in order to define the DW modeling for providing enough information to the users. G-UML modeling approach will enhance the analysis and design approaches, especially for DW development. This is because UML is a well-known modeling and widely accepted for designing with many Computer Aided Software Engineering (CASE) tools. UML modeling has minimized the cost and the effort of designer to learn new notations and concept or methodologies for every subsystem to be modeled. Designer can use the same concept they used to apply.

In this research works, UML modeling is applied whit the Goal-oriented concept to introduce new elements for designing DW requirement modeling. The ability of goal-oriented to analyze the requirement in early requirement analysis helps the designer to analyze the intension of stakeholders and decision-makers. Combine with UML modeling which is focused on late requirement analysis give an advantage to designer to produce a stable DW requirement modeling.

1.8 Thesis Outline

This thesis is divided into 6 Chapters. The content of these chapters is organized. Chapter one (Introduction) provides a brief introduction about DW, motivation and problems for the research that contribute to produce new modeling approach that cover the research gap. Chapter two (Literature Review) discusses the related work regarding to the ER modeling, UML modeling, Goal-oriented modeling and Dimensional modeling. The related research works help to describe details modeling used and support the G-UML modeling approach produced to design DW system. Chapter three (Research Methodologies) discusses about the research method and development processes of G-UML modeling approach to design DW modeling. Chapter four (G-UML Requirement for Requirement Analysis) applied the requirement modeling to Health Care domain and Chapter five (Result of Model Evaluation) explain about the process to verify the DW modeling produced. Chapter six (Conclusions and Future Works) discusses about the conclusion and future works for DW requirement modeling.

1.9 Summary

This chapter has discusses the requirement analysis for DW development process. It comprises the background of study of DW where the main key role of DW is to improve decision making system and increase organization performance. However, the lack of a current approach to combine early and late requirement phase in the requirements analysis to the conceptual design causes many DW projects fail to meet their goals in satisfying users' needs and expectation. The other problems and limitation arise in DW development process is the lack of standards for requirement analysis to analyze early and late requirements and the lack of mechanisms to ensure that the users' requirements are well mapped into the conceptual schema and then validate the obtained schema in the conceptual phase. Another problem in DW development process is decision-maker difficult to identify and define the requirements from stakeholders and from the goal of the organization since the current DW modeling does not consider the goal-oriented concept in the DW development.

This research also focused on the objectives of research, scope and limitation of the research and research significance that need to be gained. It also describes the research domain problem with approach that needs to be used. The aim of this research is to produce a stable DW requirement modeling that comprised all the needs of stakeholders and decision-makers scope and limitation by using G-UML modeling approach. The used of goal and UML modeling approach is to resolve the problems regarding to the requirements analysis, supported the DW and data sources modeling to enhance the design of the DW requirement modeling. In this chapter, a general idea has been presented in order to provide a clear view to perform the research accordingly based on case study Rural Health Care. Chapter two discusses the literature review about the DW, goal and UML modeling.

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

This chapter present literature review related to DW analysis and design. The concepts of requirement DW and modeling DW are discussed, which elaborate the research works on entity-relationship, dimensional modeling, Goal-oriented and UML.

2.2 Data Warehouse Modeling

DW is a database that collects, organizes and manages the business transaction data into a high level of application to provide information for decision making by the organizations. Inmon (2002) defines DW as a centralized repository for the entire enterprise and a DW is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process. While Kimball (2002) defines DW as a copy of transaction data specifically structured for query and analysis. The concept evolved the dimensional modeling that focuses on ease of end user accessibility and provides a high level of performance to the DW.

The key components of the DW model designed are to make the entire environment of DW well functional, manageable and accessible to the business core (Kimball & Ross, 2011). Typically, there are four separate and distinct components of DW to support the development process as Figure 2.1 (Kimball, 2002). The components are operational source systems, data staging area, data presentation area, and data access tools. The process started with the operational data source system that comes from a relational, unstructured and semi structured database based on the user requirements are selected and extracted. At this stage, decision-maker, developer is seeking, extracting the data to be used for analysis within the DW system. Next component are known as a data staging area. At this stage, the selected data sources are going through extracting, integrating, cleansing and transforming. Data mart is created in this staging area. The data are transferred or loaded into DW structure using dimensional modeling. The final component of DW environment is the data access tools which are all data access tools query the data in the DW's presentation area.

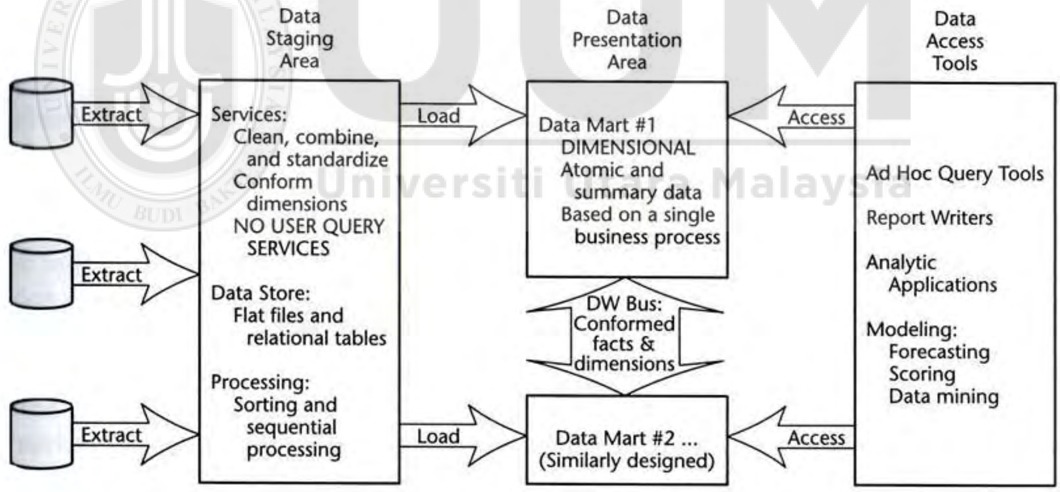


Figure 2.1: The Structure of the Data Warehouse, according to Kimball Group (Kimball, 2002)

2.3 Requirements of DW

Requirements for DW development are not an easy task to capture and analysis because it consists of stakeholders, organizations with different intention to fulfill business goals. There are few subjects need to be discovered before the exact requirements will be defined. There are functional requirements, non-functional requirements and usability requirements (Afreen, Khatoon & Sadiq, 2016).

i. Functional Requirements

In DW, functional requirements describe what the system does or expected to do. The decision-maker describe the process that the system will be required to carry out including the details of inputs into the system, details outputs that expected from the system and details of data produced when querying process (Afreen et al., 2016).

ii. Nonfunctional Requirements

Nonfunctional requirements describe the aspects of the system that concern to the how well-function of the system itself. These include the performance of updating or retrieving data, ability to cope with high level of simultaneous access by many users, anticipated volumes of data and security of data stored (Afreen et al., 2016).

iii. Usability Requirements

Usability requirements will enable decision-maker to ensure the task that users undertake, including the goals of process achieved (Afreen et al., 2016). Stakeholders involved in business transactions and decision-making process stated their requirement needs and prospects of the DW to improve their job. Analyst used Requirement Engineering (RE) process to discover the purpose, by identifying stakeholders and their demands to be analyzed by step by step process from identifying to documenting all the essentials which help to reduce failures of DW.

In this process, RE has divided into two phases which are early requirement phase and late requirement phase in DW development (Singh, Gosain & Kumar, 2009). In early requirement phase according to problem statements in Chapter One, allow analyst and decision-make r to identify and analyze the stakeholder's intention in early requirements analysis process. While late requirement phase consist of late requirement analysis process that focuses on the details of functional and non-functional requirements of the DW (Bresciani et al., 2002). The research works on requirement of DW are discussed as follows.

2.3.1 Research work on Requirement of DW

Nasiri, Zimányi, and Wrembel (2015) proposed an overview of the existing GORE-based methods in the BI domain for DW. The components of the proposed method are adopted from existing GORE based methods for BI systems. This method involves all phases of the decision-making process in the early stage of DW projects where the requirements are captured, as the final objective of DW projects is to help the decision-making process in the organisation. However, this method only covered RE for the multidimensional model, not the requirements of the operations on DW in late requirements phase.

Sarkar (2012) proposed a Business Object (BO) based requirements analysis framework for designing DW system. It is supported with abstraction mechanism and reuse capability. The BO also facilitate the process mapping of requirements descriptions into high level design components of graph called Graph Object Oriented Multidimensional Data (GOOMD) model. The BO framework starts with early requirement phase which is to identify the requirements using business process driven approach. It finally refine the requirements in further detail to map into the conceptual level DW design model using Demand-driven of Mixed-driven approach for DW requirements analysis. The GOOMD model provides a novel graph based semantic and simple but need the powerful algebra to conceptualize the multidimensional data visualization and operational model for OLAP, based on object oriented paradigm at late requirement phase.

Kumar et al. (2010) proposed Agent-Goal-Decision-Information (AGDI) model to support the early and late requirements for development of DWs from the stakeholders' perspective. This model also supports three interrelated modeling activities, namely, organization modeling, decision modeling and information modeling. In this subject field, early requirements are modeled by committing organization modeling and decision modeling activities, whereas late requirements are modeled by committing data modeling activities. The extended of this model is to support organizational modeling and goal modeling activities. However, the AGDI model needs to extend the conceptual modeling issues for development of DWs and there has no a case tool to support the modeling activities.

Winter and Strauch (2003) proposed a comprehensive method in order to determine information requirements of DW users and match these requirements with the available data sources. The activity model represents the core component of a comprehensive methodology for information requirements analysis for DW systems. All the components have been observed and applied in a several company. However, the comprehensive application of the methodology is rare, even if only a several phase has been addressed.

Farhan, Marie, El-Fangary and Helmy (2011) and Salim and Ibrahim (2011) state that it is important to design DW modeling begin with early requirement phase because incorrect or misleading data will produce wrong business decisions, and therefore, a correct design of these processes at early stages of a DW project is necessary to improve data quality. Yu (1995), P. Bresciani et al. (2004) and Lamsweerde et al. (1993) represent used the early requirements analysis phase to explain the *why* of a software system. Farhan et al. (2011) proposed a conceptual model to refresh data warehouse by using create, read, update and delete (CRUD) data using Extraction-transformation-loading (ETL) processes in the early requirement phase to avoid data redundancy and DW failure.

In summary, the related work shows that the early requirement phase is important phase in requirement engineering because any results from this process give high impact and effect to the next process. In early requirement phase, the decision-maker identifies the problems and defines the required requirements to be used. These allow decision-maker to explore the solution and alternative before design the stable DW modeling. The wrong decision or misleading information of stakeholder intention by decision-maker will cause the DW failure. However, the research gap dealing with the issues of early and late requirements phase is covered in the next chapter by using Goal-UML modeling.

2.4 Modeling DW uses Goal-Oriented Supported by Tropos Methodology and *i** Framework

Goal-oriented approach (Yu & Cysneiros, 2002) in modeling DW has been introduced to understand the early requirements of DW to meet business goals. Goal-oriented approach present notation, techniques and processes use for modeling, analyzing requirements in DW modeling. Goal-oriented also support representation and evaluation of alternatives in goal fulfillment and provides automated reasoning tools for various analysis and design tasks for DW modeling (Jiang, 2010).

Goal-oriented is support by Tropos Methodology to build goals model (Nguyen et al., 2007) for requirement analysis in DW (Giorgini et al., 2005, 2008). Tropos is a software engineering methodology (Bresciani et al., 2002; Giorgini, Kolp, Mylopoulos & Pistore, 2003) for building agent oriented software systems and derived from *i** (pronounced *eye-star*), which stands for “distributed intentionally” framework for analyzing requirements.

In DW design, Tropos showed a significant contribution in goal-oriented approach to requirement analysis in two perspectives of modeling which is organizational modeling and decisional modeling. Tropos methodology intended to support five phases of development which is Requirement Analysis (early and late), Architectural Design, Detailed Design and Implementation.

Requirements analysis in Tropos consists of two phases; early requirement and late requirement analysis. Early requirement concerned about the intentions of stakeholder that underlie in DW design, problem in decision-making process on organizational setting, exploring system solution and alternatives (Goldsby & Cheng, 2006). Meanwhile late requirement analysis defined the requirement specifications including functional and non-functional requirements for the DW design. Tropos support from early phases of requirements analysis to detailed design that focused on the understanding of the environment where the DW must operate, and communication between analyst and stakeholder in decision-making process.

2.4.1 The Development Phases

The Tropos methodology suggests spans five phases which are early requirements, late requirements, architectural design, detailed design and implementation. The Tropos methodology guides the development of agent-based systems from early requirements analysis to implementation for deeper understanding of the system-to-be (Bresciani et al., 2002; Giorgini et al., 2003).

i) Phase 1: Early requirement analysis

Early requirement analysis, developer is concerned about the goals of stakeholder that underlie in DW design, problem in decision-making process on organizational setting, exploring system solution and alternatives. This involves the identification of the domain stakeholders and modeling them as social actors. The output of this phase is list of early requirements from requirements analysis, including relevant actor, their goal and their respective dependencies.

ii) Phase 2: Late requirement analysis

During late requirement analysis, the goal-oriented approach is used to model the early requirements based on output from phase 1. There are two perspectives modeling used to model the requirement which is organizational and decisional modeling to produce late requirements. The system-to-be is introduced as another actors and specified within it operational environment, functional and qualities. It also defined the requirement specifications including functional and non-functional requirements for the design. The output of this phase is DW Requirement Modeling based on Goal-oriented Approach, Actor diagram, Relationale diagram and Extended Relationale diagram.

iii) Phase 3: Architectural design

During architectural design, the previous result is used to model conceptual schema. In this stage, the process mapping used organizational model as a bridge to decisional model to design conceptual schema. Then, the process defined the goal artifacts from previous phase, and then mapping conceptual schema to the UML class diagram by using the goal-oriented and UML notation. The output from this phase is conceptual schema and G-UML model.

iv) Phase 4: Detailed design

Detailed design introduces additional detail for each architectural component of the system such as determines multiplicity constraint which is how the goals assigned to each actor are fulfilled by association, aggregation to show the relationship between fact, dimension and measure. The output from this phase is G-UML model with table of constraint.

v) Phase 5: Verification

Expert reviews are required to design conceptual schema to produces the exactly requirement needs based on the goals of stakeholders clarified in the early and late phase of requirement analysis. The G-UML modeling approach is verified by expert review based on their knowledge on UML modeling and expert verification instrument. However, this research did not cover the development and implementation of the real DW system.

2.4.2 Modeling Activities

Modeling activities for Tropos methodology is a series of activities to collect as much information as possible about the system from early requirements until the implementation of system-to-be. There five different modeling activities: Actor modeling, Dependency modeling, Goal modeling, Plan modeling and Capability modeling (Bresciani et al., 2002; Giorgini et al., 2003).

i) Actor modeling consists of a process to identify and analyze the actors of the environment or the system actors by DW analyst and model the requirements in actor diagram.

ii) Dependency modeling consists of a process identify the actors which depend on each other to achieve goals, plans to be executed and resources to be fulfill. The diagram resulting from the actor modeling and dependency modeling is called an actor diagram where each node represents a dependency to each other between two actors.

iii) Goal modeling focuses on actors' goal analysis. The final goal of each process step is considered from the point of view of a specific actor. The goal analysis is performed by using three reasoning techniques: Means-end analysis, Contribution analysis and AND/OR-decomposition. Each analysis contributes to each aims to be achieved. Goal modeling is applied to three Tropos phases: Early requirements, Late requirements, and Architectural design.

iv) Plan modeling focuses on analysis tasks to support the goal modeling. It implies the same reasoning techniques as goal modeling.

v) Capability modeling is applied at the end of Architectural design phase when the goals and dependencies of the different sub-systems are defined. The capabilities of the sub-systems are delimited. It implies that goals and plan modeled in the previous phases become part of the capabilities. UML is generally used to represent those capability and plan diagrams.

2.4.3 The Reasoning Techniques

There are three types of reasoning techniques applied in modeling activities for analyzing the goal or plan modeling to identify the sub-goal or sub-plan for each modeling. The types are Means-end analysis (Lamsweerde, 2009), Contribution analysis and AND/OR decomposition. These techniques have their own purpose and are used for different aims based on the modeling activities. Means-end analysis is aims to discover sub goals representing means to achieve the goals of an actor (Nguyen et al., 2007). Contribution analysis (Figure 2.2) helps to identify goals that may contribute towards the partial fulfillment of the final goals. The positive or negative (++/--) influences the goals to be analyzed (Nguyen et al., 2007). AND/OR decomposition (Figure 2.3) allows for a combination of AND and OR decompositions of goal into sub goal to refining the goal structure.

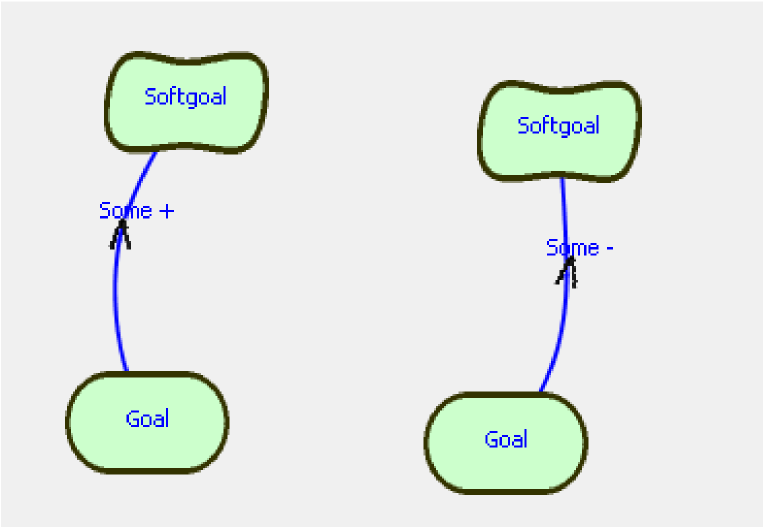


Figure 2.2: Contribution Analysis

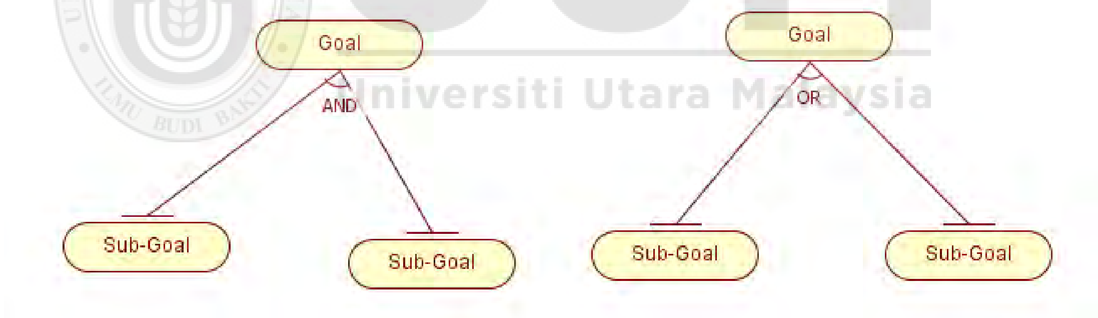


Figure 2.3: AND/OR decomposition

2.4.4 GRAnD for Requirement Analysis Approach

Based on the Tropos methodology, Goal-oriented Requirement Analysis for Data Warehouse or GRAnD (Giorgini et al., 2005) is used as a foundation of this research solution. GRAnD is a goal-oriented technique for requirement analysis offer an alternative to analyze user requirements in DWs since the current analysis approach causing failure to the DW design. GRAnD can be employed within a demand-driven framework or within a mixed supply/demand framework of DW design. Demand-driven approaches start form determining the information requirement of business users. Supply-driven approaches design the DW starting from a detailed analysis of the data sources.

There are two perspectives for requirement analysis which is organizational modeling and decisional modeling to be model accordingly that focused on stakeholders and decision-makers. Figure 2.4 shows an overview of overall diagrams that implemented on both perspectives, derived from organizational and decisional modeling, which complement each other for building the DW modeling by using GRAnD approach. Both modeling consists of three types of analyses which is goal analysis, fact analysis and attributes analysis. The output from this process is conceptual design model for DW. However, for this research, the conceptual design model is mapping to UML Class Diagram to produce a stable design for DW modeling. There are two points of view of organizational modeling and decision modeling which are early and late requirement analysis in DW based on Tropos methodology.

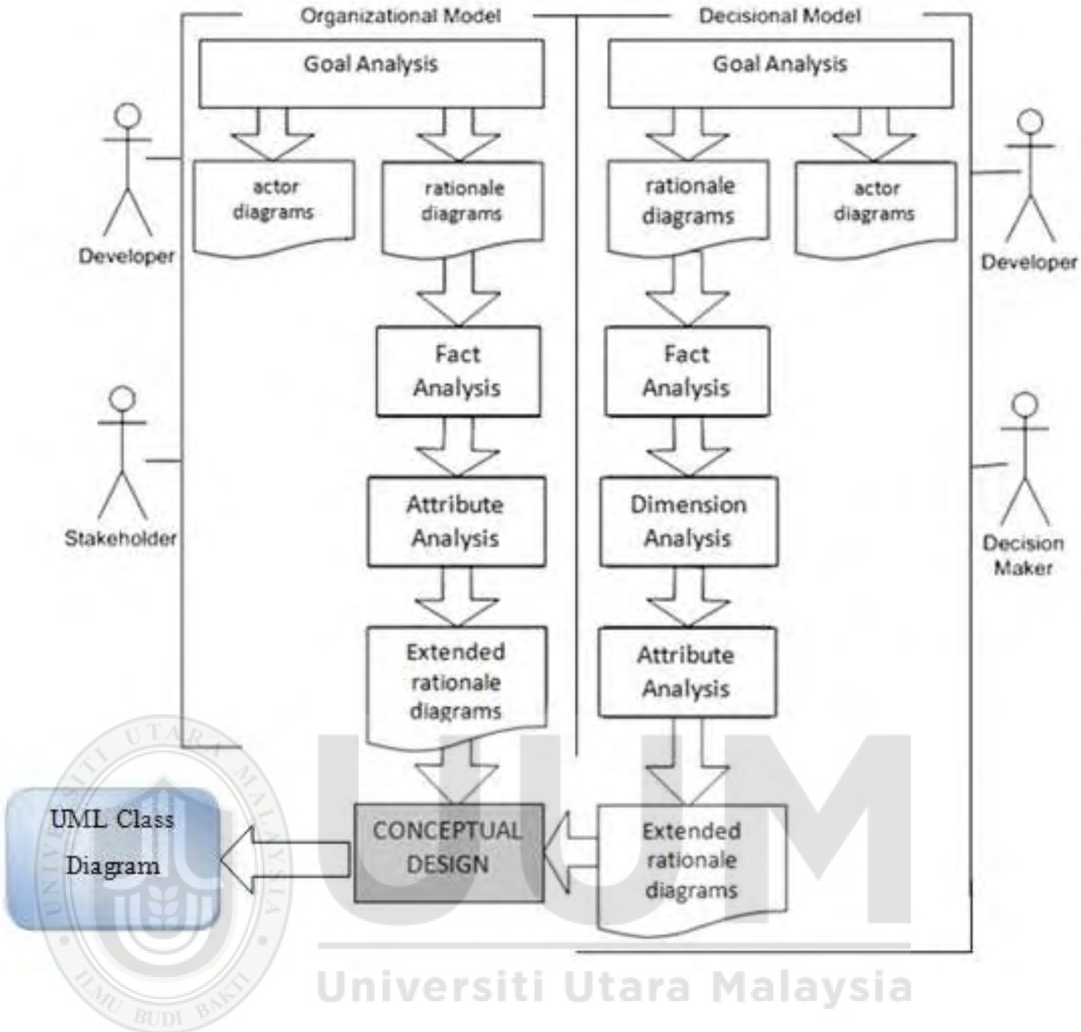
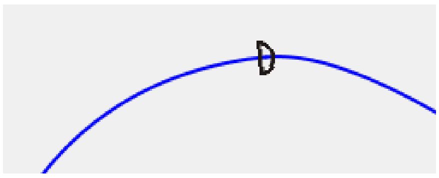


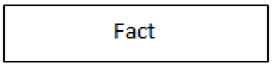
Figure 2.4: The GRAnD approach (Giorgini et al., 2005)

There are a few Tropos notations (shows in Table 2.1) that can be used in the DW design context by using GRAnD approach, such as actors, dependencies, actor diagram and rationale diagram.

Table 2.1: Tropos Notation and Description

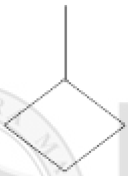
Notation	Description
 A Actor for GRAnD	Actor is representing as an enterprise stakeholder. Actor models an entity that has strategic goals and intentionality such as agents, positions or roles. An actor is symbolized as a circle.
 A Goal for GRAnD	Goal is represent strategic actors' intentions. There are two basic types of goals are considered: hard and soft goal. A goal is symbolized as an oval.
 A Dependency for GRAnD	Dependency between two actors indicates that an actor depend on another in order to achieve goals, execute a plan or exploit a resource. A dependency is symbolized as a line with an arrow in between.
	Fact is determined in both analysis

models. In organizational modeling, a fact is representing a set of events that happen when goal is achieved. In Decisional modeling, the fact represents a set of analysis for goals to be achieved. A fact is symbolized as rectangle.



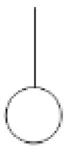
A Fact for GRAnD

Attribute is representing a field, which value is provided when a fact is recorded to fulfill a goal. An attribute is denoted as small diamonds and connected to goals.



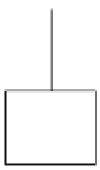
A Attribute for GRAnD

Dimension is a fact property that describes a perspective for looking at the fact to fulfill an analysis goal. A dimension is symbolized as a small circle that connected to goal.



A Dimension for GRAnD

Measure is a numerical property that represents an aggregation aspect if analysis for goal to be achieved by fact of a decision



A Measure for GRAnD

	maker. A measure is symbolized as a small square that connected to goal.

Tropos adopts the *i** modeling framework for analyzing requirements. The *i** framework is an agent-oriented modeling framework that can be used for requirement engineering. Since it supports the modeling activities that take place the system requirement are formulated, it can be used for the early and late phases of the requirements engineering.

The main concept in *i** is an actor model. The actors are described in their intentional properties and have attributes such as goals, abilities, beliefs and commitments. In actor model, an actor depends in other actors for the achievement of its goals, task to perform, and supply the resources. The actors are used to represent the system's stakeholders as well as the agents of the system to be. The *i** framework has two modeling components which is the Strategic Dependency (SD) and the Strategic Rationale (SR) model.

The SD model (actor diagram in Tropos) is a network of dependency relationships among actors. The actor diagram contain of a set of nodes (actors) and links that represents an actor depending on each other to attain some goals. Dependencies represent connecting between two actors; the *depender* and the *dependee*. The depender depends on the dependee, to deliver on the dependum. The dependum can be a goal to be fulfilled, a task to be performed or a resource to be delivered. Actor diagrams are extended during early requirements analysis by incrementally adding more specific actors (agent, roles or position), goals, and dependencies which come out from a means-ends analysis of each goal. This analysis is specified using rationale diagram.

The SR model (rationale diagram in Tropos) is for describing and supporting the reasoning that each actor goes through concerning its relationships with other actors. The SR model process elements are related by two types of links: decomposition links and *meansends* links. These links are used to model AND and OR decompositions of the process elements respectively. The goal, softgoal, task and resources involved in the system are represented according to the *i** notations which explain the relationship between the SR components. During late requirements analysis, the conceptual model developed during early requirements analysis including actors, along with the dependencies between this actors and other environments is extended to UML modeling. These dependencies define functional and non-functional requirements for the DW design.

2.4.5 Research work on Modeling DW uses Goal-Oriented, Tropos Methodology and i* Framework

In the perspective of goal-oriented approach, Park, Chung, Hong, Garrido & Noguera (2016) presented the problem-aware framework for establishing the requirement traceability that focused on goal-oriented requirement engineering. The framework proposed consists of ontology with five layers level of abstractions for early requirement. However, based on the case study used, overall observation comprised that there are difficult to recognize the relations between the problems and goals, goals and requirements, or operation softgoals and functional requirements. But by using the framework, overall, the problem-aware framework based on goal-oriented helps to reduce requirement defects hence, improving the quality of the requirements.

Supakkul, Zhao, and Chung (2016) presented GOMA, a goal-oriented modeling approach to big data analytic. GOME used to capture business goals, reason for the business situations and guide to business decision-making process. Based on their research, goal and problem modeling in early requirement phase from multiple perspectives can help to understand and identify the stakeholders and organizations business goals by using four step processes proposed.

Eridaputra, Hendradjaya, and Sunindyo (2014) proposed a generic requirement model for big data application by using i* and KAOS, a part of goal-oriented requirement engineering. Based on their research, KAOS model is better than i* model to generate the model into requirement. It is because KAOS the model is able to generate functional and non-functional requirements.

Horkoff & Yu (2012) presented the comparison between different approaches to analysis the goal-oriented requirement models, to understand the ways in which procedural design choices affect results. They were advocated goal-oriented to capture and link technical requirements to derive high-level or details user requirements using elicited goals, capture and compare alternative potential implementations. Several applications of goal-oriented techniques in different modeling frameworks, techniques, or methodology include KAOS, GBRAM, NFR, i*, Tropos, GRL, and AGORA.

Giorgini et al. (2008) proposed goal-oriented methodology for requirement analysis in a data warehouse (GRAnD), which can be used within both a demand-driven and mixed supply/demand driven design framework. The GRAnD has adopted Tropos methodology for capturing an early and late requirement and properly manage into the design process. The techniques adopt two different perspectives of requirement analysis: organizational modeling and decision modeling (Giaogini et al., 2008).

Mazón, Pardillo and Trujillo (2007) proposed a GORE approach for modeling, organizational goals that the DW supports and relating them to the information required and to use the *i** modeling framework and the model driven architecture (MDA) in order to describe (i) how to model goals and information requirements for DWs, and (ii) how to derive a conceptual multidimensional model that provides the required information to support the decision making process. Computation independent model (CIM) is specified by using *the i** modeling framework in order to model goals and information requirements for a DW and then the conceptual multidimensional model of a DW is derived from the CIM into the platform independent model, PIM (Mazón et al., 2007).

Jiang, Topaloglou, Borgida and Mylopoulos (2006) applied the techniques of early requirements analysis to produce a goal model to extended database design methodology driven by stakeholder goals. It used to explicit the space of requirement design alternatives between domain stakeholders and development stakeholders.

The research in modeling and designing DW uses the Goal-oriented approach is not new and several approaches such as KAOS, GBRAM, NFR, *i**, Tropos, GRL, and AGORA (Horkoff & Yu, 2011), Tropos methodology (Giaogini et al., 2008), CIM and PIM (Mazón et al., 2007) have attracted intention of the researchers to explore more possibilities of the approaches. However, the research gap dealing with the issues of early and late requirements, apparently the capturing and analyzing from the perspective of Goal-oriented approach still far from resolved. This will be the aimed of this research.

2.5 Entity-Relationship, Dimensional and UML Modeling

In recent years, data modeling is essential to visualization of the business world, the essence of the DW architecture and there are different approaches of data modeling presented in (Ballad et al., 1998) such as Entity-Relationship (ER) modeling and dimensional modeling that are found in a DW environment.

Dimensional modeling represents data with a 'cube' structure. The objectives of dimensional model are to make databases simple and understandable for stakeholders, decision makers, and end user to match the fundamental human need for simplicity and write queries against, and to maximize the efficiency of queries. It achieves these objectives by minimizing the number of tables and relationships between them. Dimensional modeling is different from the third-normal-form (3NF) modeling used by Innom (2002).

The 3NF modeling is design technique that looks for removing data redundancies. The industry sometimes called 3NF model as ER models, an acronym for entity relationships. ER modeling is one of the top-down approaches to design database consist of entities and relationship notation (Winter Corporation, 2005). The ER model view the specific area of interest and using three basic concepts to conceptualize the data which are entities, relationships between those entities (Ballad et al., 1998) and information or property that entities or relationships hold called attribute. ER model is simple, easy to use and understand with a minimum of training.

The conventional ER model constitutes to reduce requirements, redundancy in the data model, assist in data retrieval from difficult identifier, and optimize online transaction. However, ER suffers from a few major problems that exist in data modeling, which is there is less standard of model design and only several notations and diagram exist due to the complexity of DW model. The second problem is in DW modeling, ER diagrams are tending to be difficult to read and untidy because the DW model consists of collection databases and the third problem is to control object-oriented design issues by using an ER approach (Luján-Mora, 2005).

ER diagram represents boxes as a table and lines to communicate the relationships of tables. Same goes to dimensional modeling that can be represented as ER diagram because it consists of joined relational tables. The difference between 3NF and dimensional modeling is the degree of normalization.

Dimensional modeling is useful to represent the requirement of business user by using various approaches to design DW. The intention of the dimensional modeling is to reduce the communication gap between domain users and DW developers. In DW, dimensional modeling is a data modeling technique that separates the information into facts and dimensions. A fact is a primary table in dimensional modeling where the numerical performance measurements of the business are stored and contains the measure on specific requirement of the users. Dimension is a collection of parameter of the same types of view to analyze the fact (Ballad et al., 1998). Dimension table is integral companions to a fact table. Its contain the textual description of the business and measurement in the modeling. Dimensional modeling is developed to support the conceptual and logical data model to describe DW.

2.5.1 Fact

The fact is representing a business major. It focuses on a set of event in the business world and aimed to provide the information to the stakeholder and decision maker.

The fact is represented as a box with two sections as Figure 2.5. A row in a fact corresponds to the fact name and measurement. All facts have two or more foreign keys that connect to the dimensions primary keys.

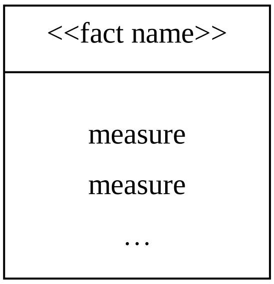


Figure 2.5 Fact Notation

2.5.2 Measure

A measurement is a row in a fact. All the measurements in the fact must be at the same grain. Measure consist of a textual measurement which is a description of the numerical lists of values for example, for fact sales, the measure for the fact sales is the total number of sales per month, total number of quantities per product sold. The measure in the fact will be defined in the second section after fact name as Figure 2.6.

2.5.3 Dimension

The dimension is integral companions to a fact table. It contains the textual description of the business. In well-design dimensional model, dimension is represented as a circle connected to the fact or as a box that have many columns or attributes as Figure 2.7. Each dimension is defined by its single primary key, designated by the PK notation. The primary key is a reference joint to the any given fact. The dimensions implement the user interface to the DW.

2.5.4 Attribute

Dimension attributes serve as the primary source of query constraints, groupings and report labels. It plays a vital role in DW. Attributes consist of textual for product description.

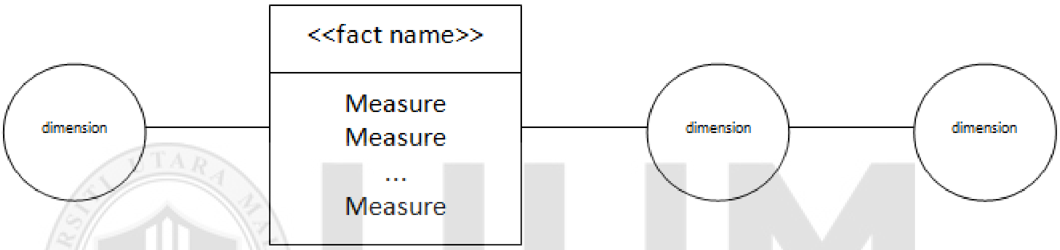


Figure 2.6 Dimension and attribute notation

2.5.5 Hierarchy Model

Hierarchy model is a data model in which data is organized in tree-like model. It consists of more than two or three dimensions and fact as a hierarchical relationship in the business as Figure 2.7. This model gives a detail explanation about the related fact. Hierarchy model help to make the query easier. The relationship between dimension could be many-to-many or many-to-one.

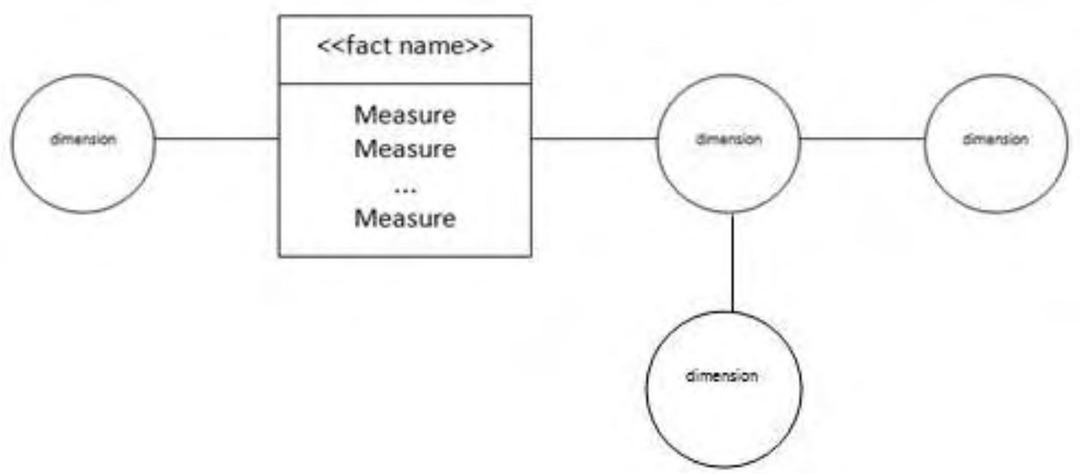


Figure 2.7 Hierarchy Model Notations

2.5.6 Research work on ER, Dimensional and UML Modeling

Chhabra, Kumar and Pahwa (2016) presented a technique for designing a relational schema from an object model and represent in UML form to transform into DW. However the technique not considers the functional and non-functional requirement. The technique used object model to design relational schema then transform it into DW. Inuwa and Oye (2015) proposed a DW database model focused on modeling and designing fro decision making. Dimensional and starER used to model DW. However, the modeling cannot support complex DW.

Kamal & Akhtar (2009) introduced UML Profile for Modeling DW in the different kinds of DW associate conceptual level. The concept uses features of UML that intended for the purpose of creating abstract and the general models. The UML Profile allows to model details of the users functional grouping. Current issues as addressed in Chapter One mention about UML modeling is not appropriate modeling, design for large DW. It demands to use another approach to design DW and UML model designed is not light to accomplish business goals.

Luján-Mora, (2005) explain that for the data modeling purpose, UML as the modeling language that follow the object oriented paradigm define a rich set of diagram such as use case diagram, class diagram, behavior diagram to resolve the ER problems. UML extension mechanism, import elements of the UML Profile for modeling DW usage aims to encompass a different ways used by DW (Stefanov & List, 2002). The UML Profile was applied to model the multidimensional model and focuses on the feature of the profile and provides an overview over the perspectives of usage such as access control, temporal intensity and temporal flexibility.

Mai, Li and Viktor (2004) proposed a met model for DW dimensional modeling by using UML to provide a foundation for modeling the logical core of DW. The traditional dimensional modeling model is a data-driven model and contains a little notation to design the model. With the UML as the foundation, the study has defined a concept to combine the UML notation to design the business process and requirements model. However, this extended model does not capture the entire business life cycle of DW.

Loján-Mora and Trujillo (2004) presented a DW development method based on UML and unified process (UP), which addresses the design and development of both DW backstage and front-end. This allows the user to tackle all the DW design stages from the operational data source to the final implementation of the DW. The researchers have extended the UML function in order to represent the different parts of the DW. Previously, Luján-Mora and Trujillo (2002) presented an object-oriented approach (extension to UML model) to design the multidimensional modeling easily. In addition, this approach has applied the extension OCL to express the well-formed riles of the new defined elements of the multidimensional models.

Silva and Castro (2002) proposed a set of UML extensions for representing organizational architectural styles based on UML modeling. This extension help to build a flexible architecture which is can change continually the requirements. Hence, it helps to realize the stakeholder demand for more flexible and complex DW. However, this extension not available in architecture design phase of Tropos methodology. In Tropos, UML only in details design phase.

Raisinghani and Mahesh (2000) proposed two modeling methods, the ER model and the dimensional model. Both models explored as primary approaches to DW modeling. Nevertheless, both models have issues in specific implementations of each model. ER and dimensional modeling need to be extended to the next level of modeling by combining with other approaches and supported by other tools in the marketplace.

Alencar et al. (2000) used i* technique to present a guideline for the integration of early and late requirement specification and depend on the precise Unified Modeling

Language (pUML), and Object Constraint Language (OCL). UML alone is not adequate to deal with different types of analysis and reasoning that are an important in requirement capture phase. i* technique is useful in DW design to ensure the DW requirement are captured according to business goals. It will help to reduce cost and time consuming in the DW development process. However, i* framework still required to handle some structuring concepts such as agent, role and position to improve the integration of organizational and functional requirements.

In summary, researchers had discovered that dimensional modeling is useful to users understood data and supports performance improvement to predict final users' intention in developing DW in previous works. Many approaches used to accomplish the conceptual design, but none of them can commit a standard dimensional modeling. There are two model uses to represent dimensional model which is Star Model and Snowflake Model.

The research in modeling and designing DW uses the UML approach is not new and several approaches have attracted intention of the researchers to explore more possibilities of the approaches. However, dealing with the issues of early and late requirements, apparently the capturing and analyzing from the perspective of UML approach still far from resolved. We urged, most of the research works was not using goal analysis together with the UML, whereas the important of early stage of requirement can be achieved through goal-oriented and late stage of requirement can be achieved through UML. Therefore, this research explored to combine both methods to model the requirements of the DW. Moreover, the emergence and maturity of goal-oriented and UML approaches have promise the solution to this issue.

2.6 Data Warehouse in Malaysia Health Care

Today, DW is not only deployed extensively in banking and finance, consumer goods and retail distribution and demand-based manufacturing, it has also become as a backbone for business intelligence in noncommercial sector, mainly in medical fields, government, military services, education and research community. The use of stand-alone systems does not integrate with information management and decision-making will have a difficult time to organize and utilize these volumes of data effectively and efficiently.

This shows that a need of implementing the DW concept in this institution's health care as a solution to integrate the system in order to collect, collate and distribute the health data. From the research problems addressed in Chapter One, this research study will be focused on Rural Health Care as the case study.

Implementing DW in Rural will give an opportunity to the clinic to get a guaranteeing high quality service and support of all the inherent activities involved in medical, clinical, pharmacy and management and running of healthcare facility. Patient information can be conveniently shared and accessed by multiple simultaneous users at different positions and experienced a collaborative between other clinicians. The purpose of this research is to study what decision-making issues are faced by stakeholders, doctors, nurses and administrator and patient of health care with the current information systems and how the decision-making might be improved within this health care setting by putting through a UML and goal-oriented modeling.

Based on previous studies by other researchers in health care area, Ado et al (2014) presents a proposed architecture for Healthcare DW for Diabetes diseases. The dimensional model has been used to support and implements querying and analysis for the purpose of decision making. They use the technology of Online Analytical Processing (OLAP) and Extract, Transform and Load (ETL) in Designing Diabetes DW. However, the use of relational design and relational database technology are not feasible implementations to support OLAP design because of the complexity of the queries and OLAP queries are not real-time queries because of the refresh cycle of data into the OLAP data repository (Wrembel & Koncilia, 2007). The ETL process is generally complicated because data must be integrated and transformed for loading the non-normalized relational schema usually associated with OLAP environments (Sheta, Osama & Eldeen, 2013).

Raghupathi and Umar (2008) on their research had explored the potential of the model-driven architecture (MDA) in health care information system development. An MDA is designed and developed for a health clinic system to track patient information. A prototype of this MDA is implemented using an advanced MDA tool.

The UML provides the underlying modeling support in the form of class diagram called Platform-Independent-Models (PIM) and rational model called Platform-Specific-Models (PSM). The separation of the PIM from the PSM enables the incorporation of open standards, leading to interoperability in healthcare DW. However, it has limits, the use of the UML for PIM and the PSM characterizes and open system approach based on standardization. While has been applied and tested sufficiently to lend itself to modeling the complex process in healthcare.

Wirtz et al, (2006) in their research had mentioned about DW for outpatient in KassenärztlicheVereinigung Bayern, Bavarian Association of Statutory Health Insurance Physicians (KVB). All the KVB data are stored in two base star schemas which are the fact-views F_Services and F_ICD since services and diagnoses are not related to each other on a Health Insurance Treatment Voucher (HITV). They contain atomic business data items: each single service or diagnosis for every single day. All other facts-views are aggregates of the base fact-views.

Previous researchers had discovered the healthcare modeling by using start schemas model, architectures modeling using the class diagram UML, OLAP and ETL in designing the DW model. However, there are a few research gap founding when design the modeling. For example, the ETL process is generally complicated because data must be integrated and transformed for loading the non-normalized relational schema.

2.7 Summary

This chapter discussed about related work on DW components that assist to design the DW modeling, such as entity-relationship modeling, multidimensional modeling, UML extensibility mechanism, and goal-oriented of requirement analysis based on Tropos methodology. In requirement engineering, there are early requirement phase and late requirement phase. Early requirement phase help the decision-maker to identifies and defines the required requirements to be used based on stakeholder intention. These allow decision-maker to explore the solution and alternative before design the stable DW modeling. The wrong decision or misleading information of stakeholder intention by decision-maker will cause the DW failure. Late requirement phase focuses on the details of functional and non-functional requirements of the DW.

There are many researches uses agent-oriented modeling approach such as KAOS, GBRAM, NFR, *i**, Tropos, GRL, and AGORA (Horkoff & Yu, 2011), Tropos methodology (Giaogini et al., 2008), CIM and PIM (Hainaut et al., 2007) to explore more possibilities of the approaches to design DW model. The dimensional modeling such as Star Scheme and Snow Flake Scheme is useful and widely known as DW model give an advantage to users more understood the data and supports performance improvement to predict final users' intention in developing DW in previous works. Many approaches used to accomplish the conceptual design, but none of them can commit a standard dimensional modeling of DW.

However, the research gap dealing with the issues of early and late requirements, apparently the capturing and analyzing from the perspective of Goal-oriented and dimensional modeling approaches are still far from resolved. There is still lack of used of early requirement phase and late requirement phase in DW development processes that give high possibility to DW failure.

In Malaysia, DW for Malaysian Health Care especially in rural health care is still far to implement and covered. Decision-maker such as medical practitioner and top-management of ministry faced the same problems every time to get right information and data to enhance their service to patients. The data and records from rural health care are no linked to hospital, pharmacy, and management, medical, other departments system and medical practitioner still used paper to record the patient's information. Their record and data are still cannot be access by multiple user.

The development of DW system for Rural Health Care give an opportunity to the clinic to get a guaranteeing high quality service and support of all the inherent activities involved in medical, clinical, pharmacy and management and running of healthcare facility. Patient information can be conveniently shared and accessed by multiple simultaneous users at different positions and experienced a collaborative between other clinicians. The purpose of this research is to study what decision-making issues are faced by stakeholders, doctors, nurses and administrator and patient of health care with the current information systems and how the decision-making might be improved within this health care setting by putting through a goal-oriented and UML modeling.

Based on the literature review, the research gap is identified and solution for the research is proposed. UML approach is not new and several approaches have attracted intention of the researchers to explore more possibilities of the approaches. However, most of the research works was not using goal analysis together with the UML, whereas the important of early stage of requirement can be achieved through goal-oriented and late stage of requirement can be achieved through UML. Therefore, this research explored to combine both methods to model the requirements of the DW. Moreover, the emergence and maturity of goal-oriented and UML approaches have promise the solution to this issue.

CHAPTER THREE
RESEARCH METHODOLOGY

3.1 Introduction

This chapter describes a method and techniques used in this research. The method is used as a guideline to conduct the research and fulfill all the requirements needs in order to achieve the objective of the research based on previous chapters. Previous chapter had discussed thoroughly the issues and problems in modeling DW.

3.2 Research Design Model

The G-UML modeling approach shown in Figure 3.1 indicates steps and activities supported by Tropos Methodology that are involved in order to achieve the research objectives. Based on stages in Tropos Methodology (Section 2.4.1), there are four major steps involve in research design model. The first step is requirement elicitation, second step is requirement modeling, third step is developing requirement model for DW by using G-UML modeling approach and last step is requirement model verification. This processes applied to Rural Health Care in Chapter Four.

The step involve in early requirement phase is requirement elicitation and requirement modeling. In early requirement phase, the process begin with identified the goals of stakeholder that underlie in DW design, problem in decision-making process on organizational setting, exploring system solution and alternatives. This involves the identification of the domain stakeholders and modeling them as social actors using goal-oriented. The output of this phase are actor diagram, relationale diagram and predefined template which is used to store the requirements analysis result, including relevant actor, their goal and their respective dependencies.

The development process in step 3 used G-UML modeling approach to mapping the result from early requirement phase. The process mapping used to map the conceptual schema to G-UML model. The result produce in step 3 are verified by expert review in Chapter 5.

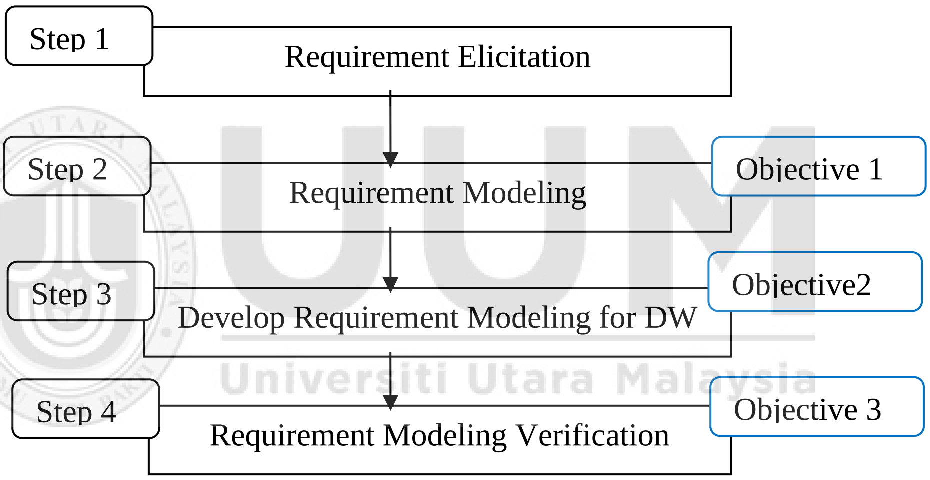


Figure 3.1: G-UML modeling approach

The G-UML modeling approach used the design DW model extended from GRAnD approach (Figure 2.4.4). The processes are based on the GRAnD approach (Giorgini et al., 2005) which is focused on organizational modeling and decisional modeling. The steps involved in Figure 3.1 and Figure 3.2a are details illustrated as Figure 3.2b. Each phases involved are discussed in next subsection.

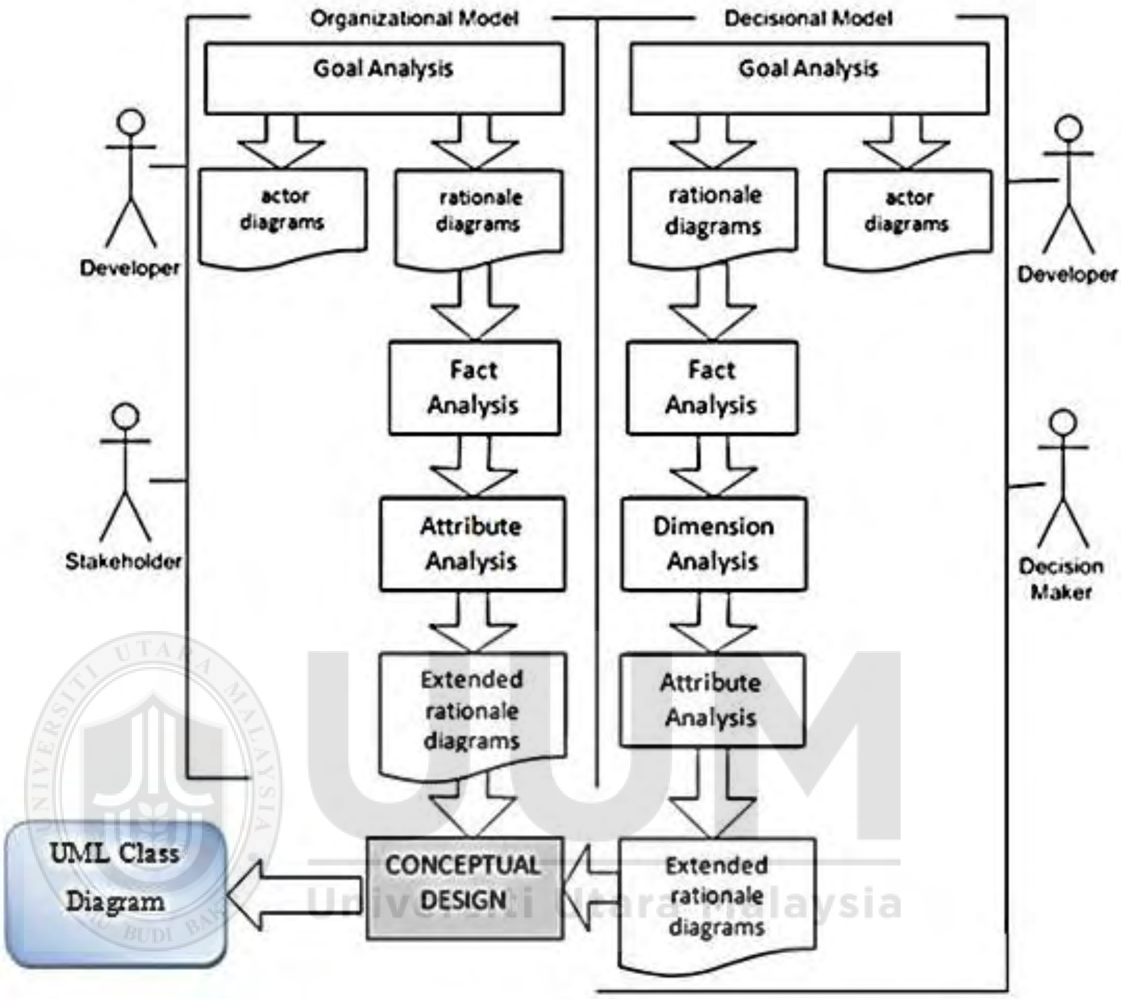


Figure 3.2a: The Extended Research Design Model from GRAnD

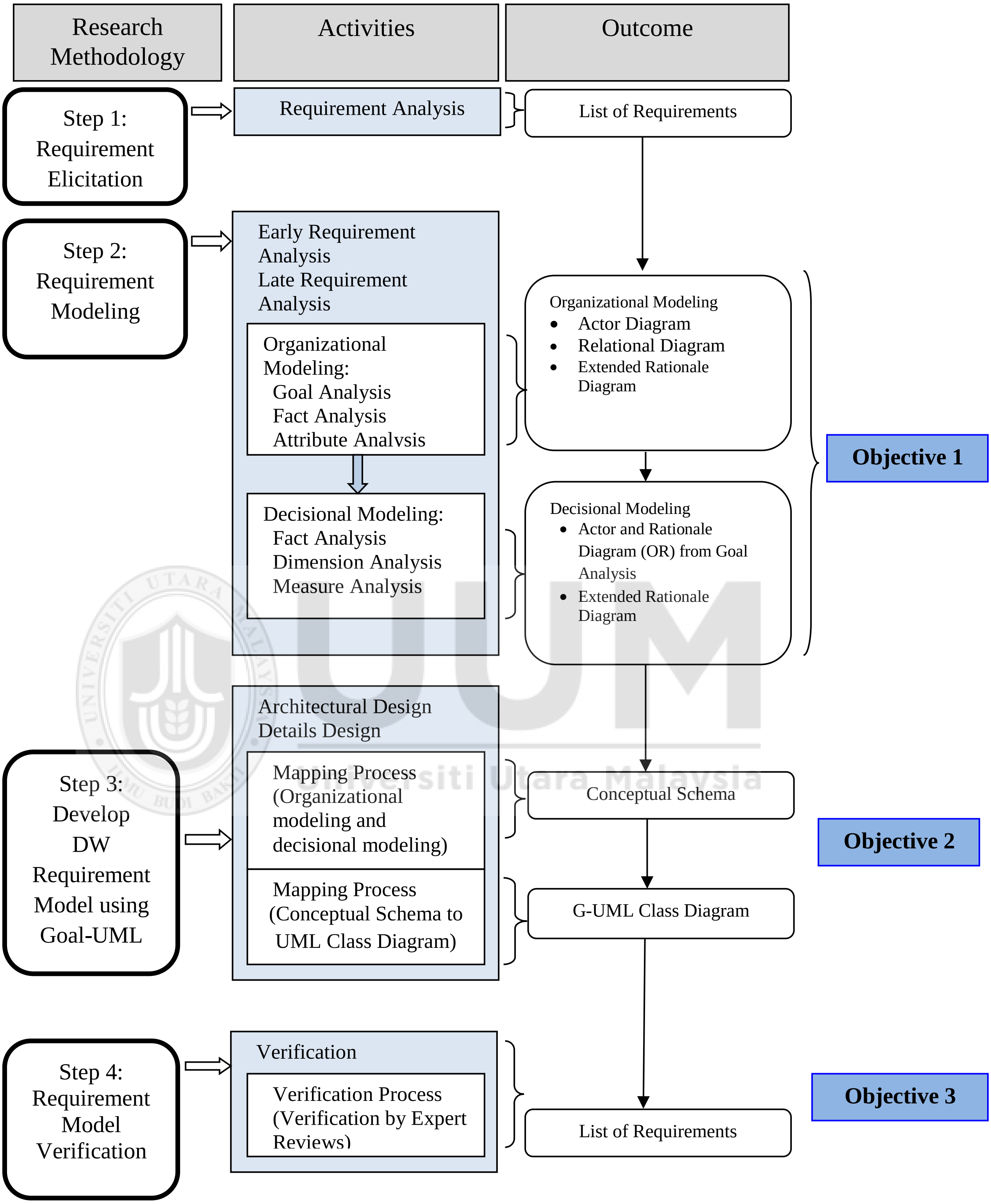


Figure 3.2b: The G-UML modeling approach details illustrated

3.3 Step 1: Requirement Elicitation

Requirement elicitation is a first step where decision-maker and analyst need to identify what a new DW should be able to do with its development. In specification of what the DW must do is based on the users' requirement, and gathering this information's from users is a key role of the DW system analyst or business analyst. Requirements include what the existing DW does, and what the new DW has to do that the existing doesn't do.

The qualitative research method is used to find out what is the goal of organizations. There are several types of fact-finding used to investigate and collect requirements. There is organization background reading, interview, observation and document sampling. The subjects of the investigation include many stakeholders, decision-makers, business analyst and end-users. Tropos methodology and goal-oriented approach are use in this stage to analyse all the requirements gathered.

i. Background Reading

Background reading helps the analysts to get an understanding of the organization before meeting people who work there. It also helps with being aware the business objectives of the organization. The type of documentations that is perfect to get sources and information, including company reports, organization charts, policy manuals, job description, annual reports, and documentation of existing system, journals, and websites that consistent to business organizations.

ii. Interviewing

Interviews are used to gather information in depth from management about their objectives for the organization and people’s requirements for the new information’s system. While conducting an interview, the analyst also can use the opportunity to get the documents that the interviewees use in their work.

iii. Observation

Observation help analyst to provide a better understanding the flow of the job process, what the information people use to carry out their job, the document uses to record the data, and how well the current system handles their needs. Observation allows the analyst to follow the entire process through the start till the end.

iv. Document Sampling

Analyst collects copies of blank or completed documents during the interview observation sessions. The documents, sampling use to determine the information that is used by people in their work from the input and output of the processes. The paper - based document gives give the actual view of the process flow.

All of these requirements are analysed to identify what is the main goal of organization, task to be done, and resource to be fulfilled. The result for requirement modeling is listed in a table. The analyst elicited and analysed user requirements because of its potential to recognize the current organizational situation and disclose the business goals of the functional and non-functional requirements. This requirement elicitation is applied to Malaysian Health Care focused on Rural Health Care in the next chapter. The details output of this step 1 is on Chapter 4, section 4.2, page 90.

3.4 Step 2: Requirement Modeling

Based on the previous result in step 1, the process begins with design the DW modeling using Goal-oriented approach to produce organizational and decisional modeling. Both modeling is used to mapping the organizational goals and decision-maker perspective. The process modeling starts with design the organizational modeling. The organizational modeling represents the main data in the organization and comprise most relevant attribute to data sources.

3.4.1 Organizational Modeling

Organizational modeling consists of three types of analyses which is goal analysis, facts analysis and attributes analysis and produced lists of facts and attributes. (i) goal analysis, in which actor and rationale diagrams are produced; (ii) fact analysis, in which rationale diagrams are extended with facts; and (iii) attribute analysis, in which rationale diagrams are further extended with attributes. Each phase is a different iterative process taking in input the diagrams produced by the previous one.

3.4.1.1 Goal Analysis

The first step for goal analysis is to represent the intentions of stakeholders for the organization and their social dependencies. In this step, requirements are analyzed by using goal analysis to produced actor diagram and rationale diagram according to the goals need to be achieved by the stakeholder and decision maker. Then, based on the diagrams produced, the analyst navigates the rationale diagram of each actor and extends it by associating goals with the facts. Facts analysis is carried out according to top-down approach.

i. Actor Diagram

Based on predefined template, actor diagram and relationale diagram are used to model early requirements and use notation as Figure 3.3. An actor diagram is a graph of actors, a diagram that shows the strategic dependencies among each other. It used to model the dependencies between each actor as shows in Figure 3.4. The diagram shows relation the why and how the actors are related. The dependency represents the dependum between two actors, the depender and dependee. The depender depends on the dependee, to deliver the dependum. The dependum can be goal to fulfill, tasks to be achieved, and resource to be delivered. The output of this actor diagram is on Chapter 4, section 4.3.1, page 97.

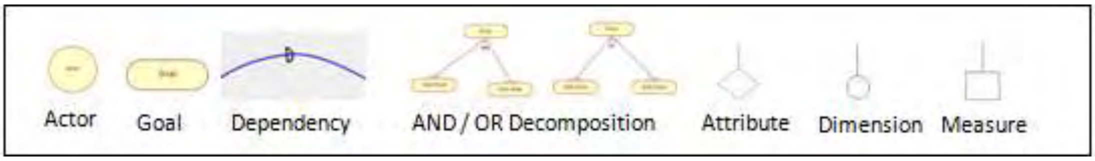


Figure 3.3: Notation for actor and Rationale diagrams

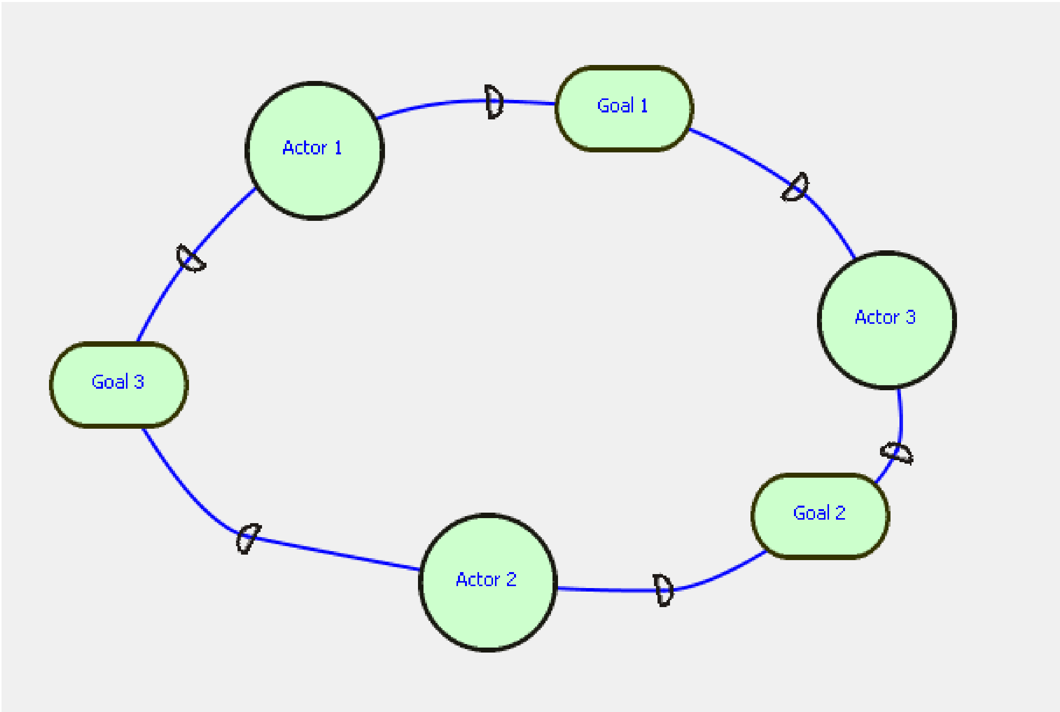


Figure 3.4: The Actor Diagram

ii. Rationale Diagram

Second diagram is called rationale diagram. Rationale diagram is used to represent the logical foundations for rules applied in the relationships between the actors. The rules are applied for decomposing the goals to sub goals by using several techniques which are MEANS-END, AND/OR, Contribution.

Rationale diagram appears as a circle or called boundary (dashed) where the goals of specific actor are analyzed and dependent with another actor are established as Figure 3.5. Goals are decomposed into sub goals and positive/negative contributions of sub goals to goals are specified. During late requirements analysis, actor diagram and relationale diagram developed during early requirements is extended to include a new actor, along with the dependencies between this actor and others in the environment by using organizational modeling and decisional modeling. The output of this rationale diagram is on Chapter 4, section 4.3.1, page 101.

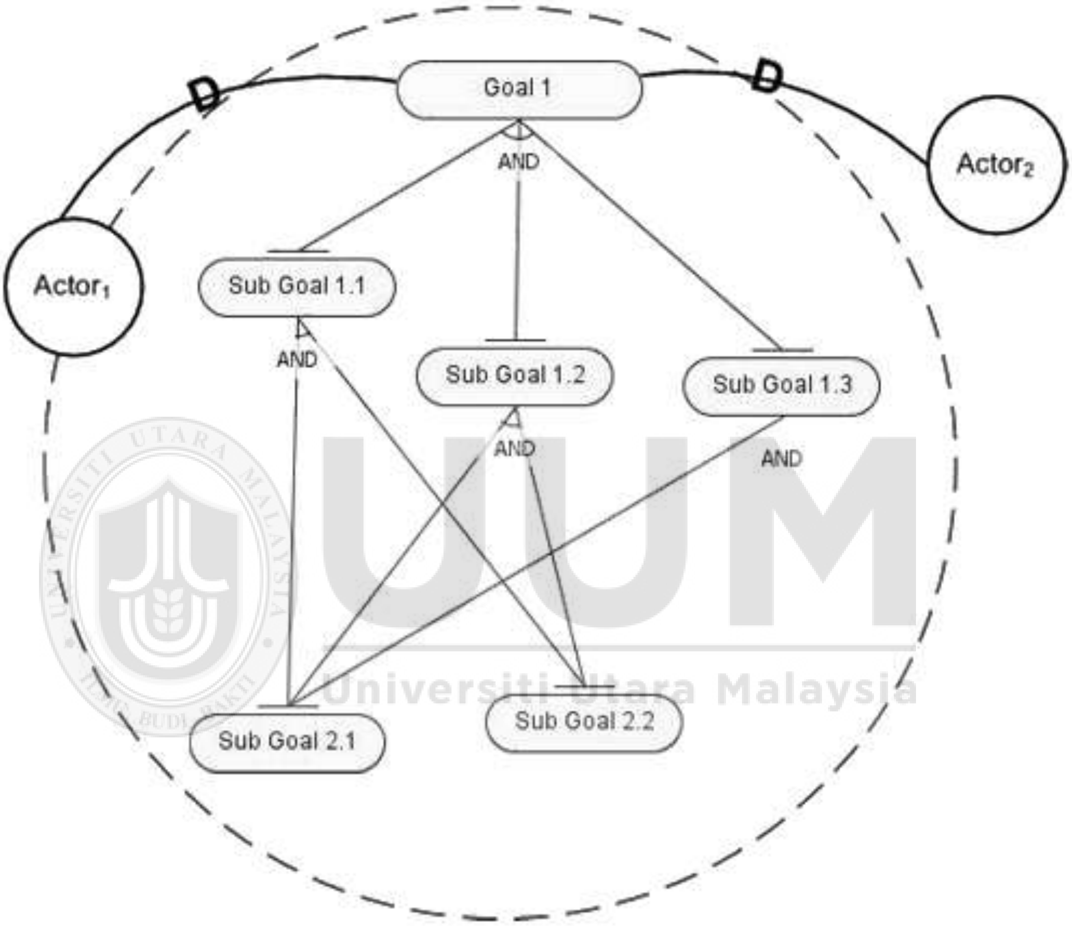


Figure 3.5: Rationale Diagram

3.4.1.2 Fact Analysis

The analysts start with the top goals going through the leaf goals to identify the relevant fact. Usually, the facts associated with the leaf goals are not considered. The rationale diagram also induces some relationship between facts such as fact merge or fact split that used in particular of facts that related to the sub goal are becoming a subject of the fact that associated with the goals. This operation of facts is useful for defining the final conceptual schema for DW. Facts analysis produces two different types of template which is (Fact, Description) for each fact and (Goal, Facts) if there are many facts associated with goals. The output of this fact analysis is on Chapter 4, section 4.3.2, page 103.

3.4.1.3 Attribute Analysis

After fact analysis, the process continues to attribute analysis which is to identify all the attributes that give a value when facts are recorded. Analyst used the extended rationale diagram from fact analysis to identify all the attributes without specifying their possible role as dimensions or measures as Figure 3.6. All the attributes are listed in the template (Attribute, Goal, and Fact). All goals, facts, and attributes are defined in the context of organizational setting. The output of this attribute analysis is on Chapter 4, section 4.3.3, page 106.

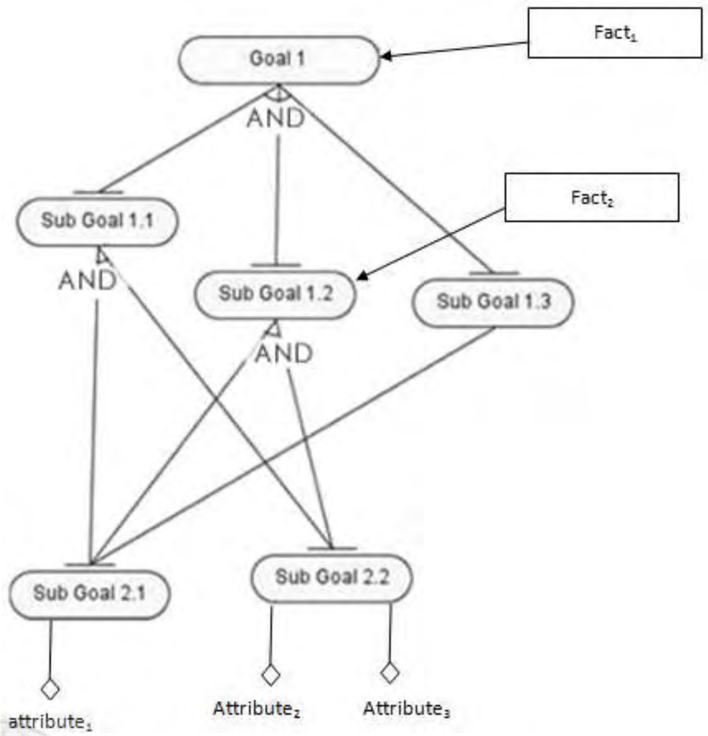


Figure 3.6: Extended Rationale Diagram for Organizational perspective

3.4.2 Decisional Modeling

Decisional modeling is focused on how the DW supports the decisional process of the organization, and focus on the requirements of the DW from the perspectives of the decision maker. The previous diagram from organizational modeling is used to support the identification of the facts that to be associated with the decision maker goals. Decisional modeling consists of four types of analyses after the decision makers are identified which is (i) goal analysis, in which produced rationale diagram, (ii) facts analysis, in which where the rationale diagrams are extended with facts, (iii) dimension analysis, in which where the rationale diagrams with the facts is extended by connecting dimensions to the goals, and (iv) measures analysis to support the decision making and produced list of facts, dimensions and measures.

3.4.2.1 Goal Analysis

Decisional modeling starts with analyzing the actor diagram for decision makers from previous Step 1. Analyst identified the decision makers and initial dependencies between them. From actor diagram, the goals that associated to each decision maker (Actor) are then decomposed to OR-decomposed and analyzed in details to produce a set of rationale diagram. The same template is used as organizational modeling to collect and organize the information. From the template, rationale diagram is designed as Figure 3.7. The output of this goal analysis is on Chapter 4, section 4.4.1, page 112.

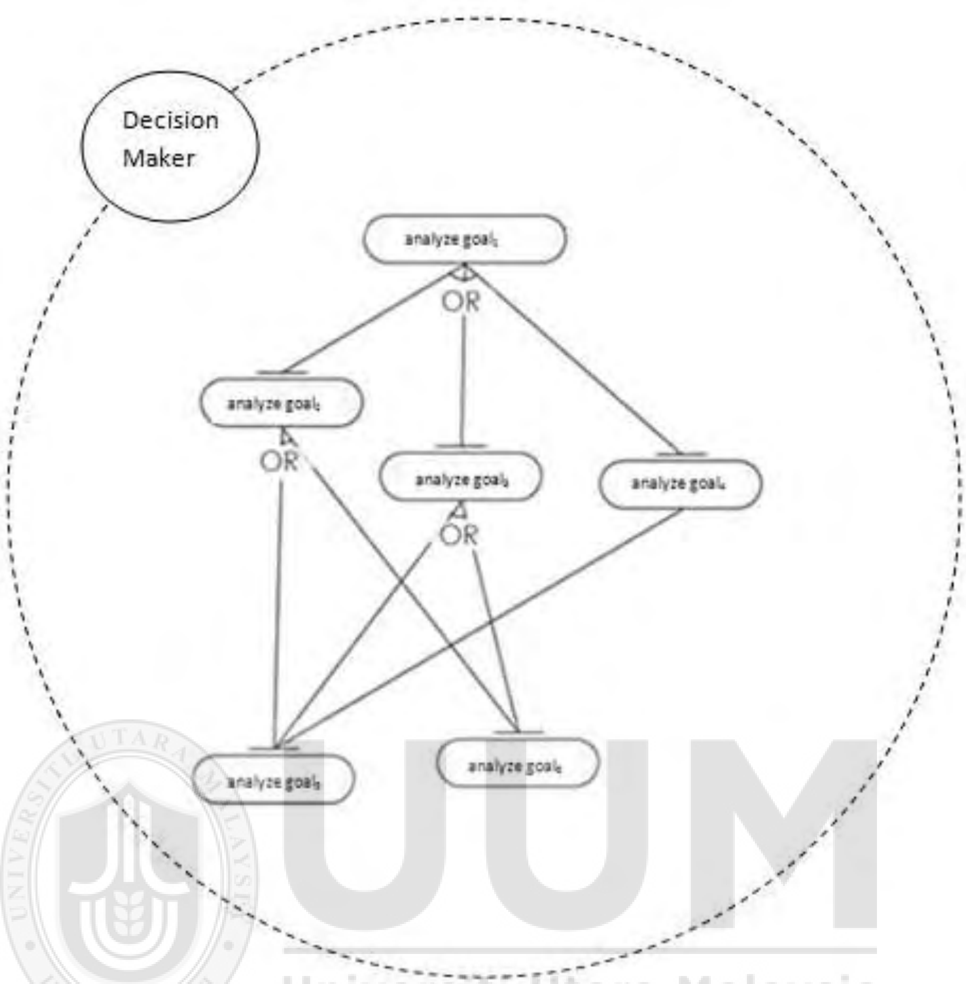


Figure 3.7: Rationale Diagram for Decision Maker from the decisional perspective

3.4.2.2 Fact Analysis

After goal analysis, the process continues with fact analysis where the relational diagram is extended with facts. In this process, analyst identified and import facts from relational diagram from previous modeling and associate to the goals of decision maker. The analyst also can produce new facts by directly analyzing the decision maker rationale diagrams. The output of this fact analysis is on Chapter 4, section 4.4.2, page 116.

3.4.2.3 Dimension Analysis

The next process is dimension analysis. In this phase, for each facts identified from fact analysis is consider necessary in order to satisfy their decisional goals. During this phase, analyst identified the dimensions from the leaf goals of the relationale diagram of decision makers and relevant facts that associated to upper level goals. Interaction between analyst and decision maker are important here to capture the possible perspective of analysis. The information for goal analysis, facts analysis and dimension analysis are collected by using template (goal, fact, dimension) and the second template are used to describe the dimensions (dimension, description). The output of this dimension analysis is on Chapter 4, section 4.4.2 Table 4.9 on page 117, and Table 4.10(i) on page 119.

3.4.2.4 Measure Analysis

Finally, after dimension analysis, the process continues with measure analysis. In this phase, analyst identified the quantitative aspect which is described numerical property of a fact that is relevant for decision making by analyzing the dimensions and facts. The measures, then associate to the facts previously identified. All goals, facts, dimensions, and measures are defined in the context of decision-maker setting. The output of this measure analysis is on Chapter 4, section 4.4.2 and Table 4.10(ii) on page 120.

3.5 Step 3: Develop Requirement Modeling for DW

Based on previous steps, the proposed approach produces two interrelated modeling activities, namely organizational modeling and decision modeling. The organizational model produced represents the main data that comprises the most relevant attributes that are part of the source database. On the other hand, the decisional model produced describes the decision maker's needs, which summarizing the role played by an actor associates to the facts, dimensions and measures. The requirements derived from organizational and decisional modeling are mapped to the conceptual schema for DW. Conceptual schema is used to help clarify ideas, facilitate communication between people from different domains, without anticipating design decisions (Stefanov, 2007). Based on GRAnD (Giorgini P et al., 2008) there are three phases involved, which are requirement mapping and hierarchy construction.

3.5.1 Requirement Mapping

Requirement mapping is a process where facts, dimensions and measures identified during decisional modeling are mapped onto entities in the source schema or conceptual schema. In this process, the Dimensional Fact Model (DFM) will be produced (Figure 3.8). DFM used to support a conceptual schema of DW. The purpose of using the DFM is to provide efficient support for communication between designer and end user to draw the user's query specification requirements that can be summarized in a draft (to create a stable platform for the target model logic).

A DFM schema of a DW consists of a set of *fact schemata.* A *fact schema* is structured as a tree whose root is a fact. A fact (*F*) is a concept relevant to decision-making processes; it models a set of events (ex: in a company: sales, shipments, purchases, etc.), It has dynamic properties or evolves in some way over time, and it has one or more numeric and continuously valued attributes which measure the fact from different points of view. A measure (*m*) is a numerical property of a fact and describes a quantitative fact aspect that is relevant to the analysis. A dimension (*d*) represents as *nodes* or a circle are a fact property with a finite domain and describes analysis axes of the fact.

3.5.2 Hierarchy Construction

The dimension in which a hierarchy is rooted defines its finest aggregation granularity; the other dimension attributes (*d-a*) define progressively coarser granularities. The hierarchy construction is a process where a basic conceptual schema from the requirement mapping process is generated by navigating the source schema. Hierarchies are sub-trees rooted in dimensions. Hierarchies are made up of discrete dimension attributes linked by many-to-one relationships, and it determines how facts may be aggregated and selected significantly for decision-making processes.

A hierarchy of the dimension (d_n) will probably include the dimension attributes (d-a_n) and represent as *nodes* or circles. Hierarchies may also include non-dimension attributes (*non d-a*). A non-dimension attribute contains additional information about a dimension attribute of the hierarchy, and is connected by a many-to-one relationship (e.g., the department address); unlike dimension attributes, non-dimension attribute cannot be used for aggregation.

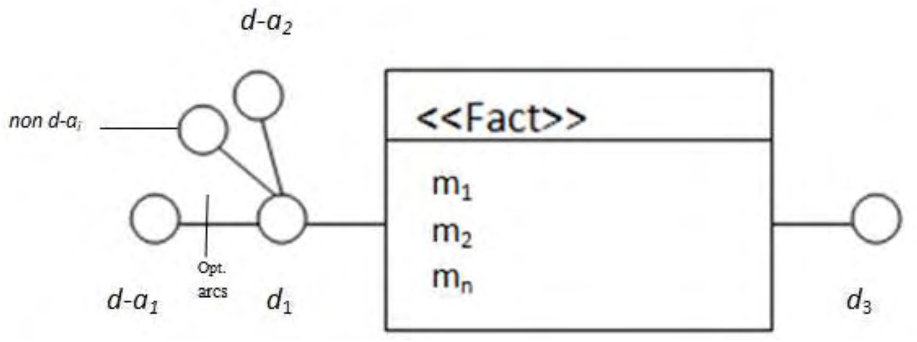


Figure 3.8: Preliminary Fact schemata for <<Fact>>

For each Fact, *F*, that identified in the decision model are mapped onto relation *F* in the source schema with the many-to-one associations. The relation expressed in the source schema by foreign keys (FK) is iteratively navigated. It starts from *F*, to build the attribute hierarchies and create the basic conceptual schema in form of Fact Schema. Relation, R, represent as *arcs* to show relationships between pairs of attributes.

Different facts are represented in different fact schemata. Two fact schemata are compatible if both facts, share at least one dimension attribute. Both facts may be overlapped to create a new fact schema. Without conflict between attribute dependencies in two facts, the set of fact attributes in a new fact schema is the union of the sets in two current facts. For example, H, F, and G is a different fact, the set of the fact attributes in H is the union of the sets in F and G. Then, the dimensions in new fact are interaction of those in two current facts; at least one attribute is shared. For example, dimension in H are intersect for those F and G fact. And each hierarchy in new fact includes all and only the dimension attributes included in the corresponding hierarchies of both current facts. For example, each hierarchy in H includes all and only the dimension attributes included in the corresponding hierarchies of both F and G.

3.5.3 The mapping process to conceptual schema

The mapping process starts from extended rationale diagrams produced by organizational modeling and decisional modeling. The goal of the source conceptual schema is to know what data is available for the DW. There are three phases involved which are requirement mapping and hierarchy construction based on GRAnD (Giorgini P et al., 2008). The process mapping will use the required notation as below Table 3.1. The process requirement mapping started like Figure 3.9. Figure 3.9 shows the flow chart of the mapping process to conceptual schema. The output of this mapping is on Chapter 4, section 4.5, Figure 4.13 on page 126, and Figure 4.14 on page 127.

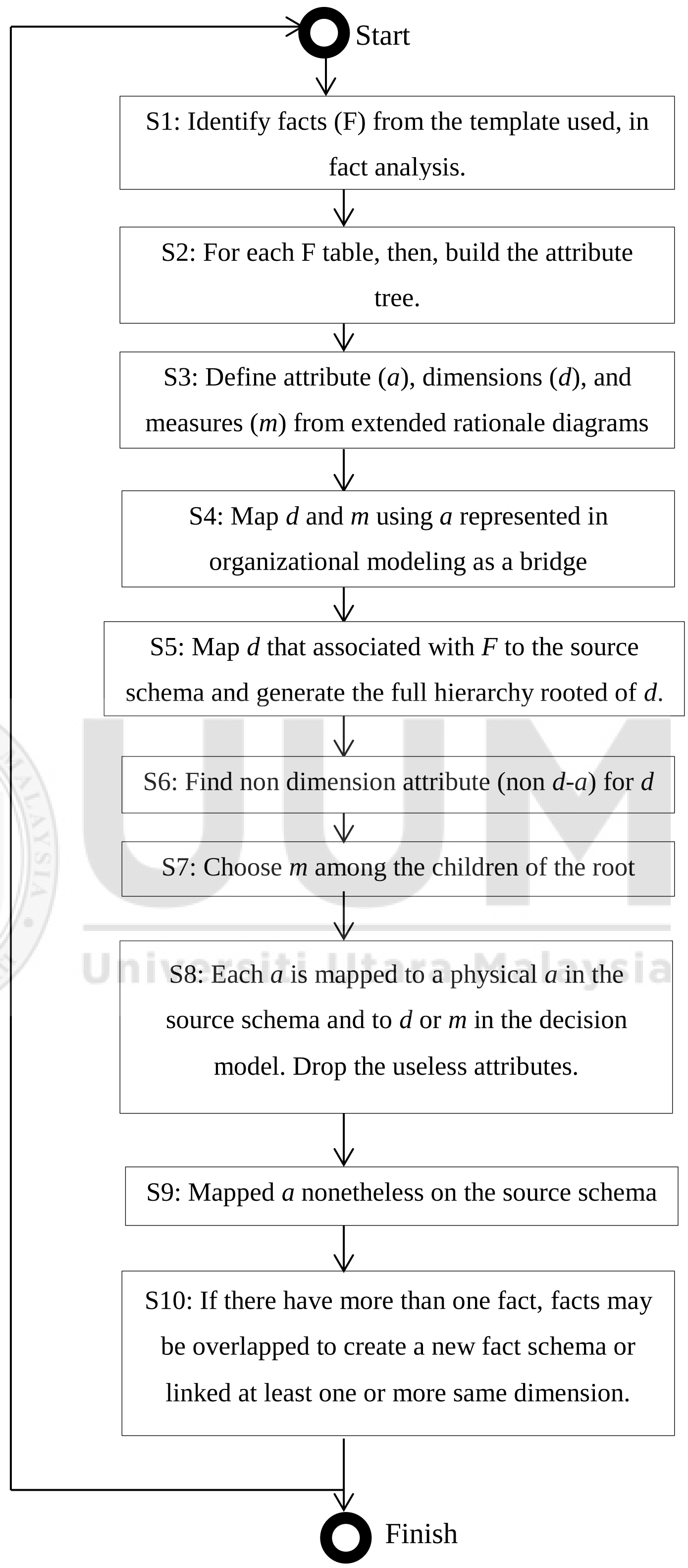


Figure 3.9: The flow chart of the mapping process to conceptual schema.

Step 1: Identify facts from the template used, in fact analysis. A fact graphically represented as a rectangle in the decision model. For each fact, navigate the functional dependencies. Fact (*F*) is represented by a table in the conceptual schema. For every fact defined becomes the root of an attribute tree.

Step 2: For each F table, then, build the attribute tree. Each node or circle of the attribute tree corresponds to one or more attribute in the decision model. The root of the attribute tree corresponds to the primary key (PK) of *F*.

Step 3: Then, define attribute (*a*), dimensions (*d*), and measures (*m*) from extended rationale diagrams produced by decision modeling and organizational modeling. A dimension is a fact property that describes a possible coordinates of analysis and represented as a small circle that connected to a goal. Meanwhile, measure is a numerical property of facts that describe a quantitative aspect for decision making and represented as small squares that connected to the goals.

Step 4: Map *d* and *m* using *a* represented in organizational modeling as a bridge (Figure 3.10).

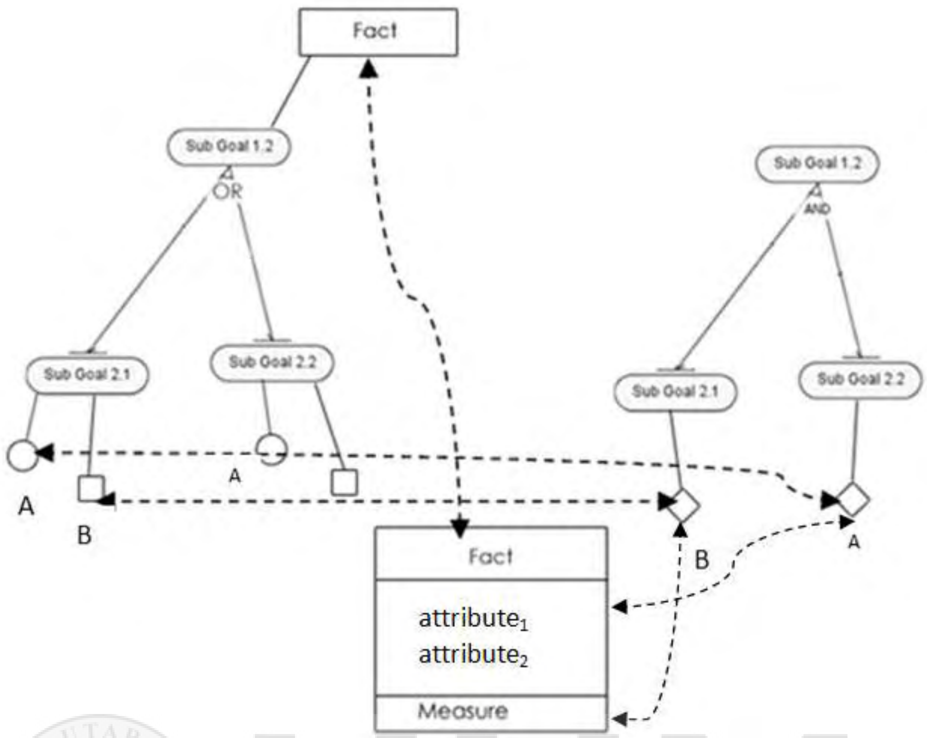


Figure 3.10: Mapping from the decision model (left) and organizational model (right) to the conceptual schema

Step 5: Map *d* that associated with *F* to the source schema and generate the full hierarchy rooted of *d*. Multiple values of d also can be related to a single instance of the F.

Step 6: Then, find non dimension attribute (non *d*-*a*) for *d*.

Step 7: Choose *m* among the children of the root. Every *m* that associated with a goal related to *F* and successfully mapped to conceptual model is included, but there no hierarchy is generated for it. A fact may have no measure. An attribute cannot be both a measure and a dimension.

Step 8: Each *a* is mapped to a physical *a* in the source schema and to *d* or *m* in the decision model. Drop the useless attributes.

Step 9: Attributes in the organizational model that not mapped into dimensions or measures in the decision model are mapped nonetheless on the source schema. Attributes are the field whose value is provided when a fact is recorded to fulfill the goal and represented as a small diamond that connected to the goals.

Step 10: If there have more than one fact, facts may be overlapped to create a new fact schema or linked at least one or more same dimension. The mapping process finish once the conceptual schema is build and normalize.

Table 3.1 List of Goal and conceptual notation

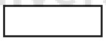
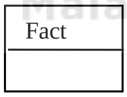
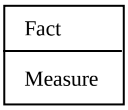
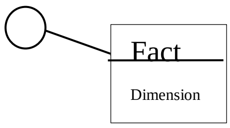
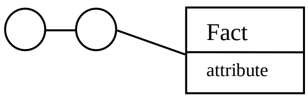
No.	Goal	Goal Notation	Conceptual Notation
1	Fact		
2	Measure, *m*		
3	Dimension, *d*		
4	Attribute, *a*		

Figure 3.11 is an example of preliminary fact schema obtained for fact registration once the process mapping is finish. Consistently with the DFM, the fact is represented as a box containing the measures; dimensions are circles connected to the fact; hierarchies are represented as trees rooted in dimensions. This figure shows that for fact registration, the dimensions are customer and title. Gender, category, length, director and main Actor are a dimension attribute. The hierarchy rooted in title has been built by navigating the source schema.

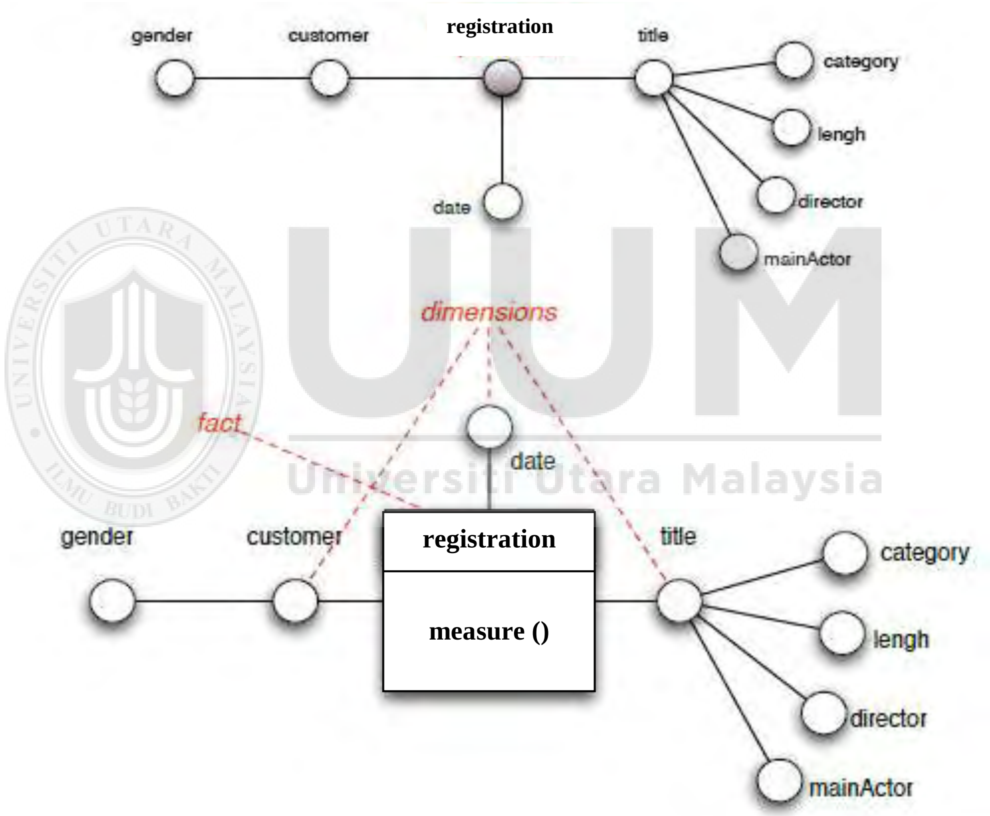


Figure 3.11: Example of Fact Schema

3.5.4 The mapping process from conceptual schema to UML Class Diagram

Next process is a mapping process from the conceptual schema to UML class diagram. The process uses the previous conceptual model of DW produced and map to UML class diagram that represents the relational model of the DW. The analysis class diagram contains classes that represent the more permanent aspects of the DW.

Class diagram uses UML classes and associations. The usual symbol for a class is a box with three compartments. The top compartment contains the class name which means the name of the entity is in the upper box. The second compartment contains the attribute names. Attributes are the essential description of the data belongs to the class. And the third compartment will contain the operations. The primary key candidates are underlined. In the class diagram data types must be defined for each attribute before implement database at this point.

In a conceptual schema shown in Figure 3.12, the facts and dimensions are all composed of UML classes. The measure attributes and other descriptive attributes are defined in the fact class. The relationship between the fact class and the root class of each dimension is described using a UML association, and the multiplicities of the two associated class roles are 0 .. * And 1 respectively. The cardinality of fact class role is 0 .. * (any), indicating that each dimension object relates to zero, one, or more fact object instances. The minimum cardinality of dimension class roles is 1, indicating that a fact objects relates to all dimensions. The relationship between a lower level ($level_i$, i>0) class in a dimension hierarchy and a higher level ($level_{i+1}$, i>0) class in the same hierarchy is described using a UML aggregation, and the multiplicity of aggregated class roles is 1.. * and 1 respectively as shown in Figure 3.13.

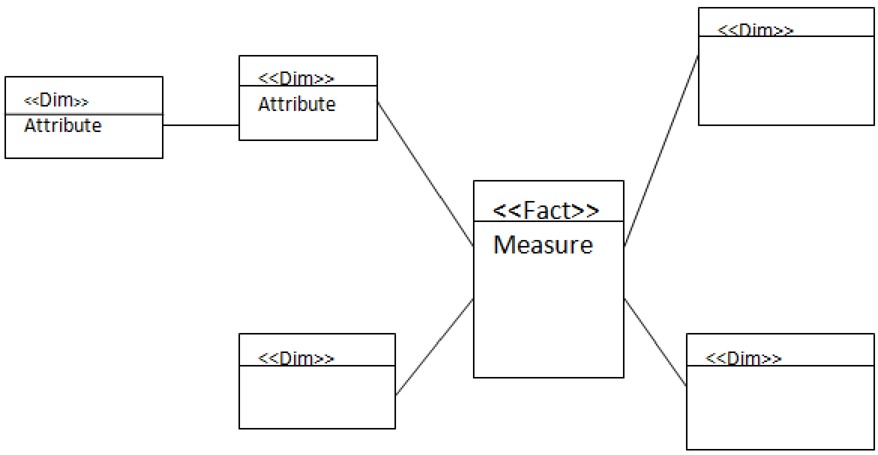


Figure 3.12: An example of Conceptual Schema

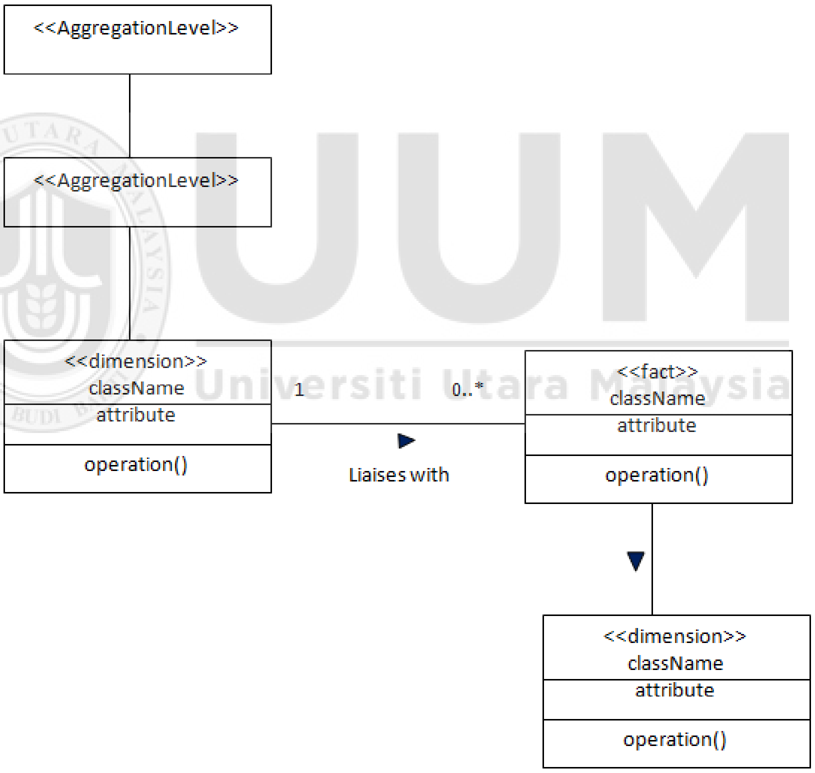


Figure 3.13: An example of conceptual schema in G-UML notation

Figure 3.14 shows the concept of process mapping from the conceptual schema to G-UML class diagram. The process starts with the conceptual schema from the previous process. The conceptual schema is then mapped to class diagram by using the G-UML notation in the Table 3.1. The mapping starts like Figure 3.14. Figure 3.14 shows the flow chart of the mapping process from conceptual schema to UML Class Diagram. The output of this mapping is on Chapter 4, section 4.6 Figure 4.17 on page 134.

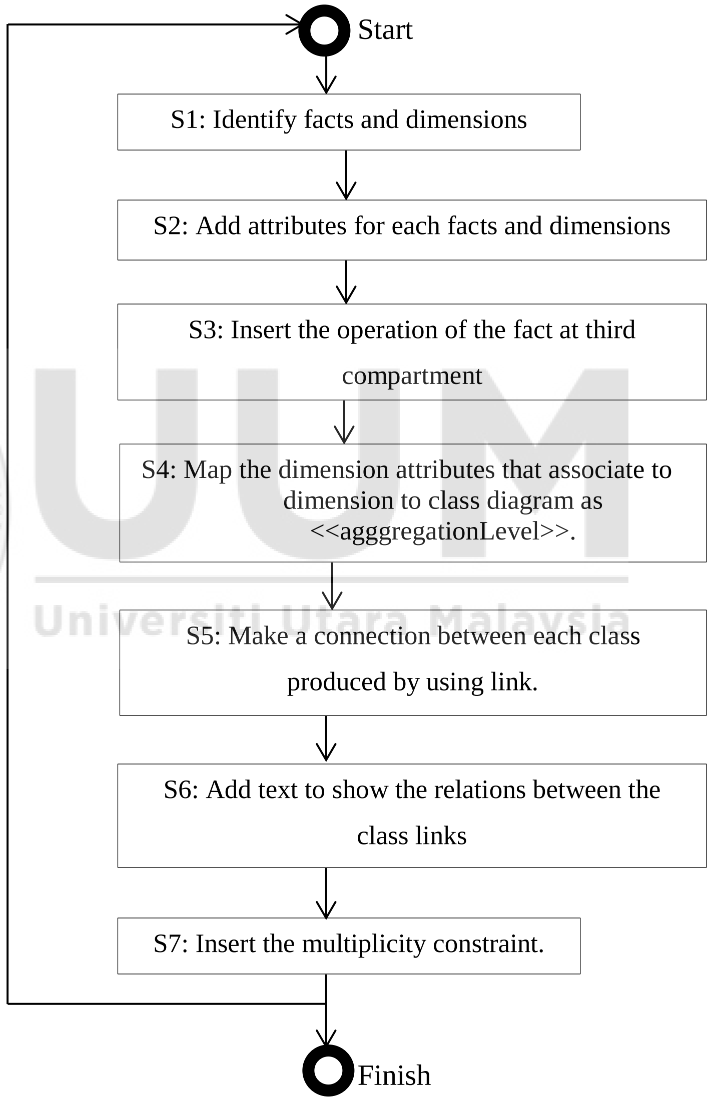


Figure 3.16: The flow chart of the mapping process from conceptual schema to UML Class Diagram

Step 1:	Identify facts and dimensions. The facts and dimensions will be mapped onto entities in the UML class diagram. Facts are graphically represented as a rectangle or a table that connected to the dimensions, represented as a node or a circle, map to entity (class) in UML is drawn as boxes (Figure 3.15).
Step 2:	Then, add attributes for each facts and dimensions to the entity (class) respectively in second compartment beginning with a lower-case letter (Figure 3.15).

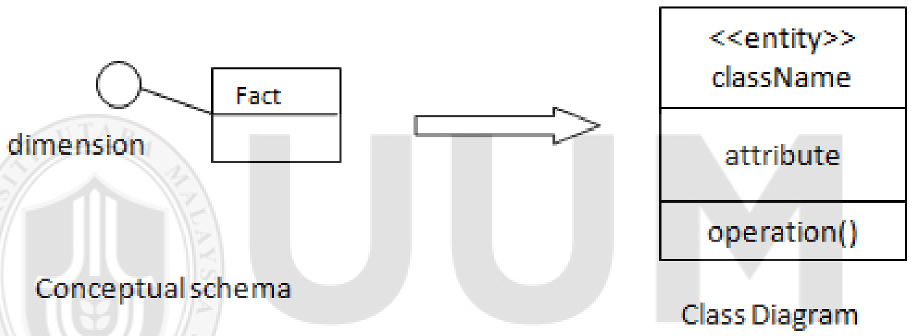


Figure 3.15: Process mapping from fact and dimension from conceptual schema to class diagram

Step 3:	Then, insert the operation of the fact at third compartment. The measure is an operation in the class diagram. As for attributes, operation names are written beginning with a lower-case letter (Figure 3.15).
Step 4:	Map the dimension attributes that associate to dimension to class diagram as <<agggregationLevel>>.
Step 5:	Make a connection between each class produced by using link. A link is a logical connection between two or more objects.

Step 6: Add text to show the relations between the class links. The text at the association end gives a name to the role that the instances of the class at the end of the association (Figure 3.16).

Step 7: Insert the multiplicity constraint.

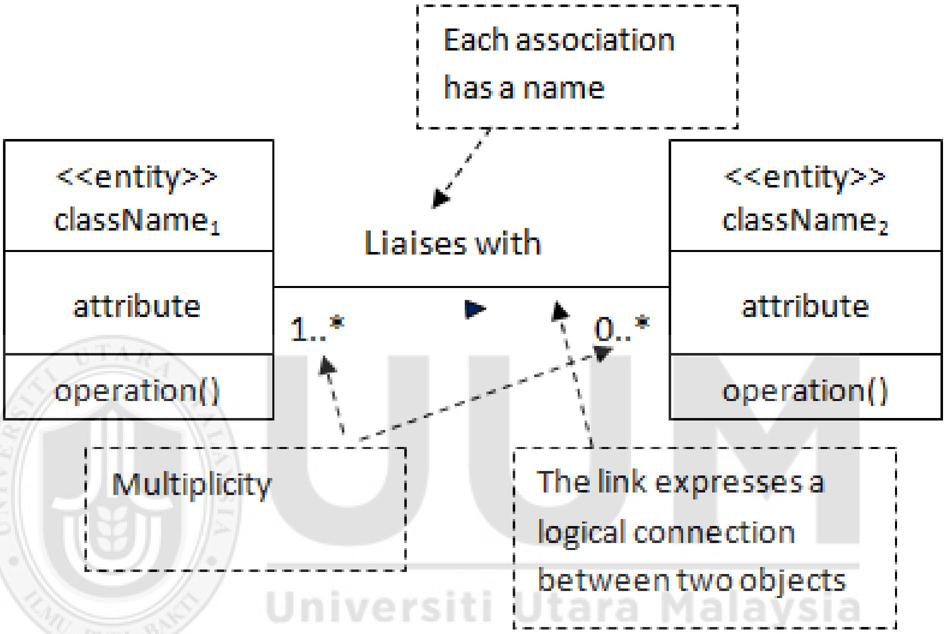
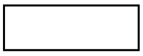
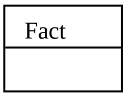
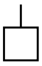
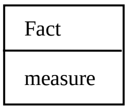
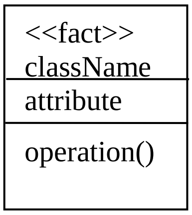
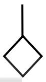
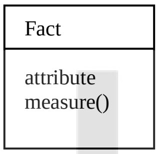
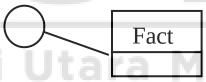
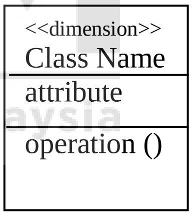
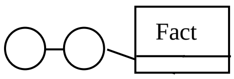
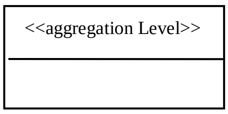


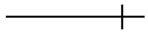
Figure 3.16: An example of class diagram

There are different reasons for selecting the right mapping, but there exist some mappings that occur based on general relational database approaches. Table 3.2 shows the notation involved in the mapping classes from goal to conceptual modeling and from conceptual modeling to UML modeling.

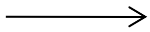
Table 3.2: List of Goal notation and UML notation in process mapping

No	Goal	Goal Notation	Conceptual Notation	UML Notation
1	Fact			
2	Measure			
3	Attribute			
4	Dimension		 dimension	
5	Dimension attribute		 *d-a*	
6	Relation	Dependency 	Association	Association

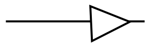
Decomposition



Contribution



Mean-ends



3.6 Step 4: Requirement Model Verification

DW model verification is performed by using case study in context of Malaysia health care which is rural health care. These verification processes support the third objective of this research. The main purpose of verification process is to examine the correctness of the proposed DW requirement modeling. The verification process consists of three parts. First part required expert review to design DW model using conceptual schema and compare the design with DW model produced by G-UML model.

The comparison is based on fact test (Golfarelli & Rizzi, 2009). Fact test is used to verify the correctness of the user requirement expressed during requirement analysis is actually supported by the conceptual schema. The fact test can be achieved by checking, for each question reviews provided (Table 3.3), how many questions can be answered based on the design produced to the required measures that have been included in the fact schema and aggregation level that can be expressed as valid grouping set on the fact schema (Golfarelli & Rizzi, 2009).

Table 3.3: List of Question Reviews

No	Question Reviews
1.	**Population Structure** i. Total population in [year] for certain age. ii. Total number of patient register
2.	**Life Expectancy at Birth** i. Birth life estimation for male and female respectively in [year] for each state
3.	**Morbidity** The health status of a community which is focus on the incidence or prevalence of disease. i. Total number of patient antenatal get hypertension ii. Total number of emergency case reported iii. Total number of patient antenatal get diabetic
4.	**Health Facilities** Define, i. Total number of community clinics ii. Total number of health clinics iii. Total number of maternal and child clinics
5.	**Post Information** Define**,** i. Total number of medical officer ii. Total number of dental officer iii. Total number of pharmacist iv. Total number of nurse v. Total number of medical assistant

6. **Promotion Matters**

MOH usually provide many campaign and promotion to help patients maintain healthy by following campaigns arranged and scheduled by the clinic and hospital.

Define,

i. Total number of campaign per year

ii. Total number of campaign by service category

7. **Treatment Services**

There are many services provided by clinics.

Define,

i. Total number of type of treatment provided to patient diabetic per year

ii. Total number of appointment has been arranged and scheduled

iii. Total number of incidence case trend for Rural Health care

8. **Pharmacy Services**

The pharmaceutical services are a mainstay of primary health care services

i. Percentage of prescription wrongly filled and detected before dispensing

---

The second part of verification process is based on expert verification instrument. The expert review instrument is used as a guideline for the experts' to verify the DW model produced is correct based on seven requirements metrics that provide all the requirements need from the stakeholders and help decision-maker in decision making processes. The metrics are: Explicit Hierarchies in dimensions, Symmetric treatment of dimensions and measures, Multiple Hierarchy in each dimension, Support for correct aggregation, Non-strict hierarchy, Many-to-many relationships between facts and dimensions, handling uncertainty. According to Pederson and Jesen (1998), proposed DW model must meet the listed metrics for validity. The responses are collected through items with 5-point Likert scale of 1 (Not satisfactory), 2 (Fairly Satisfactory), 3 (Neutral), 4 (Satisfactory) and 5 (Very Satisfactory). Table 3.16 presents the expert verification metrics.

Table 3.4: Expert Verification metrics

No.	Metrics	Description
1	Explicit Hierarchies in dimensions	The relation of hierarchies in the dimensions should be captured explicitly in the data model, so then can identify the different level of hierarchy in dimension
2	Symmetric treatment of dimensions and measures	The data model should allow the fact measures to be treated as dimensions.
3	Multiple Hierarchy in each dimension	In one dimension, there can be more than one path along which to aggregate data
4	Support for correct aggregation	The data model can produce the result according to the condition of requirements need when aggregating data. One aspect of this is to avoid

		double-counting of data.
5	Non-strict hierarchy	The data model shows the hierarchies in a dimension are not always strict; it may have many-to-many relationships between the different levels in a dimension.
6	Many-to-many relationships between facts and dimensions	The data model show the relationship between fact and dimension is not always the classical many-to-one mapping.
7	Handling uncertainty	The data model should allow grouping a few requirements in a result.

The last part of verification process is requirements verification by expert review from Health Care domain. The verification required the expert review to answer the question provided (Appendix D) to verify the late requirements produced by proposed DW requirement modeling. The questions are based on the question reviews (Table 3.3) and the late requirements. The responses are collected through items with 5-point Likert scale of 1 (Not satisfactory), 2 (Fairly Satisfactory), 3 (Neutral), 4 (Satisfactory) and 5 (Very Satisfactory).

The finding of DW model verification process by expert reviews is further discussed in Chapter 5. Also the design produced by expert review in verification process is further discussed in Section 5.2. After the verification is conducted, the results are analyzed.

3.7 Tool for Goal-oriented modeling

This research utilizes three types of tools which is Organization Modeling Environment (OME), DW Design Tools for Goal-oriented and MySQL Workbench for UML, which can support both the high-level requirements process artifacts. First tools are OME. OME is goal-oriented modeling tools that provide the developer with a graphical user interface to develop models, and supports access to a knowledge base that allows for advanced model analysis. Second tools are DW Design Tools (as shown in Figure 3.17, 3.18, and 3.19). DW Design Tools are a DW modeling tools used for modeling and designing the DW systems from the organizational modeling toward the decisional modeling. The modeling and analysis tasks produce the DW schemata used for designing the DW system accordingly (Giorgini et al., 2008). DW-Tool provides goal analysis; fact analysis and dimension analysis function to design organizational modeling.

DW-Tool also provide goal analysis, fact analysis, dimension analysis and measure analysis function to design decisional modeling. The MySQL Workbench for UML helps to design UML Class Diagram with standard UML notation.

Figure 3.17: DW-Tool for goal-oriented approach

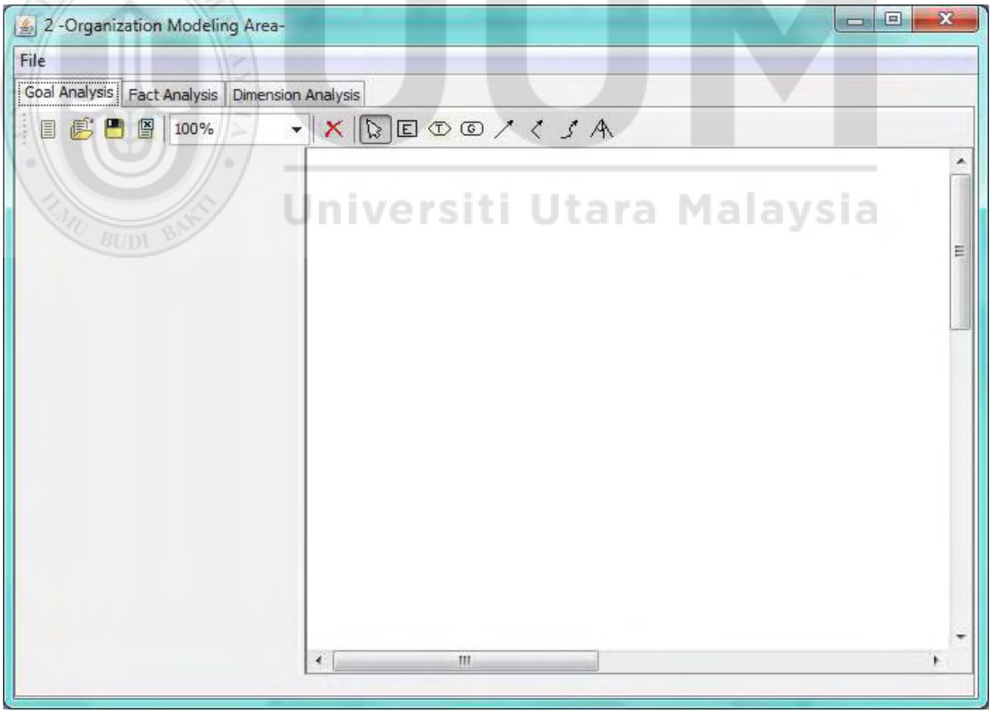


Figure 3.18: DW-Tool for organizational modeling

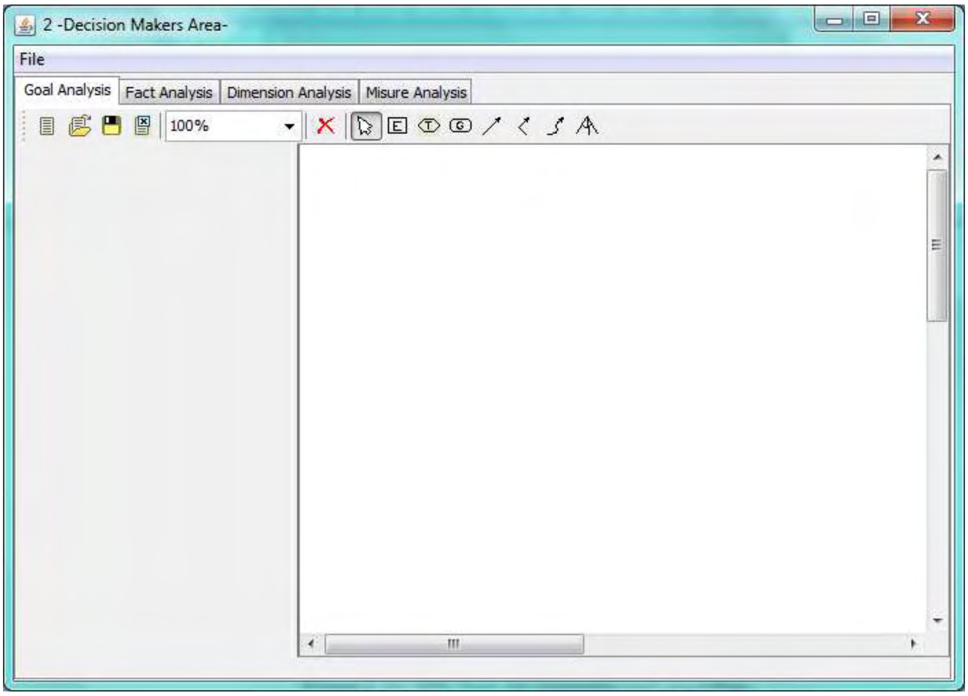


Figure 3.19: DW-Tool for decisional modeling

3.8 Summary

This chapter explains the process of developing requirement model involved based on G-UML model. The processes consist of three steps which is start with requirement elicitation, requirement modeling by using organizational modeling and decisional modeling, and requirement mapping process from conceptual schema to UML Class Diagram. In requirement elicitation, requirements and intentions from stakeholders and decision-maker are identified and analyzed qualitative research method including background reading, semi-structured interview, observation and document sampling. List of requirements is produced from this phase.

Then, the process requirement modeling begins with designing organizational modeling and decisional modeling based on result in requirement analysis. The output produced from this process is actor diagram and relationale diagram based on organizational perspective and decision-maker perspective. In requirement mapping process, the modeling produced in requirement modeling is map from conceptual schema to UML Class Diagram. All the processes development uses Rural Health Care case study and DW Tools and OME. Based on the G-UML model presented in Figure 3.2, Table 3.5 highlights the result after implementation G-UML modeling approach to design requirements of DW modeling is further discussed in Chapter Four. Each objective is supported by Tropos methodology, research method and activities to produce output and results.

Table 3.5: The G-UML Model tasks

Tropos Methodology	Research Method	Activities	Output (Page No.)
Objective 1: To identify the early and late requirements for designing DW systems			
Phase 1: Early Requirement Phase	Requirement Elicitation	Identify and analysis all the requirements gathered	• List of Early Requirements for DW for Health Care

Phase 2: Late Requirement Phase	Requirement Modeling	Organizational Modeling:- -Goal Analysis -Fact Analysis -Attribute Analysis	DW Requirement Modeling based on Goal-oriented Approach:- -Actor diagram -Relationale diagram -Extended Relationale diagram
		Decisional Modeling:- -Goal Analysis -Fact Analysis -Dimension Analysis -Measure Analysis	
Objective 2: To develop DW requirement model by using G-UML modeling approach.			
Phase 3: Architecture Design Phase Phase 4: Details Design Phase	Develop DW Requirement Model using Goal –UML Approach	Design DW Conceptual Schema and mapping to UML Class Diagram	• Conceptual Schema of Health Care • UML Class Diagram of Health Care
Objective 3: To verify the correctness of DW requirement model developed.			
Phase 5: Implementation Phase	Requirement Model Verification	Expert Review Verification	• List of Late Requirements for DW for Health Care

The verification of requirement modeling is further discussed in Chapter 5 by using expert reviews based on Health Care case study. However, for this research, the implementation phase is excluded the process of DW development for health care case study.

CHAPTER FOUR
G-UML MODEL FOR REQUIREMENT ANALYSIS

4.1 Introduction

This chapter describes a method and techniques used that applied to Rural Health Care. The method is used as a guideline to conduct the research and fulfill all the requirements needs in order to achieve the objective of the research based on previous chapters. Previous chapter had discussed details the method and techniques used for this research.

4.2 Results of Requirement Elicitation

Based on the real case study which is the Malaysian Health Care focused on Rural Health Care, the system used in hospital and clinic is not integrated with each other. This causes the management of information and decision-making between medical practitioners and staffs has a difficult time to organize and utilize these volumes of patient's data effectively and efficiently. This indicates the necessity for implementing the DW concept in this institution's health care as a solution to integrate the system in order to collect and distribute the health data.

The elicitation process is based on qualitative research method for data collection such as semi-structured interviews, document sampling and observation. All of these requirements are analysed to identify what is the main goal of organization, task to be done, and resource to be fulfilled. The semi-structured interview medical practitioners and staffs such as medical officers and nurses at Klinik Kesihatan Mempaga, Klinik Kesihatan Bandar Alor Setar, staff nurse at Klinik Kesihatan Tunjang and Klinik Kesihatan Gua Musang and study on current system documentations, sample data and websites for details information. The result for requirement modeling is listed in a table. The analyst elicited and analysed user requirements because of its potential to recognize the current organizational situation and disclose the business goals of the functional and non-functional requirements.

i. Background Reading

Background reading start with finding the documentations of Malaysian Health Care to get sources and information, including website MOH, hospitals and clinics reports, organization charts, policy manuals, medical practitioner job description, annual reports, documentation of existing system, journals, and websites that consistent to business organizations.

ii. Interview

The interview session has been conducted with medical practitioners and staffs such as medical officers and nurses at Klinik Kesihatan Mempaga, Klinik Kesihatan Bandar Alor Setar, staff nurse at Klinik Kesihatan Tunjang and Klinik Kesihatan Gua Musang. As a guideline for the interviewer, there are four questions have been prepared. The questions are as below. The analyst analyses all the answers are listed in Table 4.1a.

1. What is your position and area do you cover the clinic?
2. What are the tasks do you performed?
3. What is the action do you take to implement the tasks?
4. What are the data do you need to support your decisions that you made during implementation of the tasks?

Table 4.1a shows the analysis of data recorded during interviewed session and the conclusion of each position. The data divided into four columns. The first column is position and area covered by respondents, second is tasks, third is action taken or decision that they had to take, and four is data required to be recorded.

Table 4.1a: Interview result

Position	**Task**	**Action Taken**	**Data Record**
Medical Officer -OPD Department -MCH Department	1. Provide general health services to patients. 2. Identify certain cases to be referred and follow up by specialist at hospital. 3. Call officer or other medical officer at hospital or other clinic to	- listen to patient complaint and record in treatment card. - review past medical history - review past medicine history - assign to take test such as blood test and urine test - decide for next	Details Patient record, Medicine record, Patient history record, Result test record, Treatment record.

	collect patient history record for further action. 4. Make a reference letter to follow up at hospital. 5. Prepare the routine report of medical description. 6. Prepare the medicine description to patient.	appointment date for follow up - assign the action supposed to do for follow up	
Staff nurse	1. Assist medical officer in treatment. 2.Assist medical officer in patient record 3. Provide test for patient 4. Visit patient at their home (baby and mother) 5. Register and arrange patient for next appointment 6. Escort patient to	- listen to patient complaint - review past medical history for next appointment - test sample blood test and urine in the lab - arrange and schedule for next appointment date for follow up - take an action for patient test request by medical officer	Details Patient record, Medicine record, Patient history record, Result test record, Treatment record, Test record.

	hospital 7. Assist in lab test to test the sample given by patient such as blood, urine. 8. Provide result for the test.	- sterilize treatment equipment	
Staff Administrative / Assistant Nurse	1. Register patient 2. Record payment 3. Ask patient disease	- produce receipt for payment - record patient details - provide treatment card to medical officer	Details Patient record, Patient history record, Bill and finance record
Pharmacist	1. Provide medicine to patient based on medicine description provided by medical officer 2. Help patient in medicine daily taken 3. Record medicine	- select medicine - record in/out medicine - construct patient for daily taken	Medicine record

Based on the result interview from Table 4.1a, the data was analyzed to design DW requirements that focused on OPD and MCH Services by using actor diagram and rationale diagram in organizational modeling and decisional modeling in the next subsection.

iii. Observation

Observation in hospital and clinic procedures help analyst to provide a better understanding of the flow of the medical practitioner job process, what the information people use to carry out their job, the document uses to record the data, and how well the current system handles their needs. Observation allows the analyst to follow the entire process through the start till the end. Figure 3.3 shows an example of the standard operation procedures that clinics used to treat the patients.

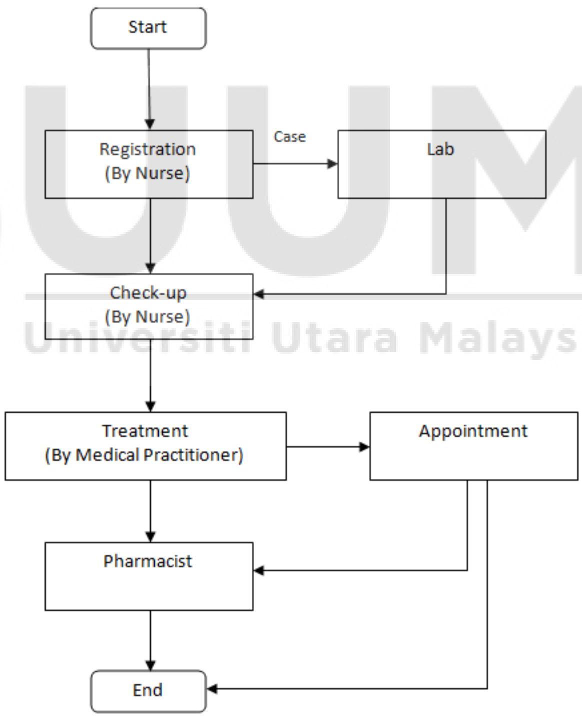


Figure 4.1: An example of Flow of standard operation procedure used by clinics

iv. Document Sampling

The second type of data collection is document sampling. For clinic that used traditional methods to record patient data, this document is important as a reference when the patient need to be refer or follow up at hospital. During the interview session, respondents show how the patient data was recorded by using document provided by MOH. The documents used is listed as below as Table 4.1b and Appendix C.

Table 4.1b: Document Sampling

No	Document's Name	Description of Document
1.	Treatment card -Perubatan 96-Pin. 1/78)	Used to record details OPD patients, treatment given, test need to do by patient, medicine, and appointment.
2.	Treatment book - KIK/1(a)/96(Pind.2012) - KIK/1(b)/96	Used to record maternity health.
3.	Treatment book	Used to record baby health.

4.3 Results of Requirement Modeling

Organizational modeling represents the main data in the organization and comprise most relevant attribute to data source. Organizational modeling consists of three types of analysis which are goal analysis, fact analysis and attribute analysis.

4.3.1 Goal Analysis

In Rural Health Care, goal analysis used to represent the intentions of stakeholders for the organization and their social dependencies. In this step, requirements from data collections such as background reading, interview, observation and documents sampling are analyzed by using goal analysis to produced actor diagram and rationale diagram according to the goals need to be achieved by the stakeholder and decision maker. Then, based on the diagrams produced, the analyst navigates the rationale diagram of each actor and extends it by associating goals with the facts. Facts analysis is carried out according to top-down approach.

i. Actor Diagram

Figure 4.2 shows the actor diagram for Malaysian Health Care. It starts with the high level of goal from actor Ministry of Health (MOH). The goal of MOH is depended on the State of Health Department and District of Health Office to be achieved. Then, proceed to decomposing the goal into sub-goals that needs to be achievable by State of Health Department and District of Health Office as an actor. Table 4.2 shows the list of main actors that support their main objectives to be achieved.

Table 4.2: List of Main Actors

Main Actor

Actor	Objectives
Ministry	• Provide Quality of Life of an Advanced Nation • Facilitate & support People for High Quality of Health
Hospital	• Provide High Quality of Health System
Clinic	• Provide Public Health

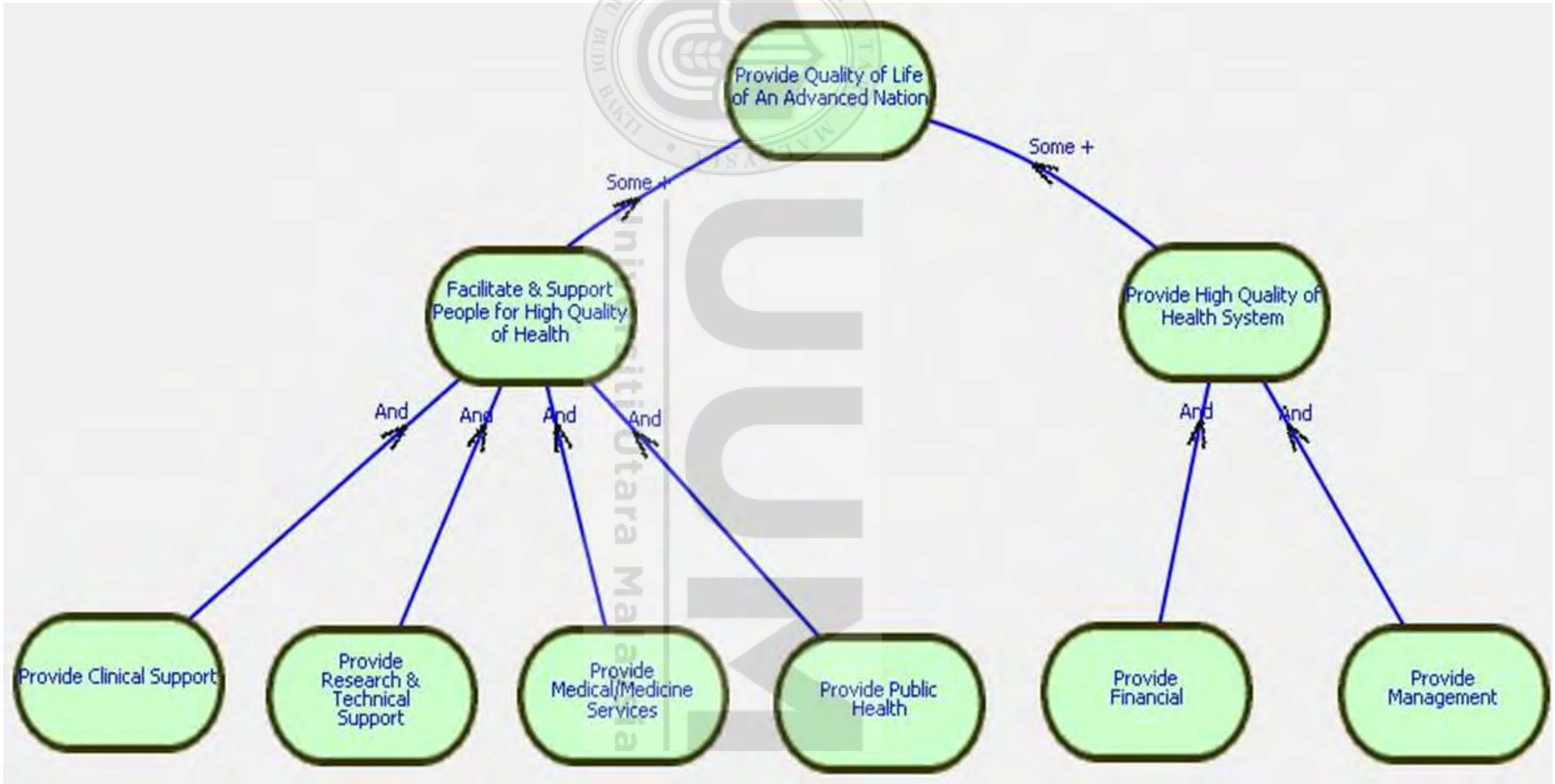


Figure 4.2: An Actor Diagram for Malaysian Health Care

Figure 4.3 shows an actor diagram for Rural Health Care. This diagram shows a partial actor diagram for Malaysian Health Care case study. The Hospital depends on Clinic to achieve the goal provide Public Health and the Patient to achieve the goal manage In-Patient. Meanwhile, the actor Clinic depends on Patient to achieve goal manage Out-Patient. Two different types of templates are provided, which are *Sub-actor* (Sub-actor, Type, Goals) as Table 4.3 and *Dependencies* (Depender, Dependee, Goal) as Table 4.4.

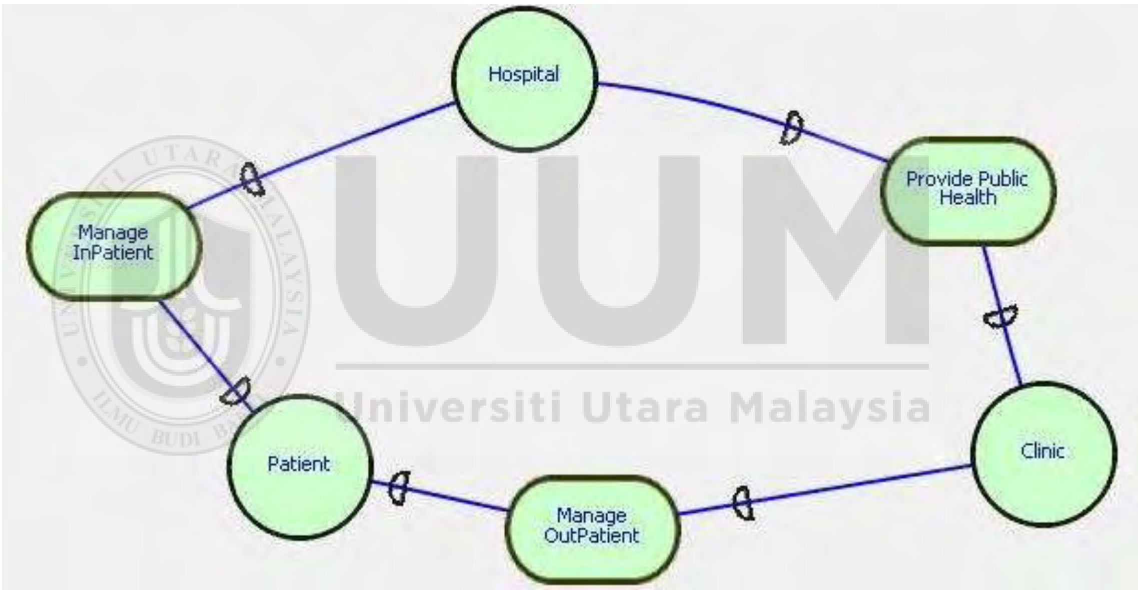


Figure 4.3: Actor Diagram for Rural Health Care

Table 4.3: List of Sub-Actors

Sub-actor

Sub-actor	**Type**	**Goals**
Hopsital	State Health Department	• Manage Inpatient
Clinic	District	• Manage Outpatient

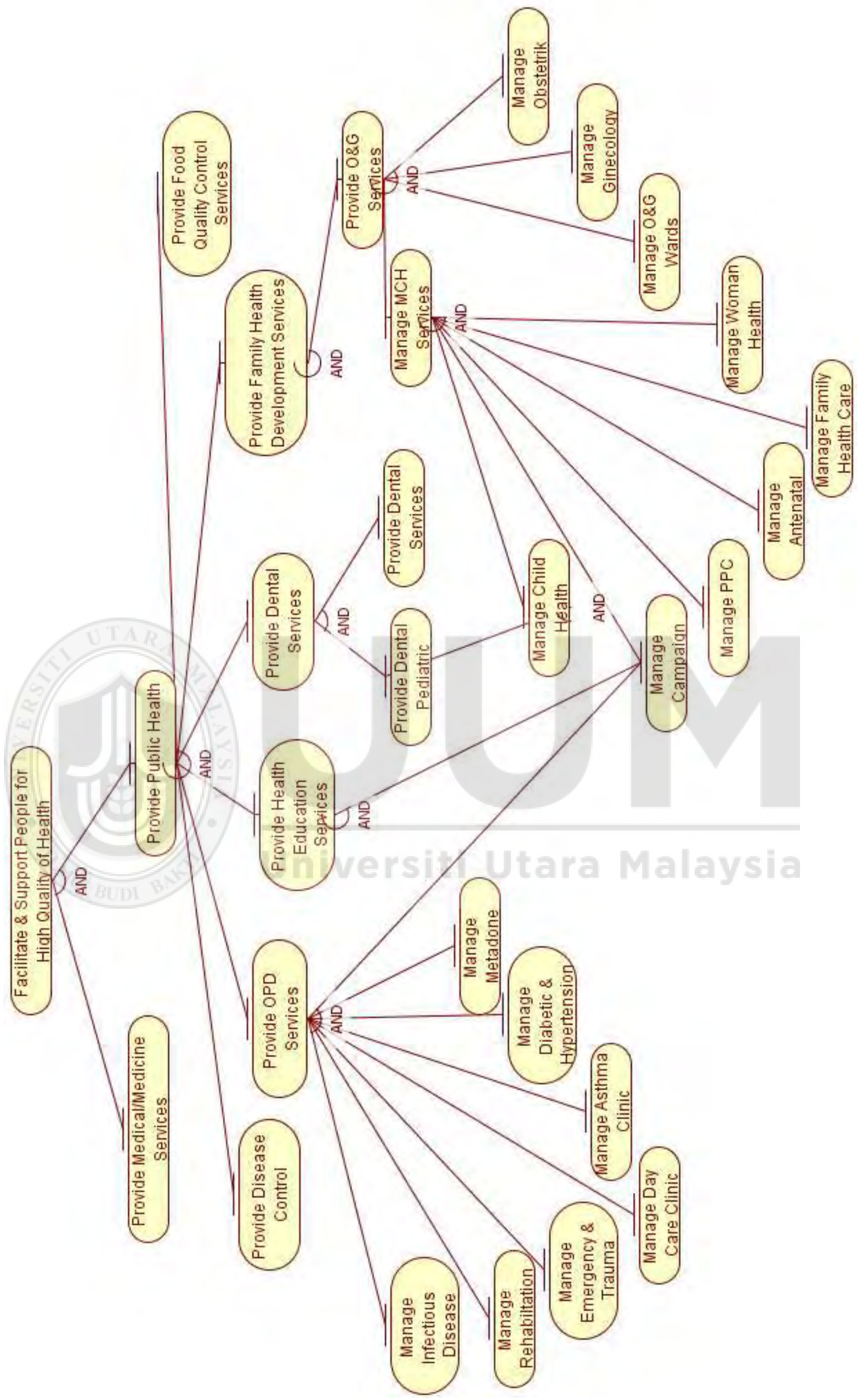
Table 4.4: List of Dependencies

Dependencies		
Depender	**Dependee**	**Goal**
Hospital	Clinic	• Provide Public Health
Hospital	Patient	• Manage Inpatient
Clinic	Patient	• Manage Outpatient

ii. Rationale Diagram

The second step produced rationale diagram for each actor in actor diagram for Rural Health Care. Goals are AND-decomposed to sub goals and contribution link between goals are discovered. The analysis ends when all the goals of each actor have been analyzed and all the dependencies among actors are established in actor and rationale diagrams. As for the first step, the information are collected and organized using a predefined template (see Table 4.2, Table 4.3, and Table 4.4). Figure 4.4 shows the partial of rationale diagram for Public Health sub-goal from Rural Health Care for organizational modeling.

The diagram focused on Hospital and Clinic as an actor. The goal provide Public Health is decomposed into a few sub-goals which are provide Education Services, provide Disease Control, provide OPD Services, provide Dental Services, provide Family Health Development Services and provide Food Quality Control Services. There are new dependencies that discovered at this point, for example, provide OPD Services and provide Family Health Development Services depend on Clinic actor to manage OPD and MCH services.


Facilitate & Support People for High Quality of Health
AND
Provide Medical/Medicine Services
Provide Public Health
Provide Disease Control
Provide OPD Services
Provide Health Education Services
Provide Dental Services
Provide Family Health Development Services
Provide Food Quality Control Services
Provide O&G Services
Manage MCH Services
Manage Obstetrik
Manage Ginecology
Manage O&G Wards
Manage Woman Health
Manage Family Health Care
Manage Antenatal
Manage PPC
Manage Child Health
Manage Campaign
Provide Dental Pediatric
Provide Dental Services
Manage Metadone
Manage Diabetic & Hypertension
Manage Asthma Clinic
Manage Day Care Clinic
Manage Emergency & Trauma
Manage Rehabilitation
Manage Infectious Disease

4.3.2 Fact Analysis

The previous rationale diagram from goal analysis is used to identify all the relevant facts for the organization. Fact focuses on a set of event in the business world and aimed to provide the information to the stakeholder and decision maker. At this phase of view, the analyst navigates the rationale diagram for each actor and extends it by associating goals with facts. Figure 4.5 shows the fact OPD Services is associated to goal Provide OPD Services. Meanwhile, Figure 3.11 shows the fact MCH Services is associated to goal Manage MCH Services. The information is collected using two different types of template which are (Fact, Description) as Table 4.5 and (Goal, Fact) as Table 4.6.

Table 4.5: List of Facts and Description

Fact	**Description**
OPD Services	Provide OPD Services to Patient at Rural Health Care
MCH Services	Manage MCH Services to Patient at Rural Health Care
Registration	Record outpatient details
Antenatal	Record mother and child record

Table 4.6: List of Goals and Facts

Goal	**Fact**
Provide OPD Services	OPD Services
Manage MCH Services	MCH Services

Figure 4.5 shows the extended rationale diagram for OPD Services of Rural Health Care in organizational modeling for Medical Practitioner and Out-Patient actor. The fact OPD Services is associated with the main goal provides OPD Services. Then, by analyzing the sub-goals of the goal provide OPD Services, introduced the a few attributes that is associated with each sub-goals (Figure 4.5).

Meanwhile, Figure 4.6 shows the extended rationale diagram for MCH Services of Rural Health Care in organizational modeling for Medical Practitioner and Out-Patient actor. The fact MCH Services is associated to the main goal manage MCH Services. Attributes that is associated to each sub-goals is defined in attribute analysis.

4.3.3 Attribute Analysis

Starting from previous diagram (Figure 4.5 and 4.6), the entire related attributes for the sub-goals is explored. The attributes are simply data that associated to the goals. All the attributes that are given a value is identified when facts are recorded such as Figure 4.7. Figure 4.7 shows the attributes that exist for fact Registration. The information is collected by using table (Attribute, Sub-Goal, Goal, and Fact) as Table 4.7 and Table 4.8 for diagram Figure 4.5 and Figure 4.6.

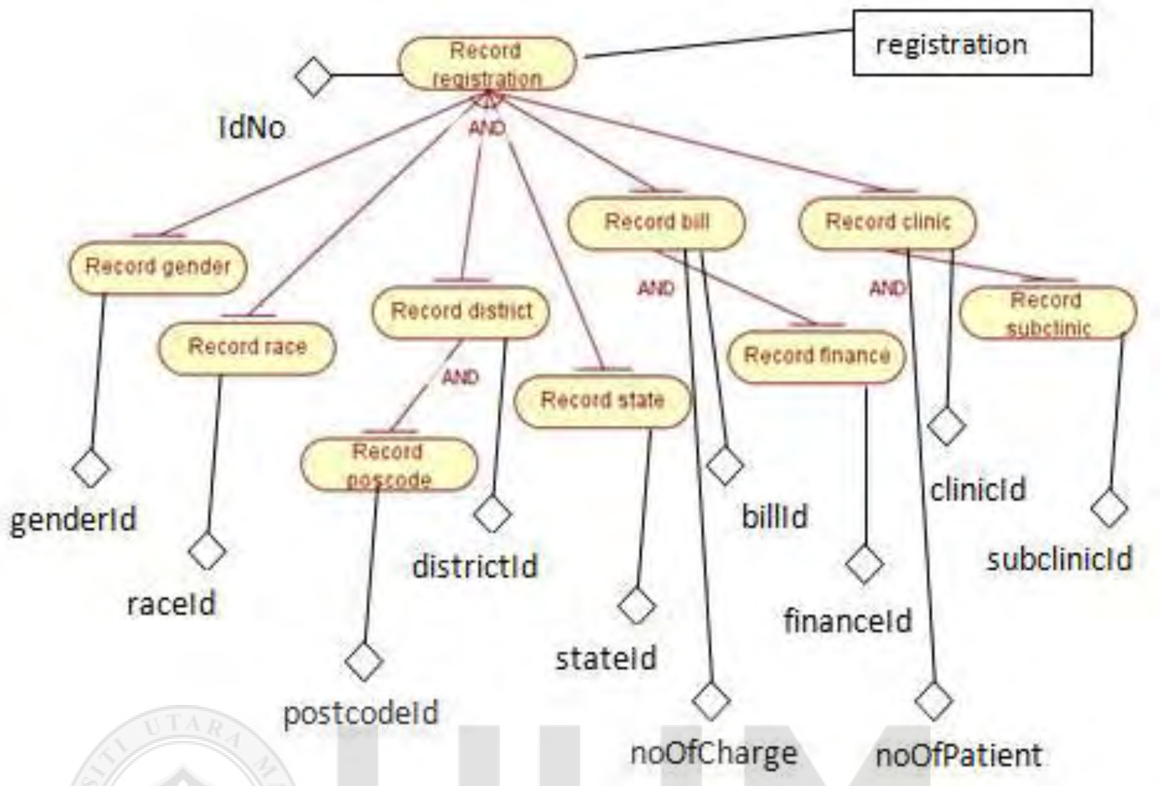


Figure 4.7: Extended Rationale Diagram for Organizational perspective for attribute analysis by using Fact Registration

Table 4.7: List of Attributes for Goals OPD Services

Attribute	Sub-Goal	Goal	Fact
- campaignId - NoOfCampaign	- Analyze campaign	Manage Campaign	OPD Services
- idDid - tti_Id - ti_Id	- Analyze infectious_disease - Analyze type_of_test_infection - Analyze type_of_infectious	Manage Infectious Disease	OPD Services

- rehabilitationId - rtypeId	- Analyze rehabilitation - Analyze type_of_rehabilitation	Manage Rehabilitation	OPD Services
- erId - typeoferId - erlevelId - diagnosistId	- Analyze emergency&trauma - Analyze type_or_er - Analyze of_er_level - Analyze diagnosist	Manage Emergency & Trauma	OPD Services
- diabeticId - dhlevelId - htestId	- Analyze diabetic - Analyze diabetic_level - Analyze diabetic_test	Manage Diabetic & Hypertension	OPD Services
- treatmentId - t_treatmentId - NoOfTreatment	- Analyze treatment - Analyze type_of_treatment	Manage Day Care Clinic Manage Asthma Clinic	OPD Services
- idNo - genderId - raceId - districtId - stated - billId - clinicId - NoOfPatient	- Analyze registration - Analyze gender - Analyze race - Analyze district - Analyze state - Analyze bill - Analyze clinic		registration

- NoOfCharge			
- appId - typeofappId - NoOfAppointment	- Analyze appointment - Analyze type-of_appointment		OPD Services
- staffId	- Analyze staff		OPD Services
- medicineId - pharmasistId	- Analyze medicine - Analyze pharmacist		OPD Services
- hypertensionId - hlevelId - htestId	- Analyze hypertension - Analyze hypertension_level - Analyze hypertension_test	Manage Hypertension	OPD Services
- prId	- Analyze patient_reference		OPD Services
- diagnosistId - charted	-Analyze diagnosist		OPD Services

Table 4.8: List of Attributes for Goals MCH Services

Attribute	Sub-Goal	Goal	Fact
- campaignId - NoOfCampaign	- Analyze campaign	Manage Campaign	MCH Services
- staffed	- Analyze staff		MCH Services
- birthdRecordId - IdNo - bpId - juandisId - appId - isId - antenatalId - complicationId - giId	- Analyze child_health - Analyze birth - Analyze general_inspection - Analyze sibling_information - Analyze complication - Analyze juandis	Manage Child Health	
- treatmentId - t_treatmentId - antenatalId - NoOfTreatment	- Analyze treatment - Analyze type_of_treatment - Analyze antanatal	Manage PPC	MCH Services
- idNo - genderId - raceId - districtId - stated - billId - clinicId	- Analyze registration - Analyze gender - Analyze race - Analyze district - Analyze state - Analyze bill - Analyze clinic		registration

- NoOfPatient - NoOfCharge - rfId	- Analyze riskFctor		
- appId - typeofappId - NoOfAppointment	- Analyze appointment - Analyze type-of_appointment		MCH Services
- diagnosistId - charted	-Analyze diagnosist		MCH Services
- medicineId - pharmasistId - tomId	- Analyze medicine - Analyze pharmacist - Analyze type_of_medicine		MCH Services
- prId	- Analyze patient_reference		MCH Services
- antenatalId - idNo - bpId - atId - mfhId - cpId - ac - daId - bpmonitoringId - mgttId - usId	- Analyze antenatal	Manage Antenatal care	Antenatal

- fmcId
- cpcId
- bsId
- birthRecordId
- pmId
- cbId
- pmcId
- heId
- mhrId

4.4 Results of Decisional Modeling using Goal-UML

The previous diagram from organizational modeling is used to support the identification of the facts that to be associated with the decision maker goals. Decisional modeling consists of four types of analyses after the decision makers are identified which is (i) goal analysis, in which produced rationale diagram, (ii) facts analysis, in which where the rationale diagrams are extended with facts, (iii) dimension analysis, in which where the rationale diagrams with the facts is extended by connecting dimensions to the goals, and (iv) measures analysis to support the decision making and produced list of facts, dimensions and measures.

4.4.1 Goal Analysis

In this analysis phase, goal analysis starts with analyzing the actor diagram for the decision maker. Decision maker is identified and initial dependencies between them were established. The goals are decomposed to produce rationale diagrams.

Figure 4.8 shows the rationale diagram for decision maker Medical Practitioner, focusing on goal of analyze OPD Services. The goal analyze OPD Services is OR – decomposed into analyze Campaign, analyze Infection Disease Treatment, analyze Rehabilitation Treatment, analyze Emergency & Trauma Treatment, analyze Diabetic & Hypertension Treatment, analyze Day Care Treatment, analyze Asthma Clinic Treatment and analyze Methadone, which in turn are further decomposed. For example, the goal analyze Infections Disease Treatment is OR-decomposed into analyze Registration, analyze Description Medicine, analyze Type of Treatment, analyze Diagnosis Result and analyze Appointment.

Figure 4.9 shows the rationale diagram for decision maker Medical Practitioner, focusing on goal of analyze MCH Services. The goal analyze MCH Services is OR – decomposed into analyze Campaign, analyze Child Health, analyze PPC, analyze Antenatal Care, analyze Family Health Care, and analyze Woman Health, which in turn are further decomposed. For example, the goal analyze Child Health is OR-decomposed into analyze Registration, analyze Description Medicine, analyze Type of Treatment, analyze Diagnosis Result and analyze Appointment.

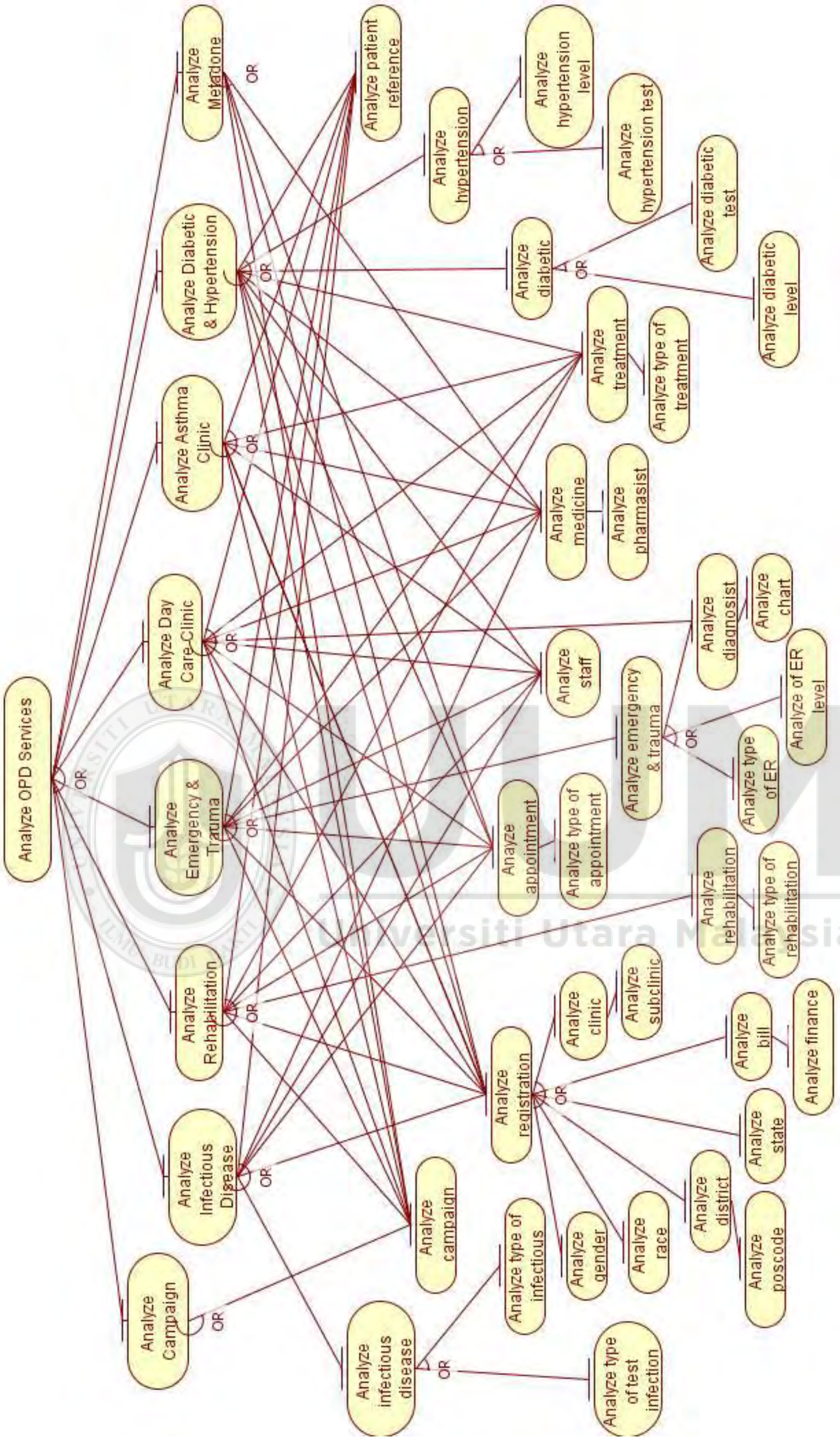
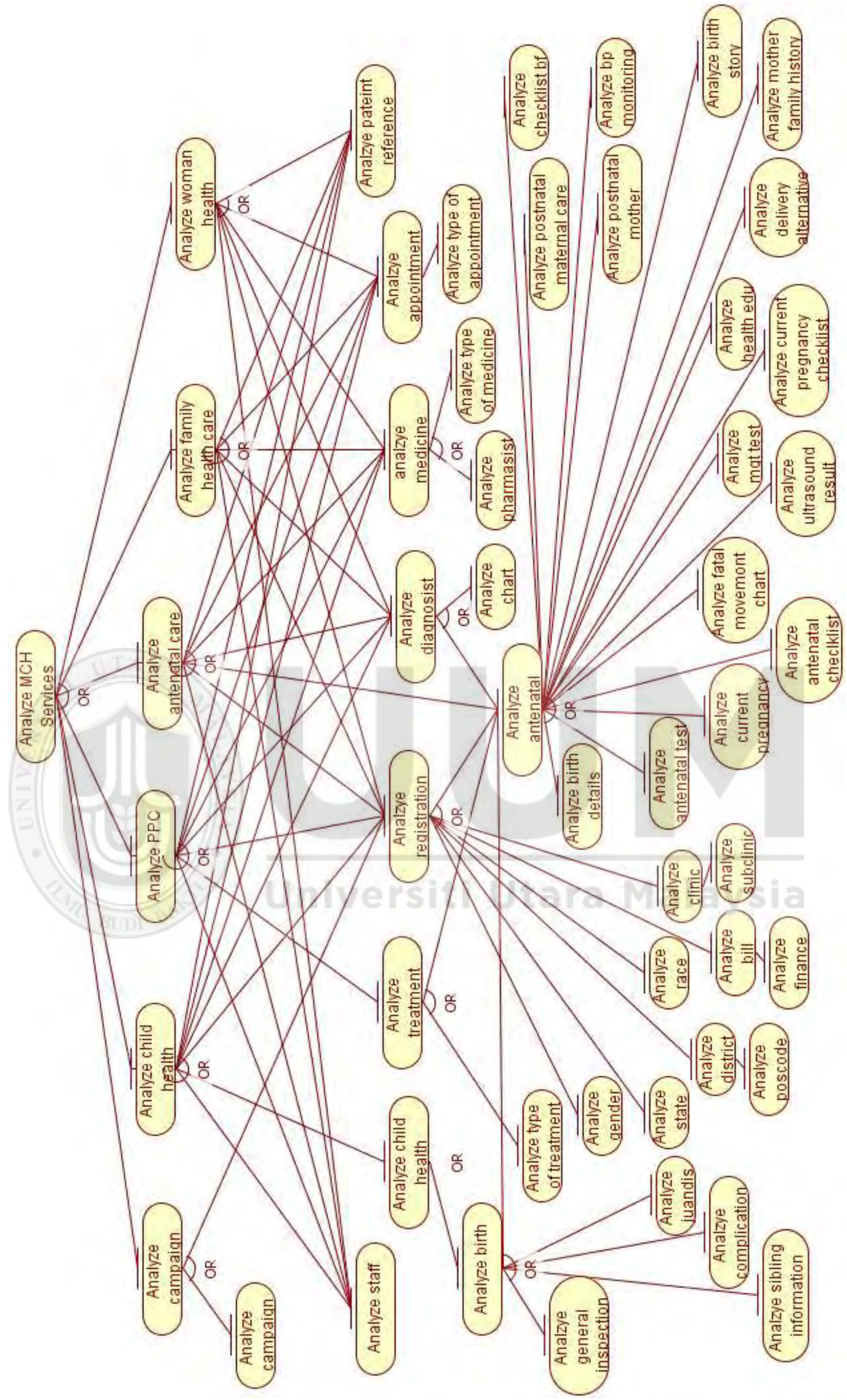

Analyze OPD Services
OR
Analyze Metadone
Analyze Diabetic & Hypertension
Analyze Asthma Clinic
Analyze Day Care Clinic
Analyze Emergency & Trauma
Analyze Rehabilitation
Analyze Infectious Disease
Analyze Campaign
Analyze patient reference
Analyze hypertension
Analyze hypertension level
Analyze hypertension test
Analyze diabetic
Analyze diabetic test
Analyze diabetic level
Analyze treatment
Analyze type of treatment
Analyze medicine
Analyze pharmasist
Analyze diagnosist
Analyze chart
Analyze staff
Analyze emergency & trauma
Analyze of ER level
Analyze type of ER
Analyze appointment
Analyze type of appointment
Analyze rehabilitation
Analyze type of rehabilitation
Analyze clinic
Analyze subclinic
Analyze bill
Analyze finance
Analyze registration
Analyze state
Analyze district
Analyze poscode
Analyze race
Analyze gender
Analyze campaign
Analyze type of infectious
Analyze infectious disease
Analyze type of test infection

Figure 4.8: Rationale di


Analyze MCH Services
OR
Analyze woman health
Analyze family health care
Analyze antenatal care
Analyze PPC
Analyze child health
Analyze campaign
Analyze campaign
Analyze staff
Analzye pateint reference
Analzye appointment
Analyze type of appointment
analyze medicine
Analyze type of medicine
Analyze pharmasist
Analyze diagnosist
Analyze chart
Analzye registration
Analyze treatment
Analyze child health
Analyze birth
Analyze antenatal
Analyze checklist bf
Analyze postnatal maternal care
Analyze bp monitoring
Analyze postnatal mother
Analyze birth story
Analyze mother family history
Analyze delivery alternative
Analyze health edu
Analyze current pregnancy checklist
Analyze mgt test
Analyze ultrasound result
Analyze fatal movemont chart
Analyze antenatal checlklist
Analyze current pregnancy
Analyze antenatal test
Analyze birth details
Analyze clinic
Analyze subclinic
Analyze bill
Analyze finance
Analyze race
Analyze district
Analyze poscode
Analyze state
Analyze gender
Analyze type of treatment
Analyze juandis
Analyze complication
Analzye sibling information
Analyze general inspection

4.4.2 Fact Analysis, Dimension Analysis and Measure Analysis

Figure 4.10 shows extended rationale diagram for medical practitioner as decision maker in OPD Services and MCH Services for fact registration from decisional perspective. The analyst associates fact registration from OPD Services and MCH services, identified during organizational modeling to goal analyzes OPD Services. Then, dimensions gender, race, district, postcode, state, bill, finance, clinic and sub clinic are connected to the goals associated to the fact and two measure is identified for goal analyze bill and clinic. Meanwhile, fact MCH Services associates to goal analyzes MCH Services. Table 4.9 shows the information for goal analysis, facts analysis and dimension analysis are collected by using template (goal, fact, dimension) and the second template are used to describe the dimensions (dimension, description) and (Fact, measure) Table 4.10.

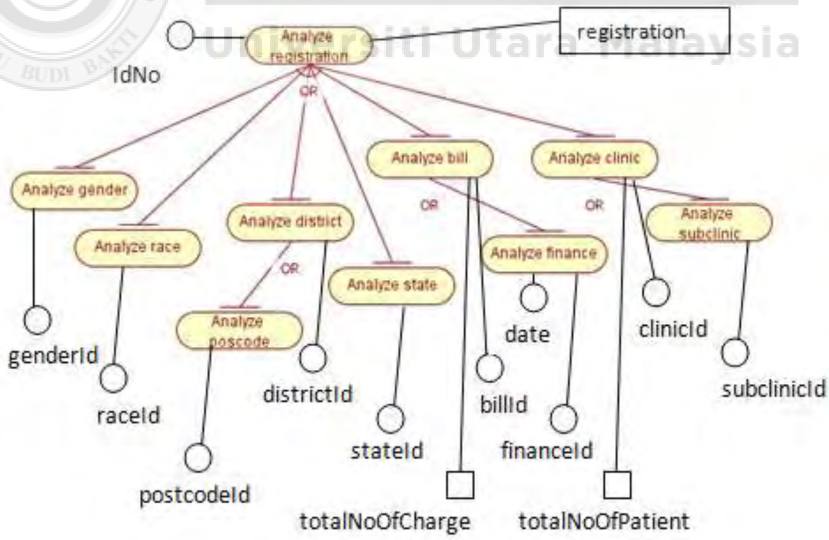


Figure 4.10 Extended Rationale Diagram for Medical Practitioner Decision Maker: OPD Services from Decisional Perspective

Table 4.9: List of Goal, Fact and Dimension for OPD and MCH Services

Goal	**Fact**	**Dimensions**	**Dimension Attributes**
Provide OPD Services	opd_Services	gender	poscode
	mch_Services	race	finance
	antenatal	district	subclinic
Manage MCH Services	registration	clinic	type_of_infectious
		bill	type_of_testinfectious
		state	type_of_rehab
		riskfactor	type_of_er
		infectiousDisease	erlevel
		rehabilitation	diabeticlevel
		emergencyTrauma	diabetictest
		diabetic	hypertensionlevel
		hypertension	hypertensiontest
		dignosist	Chart
		appointment	type_of_appointment
		treatment	type_of_treatment
		medicine	pharmasist
		Staff	type_of_campaign
		campaign	
		patientReference	

antenatal	registration
	birthDetails
	antenataltest
	motherfamilyhistory
	currentpregnancy
	antenatalchecklist
	deliveryalternative
	bpmonitoring
	mggttest
	ultrasound
	fatalmovementchart
	currentpregnancychecklist
	birthstory
	birthrecord
	postnatalmaternalcare
	postnatalmother
	checklistbf
	healthedu
	riskFactor
	motherHealthRecord

Table 4.10 (i): List of Dimensions and Description

Dimensions	Description
gender	All the information of gender
race	All the information of race
district	All the information of district
clinic	All the infiormation of clinic
bill	All the information of bill
state	All the information of state
riskfactor	All the information of risk factor
infectiousDisease	All the information of disease infection
rehabilitation	All the information od rehabilitation
emergencyTrauma	All the information od emergency and trauma
diabetic	All the information od diabetic
hypertension	All the information od hypertension
dignosist	All the information od dignosist
appointment	All the information od appointment
treatment	All the information od treatment services
medicine	All the information od medicine
staff	All the information od staff
campaign	All the information od campaign
patientReference	All the information od patient history and family
antenatal	All the information od antenatal

Table 4.10(ii): Fact and Measures tables

Fact	***Measure***
opd _services	totalNoOfTreatment
	totalNoOfAppointment
antenatal	totalNoOfPatient
registration	totalNoOfCharge
mch_service	totalNoOfCampaign
	totalNoOfBirthLife
	totalNoOfBirth
	totalNoOfPopulation
	totalNoOfClinics
	totalNoOfStaff
	totalNoOfPharmasist
	totalNoOfPatientAntenatalgetHypertension
	totalNoOfEmergencyCareReported
	totalNoOfPatientAntenatalgetDiabetic
	totalNoOfCommunityClinics
	totalNoOfMaternalandChildClinics
	totalNoOfMedicalOfficer
	totalNoOfNurse
	totalNoOfMedicalAssistant
	totalNoOfIncidenceCase
	totalNoOfWrongPerscription

4.5 Results of Mapping Process to Conceptual Schema

The mapping process starts from extended rationale diagrams produced by decisional modeling (see example on Figure 4.10) and organizational modeling (see example on Figure 4.7). The process mapping will use the required notation as below Table 3.1 (see subsection 3.5.3). The processes start with like below.

Step 1: The fact registration is identified in fact analysis based on decisional perspective. For every fact defined becomes the root of an attribute tree. Fact (*F*) is represented as a table in the conceptual schema.

Step 2: For each F table, then, attribute tree is build. Each node or circle of the attribute tree corresponds to one or more attribute in the decision model. The root of the attribute tree corresponds to the primary key of *F*.

Step 3: Then, attribute (*a*), dimensions (*d*), and measures (*m*) are defined from extended rationale diagrams produced by decision modeling and organizational modeling. The dimension of registration for fact OPD Services become union and map to the dimension registration for fact MCH Services to create new fact called Registration. The fact shared at least one dimensions and attributes produced for each other. As dimension is a fact property, it is represented as a small circle that connected to a fact.

Step 4: The *d* and *m* are mapped by using *a* represented during organizational modeling as a bridge (Figure 4.11). For example, attribute genderId for record gender in the organizational model is mapped to dimension Gender (represent as node or circle) associated to the analysis goals analyze gender in decisional model. The same attribute gender might for instance correspond, on the source schema, to an attribute genderId within the Registration table.

Step 5: The *d* that associated with *F* is mapped to the source schema and generates the full hierarchy rooted of *d*. Multiple values of d also can be related to a single instance of the F. All *d* that associated with *F* to the source schema are then generate the full hierarchy rooted of *d* as Figure 3.21.

Step 6: Then, non-dimension attribute (non *d-a*) for *d* is defined.

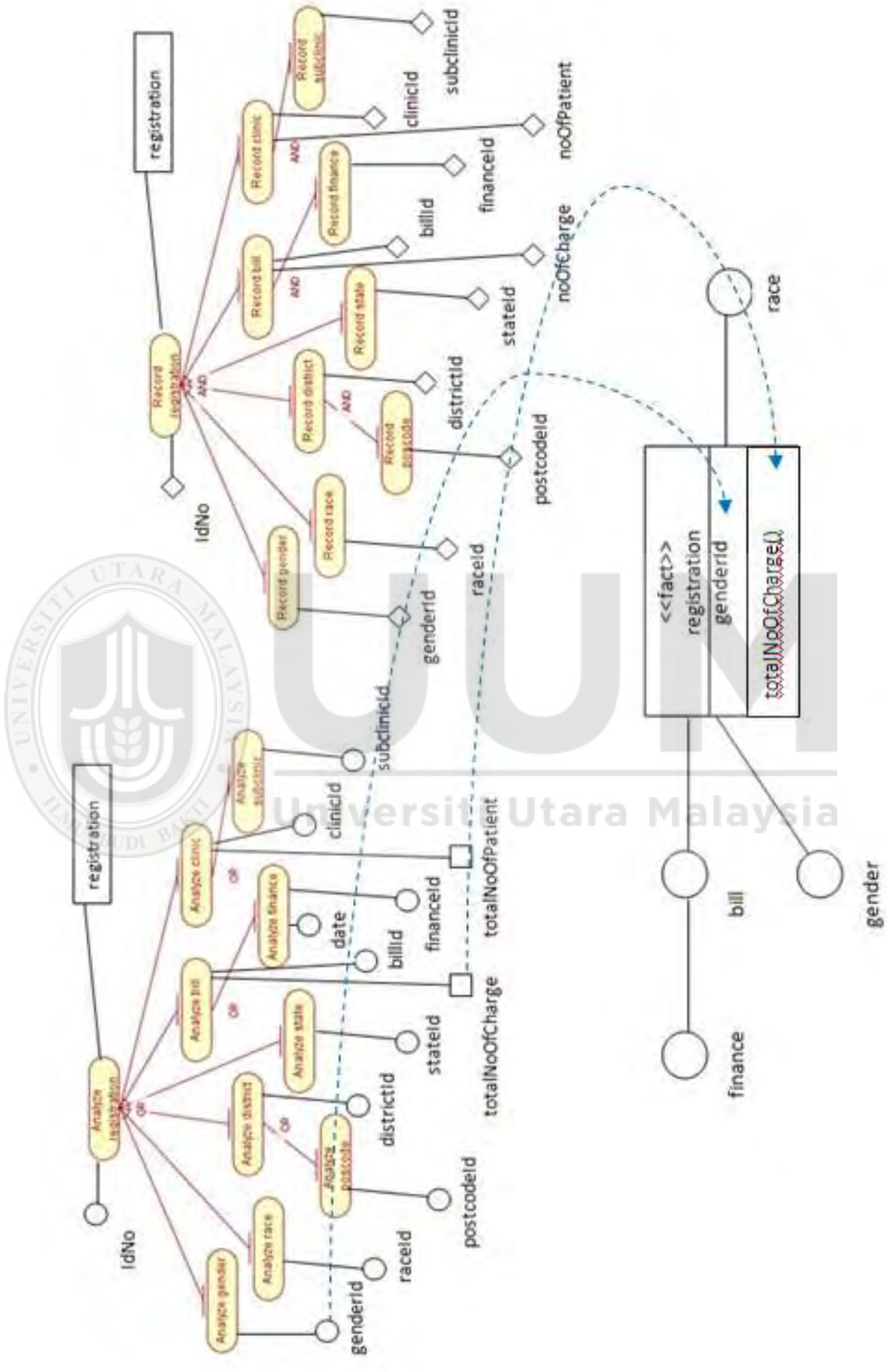
Step 7: The *m* among the children was chosen of the root. Every *m* that associated with a goal related to *F* and successfully mapped to conceptual model is included, but there no hierarchy is generated for it. A fact may have no measure. An attribute cannot be both a measure and a dimension.

Step 8: Each *a* is mapped to a physical *a* in the source schema and to *d* or *m* in the decision model.

Step 9: Attributes in the organizational model that not mapped into dimensions or measures in the decision model are mapped nonetheless on the source schema. Attributes are the field whose value is provided when a fact is recorded to fulfill the goal and represented as a small diamond that connected to the goals.

Step 10: If there have more than one fact, facts may be overlapped to create a new fact schema or linked at least one or more same dimension. For example, fact registration in OPD Services is overlapped to fact registration in MCH Services.

Figure 4.11 shows an example of requirement mapping process for fact "Registration" from organizational modeling and decisional modeling to produce fact table "Registration" on Figure 4.12. The overall mapping process including all facts, dimensions, attributes and measures produced the OPD Services conceptual schema (Figure 4.13) and MCH Services conceptual schema (Figure 4.14).


registration
Record registration
AND
IdNo
Record gender
Record race
Record district
Record postcode
Record state
Record bill
Record finance
Record clinic
Record subclinic
genderId
raceId
postcodeId
districtId
stateId
billId
financeId
noOfCharge
noOfPatient
clinicId
subclinicId
registration
Analyze registration
OR
IdNo
Analyze gender
Analyze race
Analyze district
Analyze postcode
Analyze state
Analyze bill
Analyze finance
Analyze clinic
Analyze subclinic
genderId
raceId
postcodeId
districtId
stateId
billId
date
financeId
totalNoOfCharge
totalNoOfPatient
clinicId
subclinicId
<<fact>>
registration
genderId
totalNoOfCharge()
race
bill
gender
finance

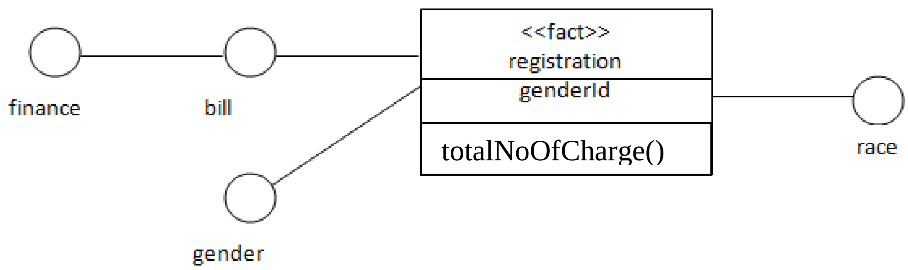


Figure 4.12: Fact Table registration

The process mapping finish when conceptual schema was designed for all facts and dimensions. Figure 4.13 shows the conceptual schema for OPD Services. It contains two facts which are OPD Services fact and registration fact connected to dimension (nodes). The conceptual schema shows how the facts and dimensions are connected to each other's by using conceptual schema notation from Table 4.10 List of Goal and conceptual notation. The same process used for MCH Services (Figure 4.14).

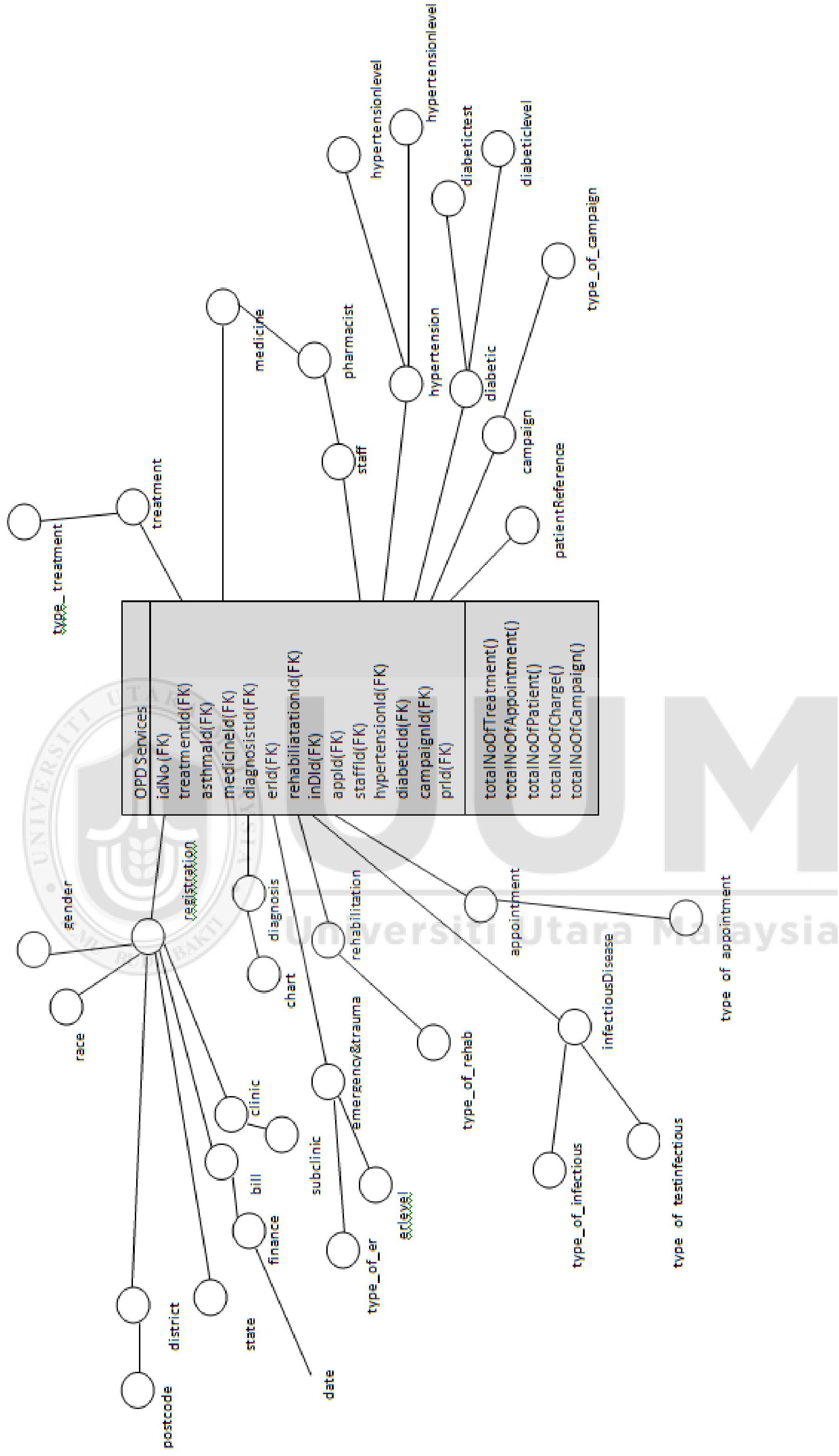


Figure 4.13: Conceptual schema for OPD Services

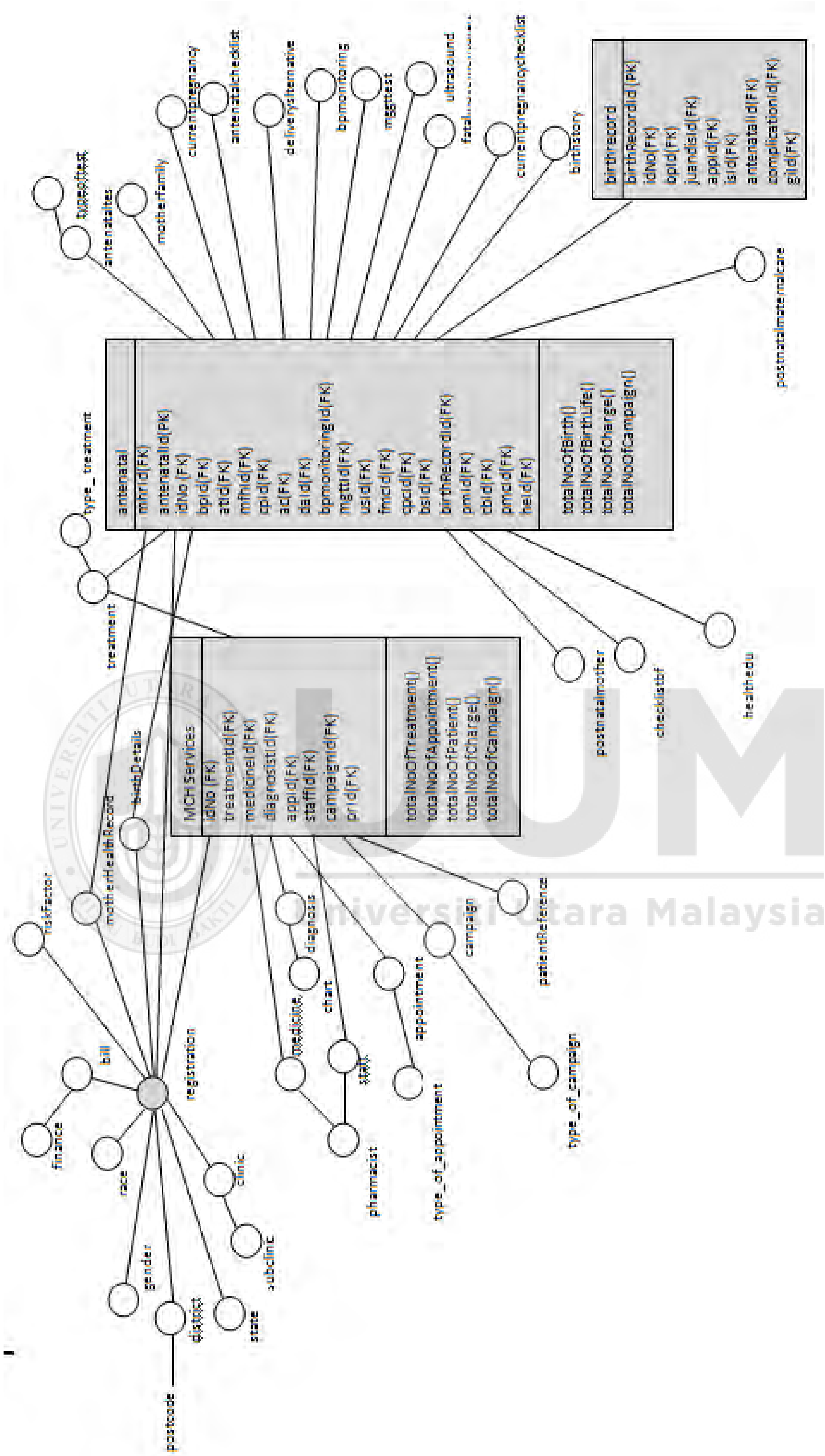


Figure 4.14: Conceptual schema for MCH Services

4.6 Mapping Process from conceptual schema to UML Class Diagram

Figure 4.15 shows the concept of process mapping from the conceptual schema to G-UML class diagram. The process starts with the conceptual schema from the previous process. The conceptual schema is then mapped to class diagram by using the G-UML notation in the Table 3.2 (see subsection 3.5.4). The mapping starts with:

Step 1: Facts and dimensions are identified. The facts and dimensions are mapped onto entities in the UML class diagram. Facts are graphically represented as a rectangle or a table that connected to the dimensions, represented as a node or a circle, map to entity (class) in UML is drawn as boxes (Figure 4.15).

Step 2: Then, attributes for each facts and dimensions are added to the entity (class) respectively in second compartment beginning with a lower-case letter (Figure 4.14).

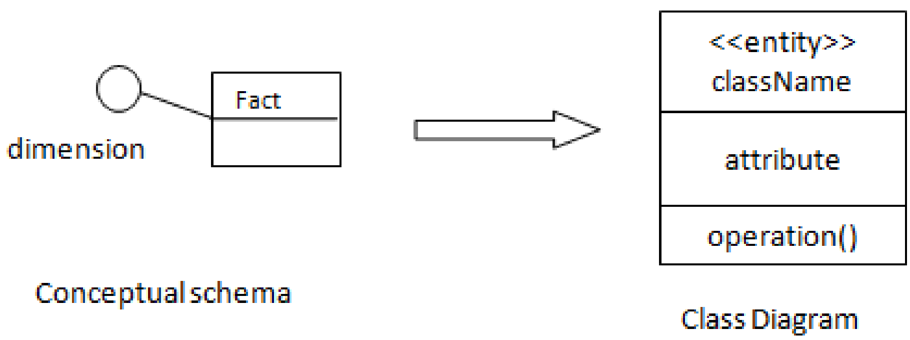


Figure 4.15: Process mapping from fact and dimension from conceptual schema to class diagram

Step 3: Then, the operation of the fact is added to third compartment. The measure is an operation in the class diagram. As for attributes, operation names are written beginning with a lower-case letter (Figure 4.15).

Step 4: The dimension attributes that associate to dimension are mapped to class diagram as <<agggregationLevel>>.

Step 5: A connection between each class produced are linked

Step 6: Text is added to show the relations between the class links. The text at the association end gives a name to the role that the instances of the class at the end of the association (Figure 4.16).

Step 7: The multiplicity constraint is defined.

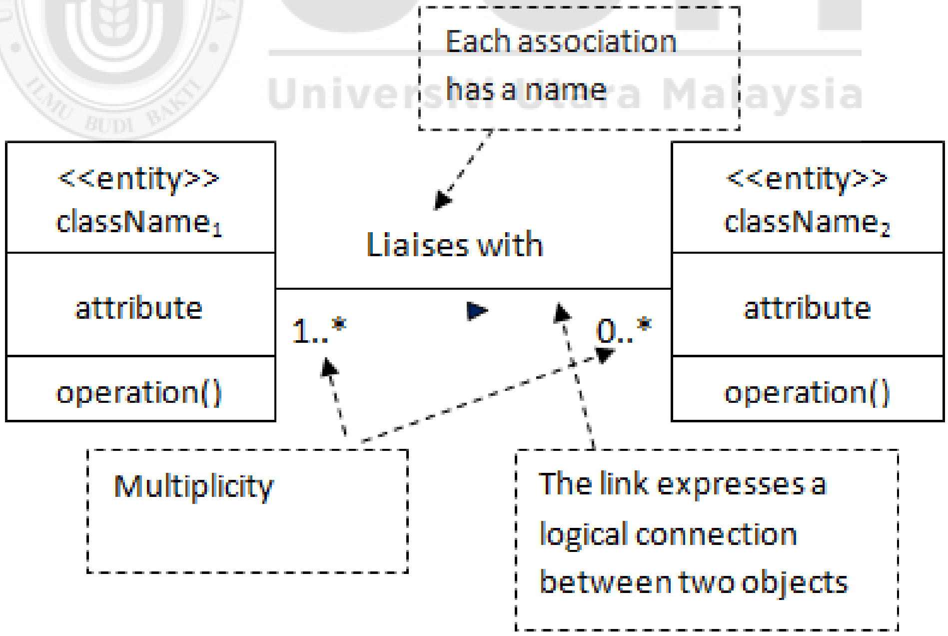


Figure 4.16: An example of class diagram

The process mapping starts with the conceptual design produced and requirements in Table 4.11. Facts and dimensions are represented by <<Fact>> and <<Dimension>> classes, respectively. Figure 4.17 shows the class diagram for Rural Health Care DW based on notation on Table 4.11. The model show that the dimensions are associated mapped to facts from decisional model to source schema with many-to-one relationship. Every measure associated mapped to a goal related to facts from decisional model to the source schema and provide many-to-one relationship. The dimension-attribute is connected as aggregation level.

In this mapping process, the classes are mapped to tables, attributes to column, types to data types, and associations to relationships. Some elements in the modeling are not being representing in the logical model because they are not stored in the database. For example, totalAmount, totalNoOfTreatment, which represent the sum of multiple columns in the database, is not stored because it is a calculation in the application and it is a derived attribute.

Table 4.11(i): Fact and dimension tables for each business process of Rural Health Care

Fact Table	**Dimension Tables**	**Dimension Attribute**
opd_mchServices	registration	gender
antenatal		race
registration		district

	poscode
	state
	finance
	bill
	clinic
	subclinic
infectiousDisease	type_of_infectious
	type_of_testinfectious
rehabilitation	type_of_rehab
emergencyTrauma	type_of_er
	erlevel
diabetic	diabeticlevel
	diabetictest
hypertension	hypertensionlevel
	hypertensiontest
dignosist	Chart
appointment	type_of_appointment
treatment	type_of_treatment
medicine	pharmasist
Staff	
campaign	type_of_campaign
patientReference	

antenatal	registration
	birthDetails
	antenataltest
	motherfamilyhistory
	currentpregnancy
	antenatalchecklist
	deliveryalternative
	bpmonitoring
	mggttest
	ultrasound
	fatalmovementchart
	currentpregnancychecklist
	birthstory
	birthrecord
	postnatalmaternalcare
	postnatalmother
	checklistbf
	healthedu
	riskFactor
	motherHealthRecord

Table 4.11(ii): Fact and Measures for each business process of Rural Health Care

Fact	***Measure***
opd_mchService	totalNoOfTreatment
	totalNoOfAppointment
antenatal	totalNoOfPatient
registration	totalNoOfCharge
	totalNoOfCampaign
	totalNoOfBirthLife
	totalNoOfBirth
	totalNoOfPopulation
	totalNoOfClinics
	totalNoOfStaff
	totalNoOfPharmasist
	totalNoOfPatientAntenatalgetHypertension
	totalNoOfEmergencyCareReported
	totalNoOfPatientAntenatalgetDiabetic
	totalNoOfCommunityClinics
	totalNoOfMaternalandChildClinics
	totalNoOfMedicalOfficer
	totalNoOfNurse
	totalNoOfMedicalAssistant
	totalNoOfIncidenceCase
	totalNoOfWrongPerscription

Figure 4.17: G-UML Class Diagram for OPD Services Rural Health Care DW

A multiplicity constraint is the association relationship between objects. Multiplicity constraints are used to check and maintain data quality. Table 4.12 shows the multiplicity constraint that associated between dimensions to fact, fact to fact and aggregrationLevel to dimension. For example, one staff manage many patient (opd_mch Services) record.

Table 4.12: Multiplicity Constraint for G-UML model

Class Name	Multiplicity	Relationship	Multiplicity	Class Name
<<Dimension>> staff	1..*	manage	1..*	<<Fact>> opd_mchServices
<<Dimension>> registration	1	had	1..*	<<Fact>> opd_mchServices
<<Fact>> antenatal	1	had	0..*	<<Fact>> opd_mchServices
<<aggregationLevel>> gender	1	associated with	1	<<Dimension>> registration
<<Fact>> birthRecord	1..*	refer to	1	<<Fact>> Antenatal

<<Dimension>> Medicine	1..*	associated with	1	<<Fact>> opd_mchServices

4.7 Summary

This chapter shows the implementation of G-UML modeling approach in Rural Health Care to develop DW requirement modeling. The processes consist of three steps which is step one, requirement elicitation, step two, requirement modeling using goal-oriented, and step 3, develop requirement modeling for DW. In requirement elicitation, requirement analysis help to analyze the requirements and intentions from qualitative research method including background reading, semi-structured interview, observation and document sampling. List of early requirements (Figure 4.1) is produced from this phase.

Then, the process requirement modeling begins using goal-oriented to design organizational modeling and decisional modeling based on result in requirement analysis step one. The output produced from this process is actor diagram and relationale diagram based on organizational perspective and decision-maker perspective. The processes in Step 3 continue with requirement mapping process, where the results produced in requirement modeling are map to conceptual schema (Figure 4.13 and Figure 4.14) and from conceptual schema to UML Class Diagram (Figure 4.17). There are 10 steps used to mapping the requirements to conceptual schema and 7 steps are used to mapping the conceptual schema to UML Class Diagram. All the processes use Rural Health Care case study and DW Tools and

OME. Based on the G-UML model presented in Figure 3.2, this chapter highlights the result after implementation G-UML modeling approach to Rural Health Care to design requirements of DW modeling.

The verification process for G-UML model using goal and UML class diagram is further discussed in Chapter 5 by using expert reviews based on Health Care case study. However, for this research, the implementation phase is excluded in the process of DW development for health care case study.

CHAPTER FIVE
RESULTS OF MODEL EVALUATION

5.1 Introduction

This chapter presents the verification process for G-UML model using goal and UML class diagram based on case study Rural Health Care. The verification process is done by an expert review to ensure the modeling of G-UML has achieved the user requirements from stakeholders and decision-makers perspective and to identify the strength and weakness of the G-UML. The discussion about the process ends with the conclusion of the chapter.

5.2 Verification for G-UML Requirement Modeling

This research used three experts review. Two of them are from DW industry that expert in DW development and modeling to verify the correctness of the proposed DW requirement modeling and one expert review from the Health Care domain to validate the requirements produced from the proposed DW requirement modeling. They have experiences almost 4 to 9 years in their field. The verification process consists of three parts. First part required expert review to design DW model using conceptual schema and compare the design with DW model produced by G-UML model.

Two experts review from DW industries are from private company. First expert review is a Software Engineer at Core Banking, Maybank Tower. The main project that he has conducted is Maybank Electronic Application Accomodation (MEAA) and Enterprise Origination System (EOS) for Maybank. For EOS, the data produce based on user requirements. EOS produced the data (automatic generated by system everyday) in text files format and the data isolated by pipeline and used in DB2. He has experience in DW development for 4 years.

Second expert review is from private company. He is a software Engineer. He has experience in DW and software development for 5 years. He has his own company at Perak. He made a collaboration with Wordpress and other open sources software where he active in developing themes, plugins and mobile applications to them. He is active in web and system development for business used. The expert review is chosen because it is a viable method of verifying and validating conceptual models in software engineering (Roger et al., 2010; Lazar, Feng, & Hochhneister, 2010).

One expert from Health Care domain have been chosen to verify the requirement produced from the proposed DW requirement modeling. She is a nurse from the Klinik Kesihatan Mempaga, Pahang. She has experience in organizing outpatient record including pharmacy record, staff record, mother and child record and others. She had experience almost 9 years in this field.

Based on the discussion in Section 3.6, the verification process started with the following steps.

i. Case study Rural Health Care as Section 4.2, interview result (Table 4.1a) and document sampling (Table 4.1b) is provided to two experts from DW industries.

ii. Expert review is required to design the DW model using the current model which is conceptual schema or UML class diagram.

iii. After DW model is designed, expert review is provided with a questions review (Table 3.3). Based on the DW model designed, expert review is required to check the DW model design whether the design can answer the question given or not.

iv. Based on the result in (iii), expert review is required to answer the Expert Verification Instrument provided (Appendix A).

v. Expert review from Health Care domain is required to verify the requirements produced. Details explanation about the requirements produced is given to better understand and one set of question is provided to the expert to answer the questions (Appendix D).

5.2.1 Result by Expert Review for Case Study

The case study required expert review to design UML class diagram or conceptual schema for Rural Health Care. The same result of interview questions and document sampling are provided to both expert reviews. Figure 5.1 shows the list of tables produced by expert review 1. The total number of tables produced is fifty-four (54) tables. Facts, dimensions and measures are listed in Table 5.1. There are two (2) facts produced which is opd_patient and mch_patient, and fifty-one (51) dimensions. However, there are no measure is mentioned at the conceptual design Figure 5.2.

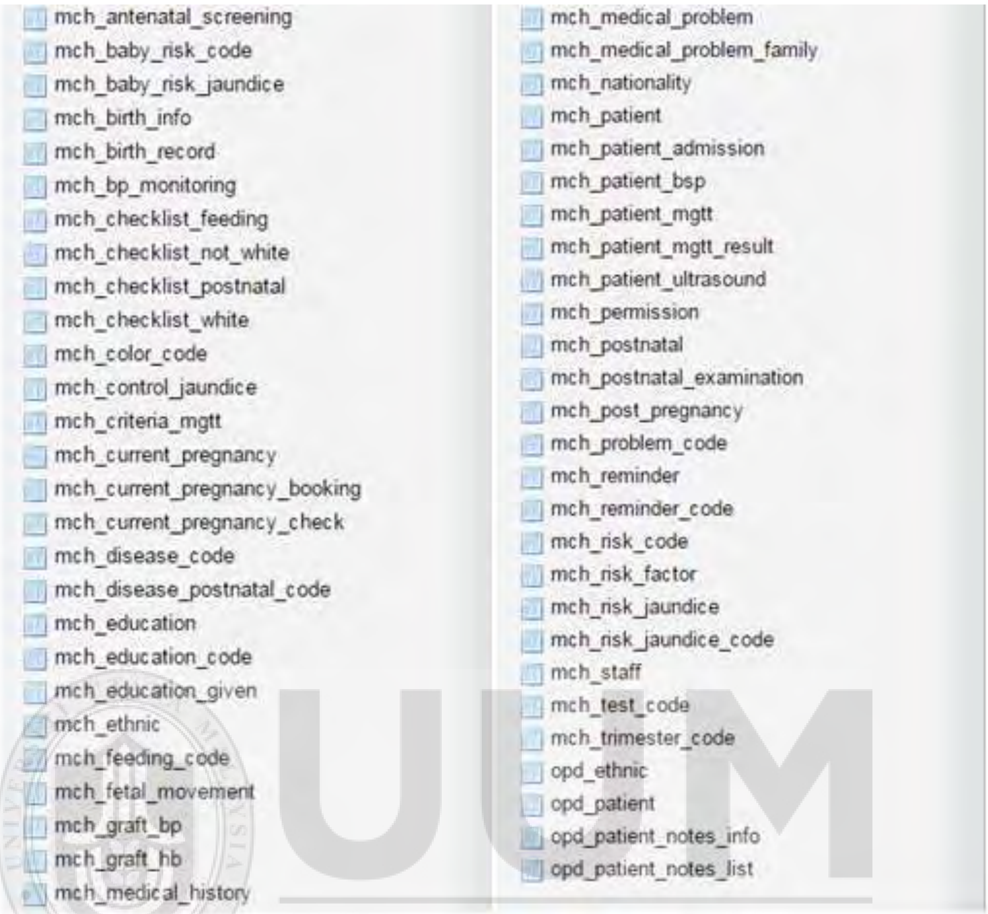


Figure 5.1: List of Tables produced by Expert Review 1 for Case Study

Figure 5.2(i) : Conceptual Design from Expert Review 1

Table 5.1: List of Fact, Dimension and Measure from Expert Review 1

Fact	***Dimension***	***Measure***
opd_patient	opd_ethic	totalNoPopulation
	opd_patient_notes_info	totalNoPatient
	opd_pateint_notes_list	
mch_patient	mch_antenal_screenin	totalNoMale
	mch_babu_rick_code	totalNoFemale
	mch_baby risk_jaunice	totalNoCommunityClinic
	mch_birth_info	totalNoHealthClinic
	mch_birth_record	totalNoMCclinic
	mch_bp_monitoring	totalNoMO
	mch_checklist_feeding	totalNoDO
	mch_checklist_not_white	totalNoPharmacist
	mch_checklist_postnatal	totalNoNurse
	mch_checklist_white	totalNoMA
	mch_color_code	totalNotreatmentdiabetc
	mch_control_jaundice	totalNoappointment
	mch_criteria_mgtt	totalNoIncidensecase
	mch_current_prenancy	

mch_current_pregnancy_booking

mch_current _pregnancy_check

mch_disease_code

mch_disease_postnatal_code

mch_education

nch_education_code

mch_education_given

mch_ethic

mch_feeding_code

mch_fetal_movement

mch_graft_bp

mch_graft_hb

mch_medical_history

mch_medical_probelm

mch_medical_problem_family

mch_nationality

mch_patient_admission

mch_patient_bsp

mch_patient_mgtt

mch_patient_mgtt_result

mch_patent_ultrasound

mch_permission

mch_postnatal

mch_postnatal_examination

mch_post_pregnancy

mch_problem_code

mch_reminder

mch_reminder_code

mch_risk_code

mch_risk_factor

mch_risk_jaundice

mch_risk_jaundice_code

mch_staff

mch_test_code

mch_trimester_code

Figure 5.3 and Figure 5.4 shows conceptual design for OPD and MCH Rural Health Care from Expert Review 2. From the OPD Rural Health Care designs shows that there are two (2) facts, fourteen (14) dimensions and no measures are produced. Meanwhile in MCH Rural Health Care design shows that there are three (3) facts, thirty-four (34) dimensions and no measure are produced. The details facts, dimensions and measures produced are listed in Table 5.2.

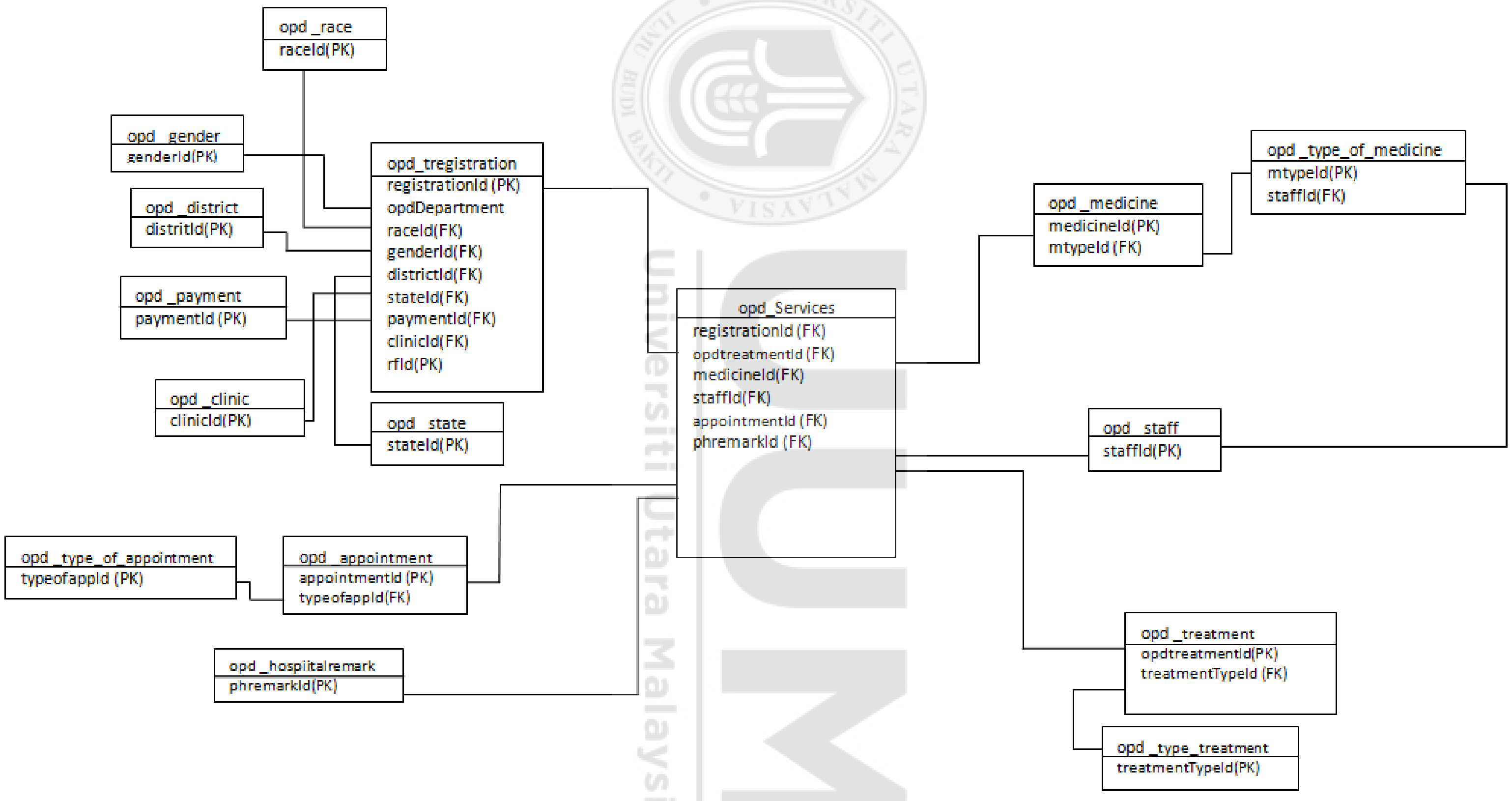


Figure 5.3: Conceptual Design for OPD Rural Health Care from Expert Review 2

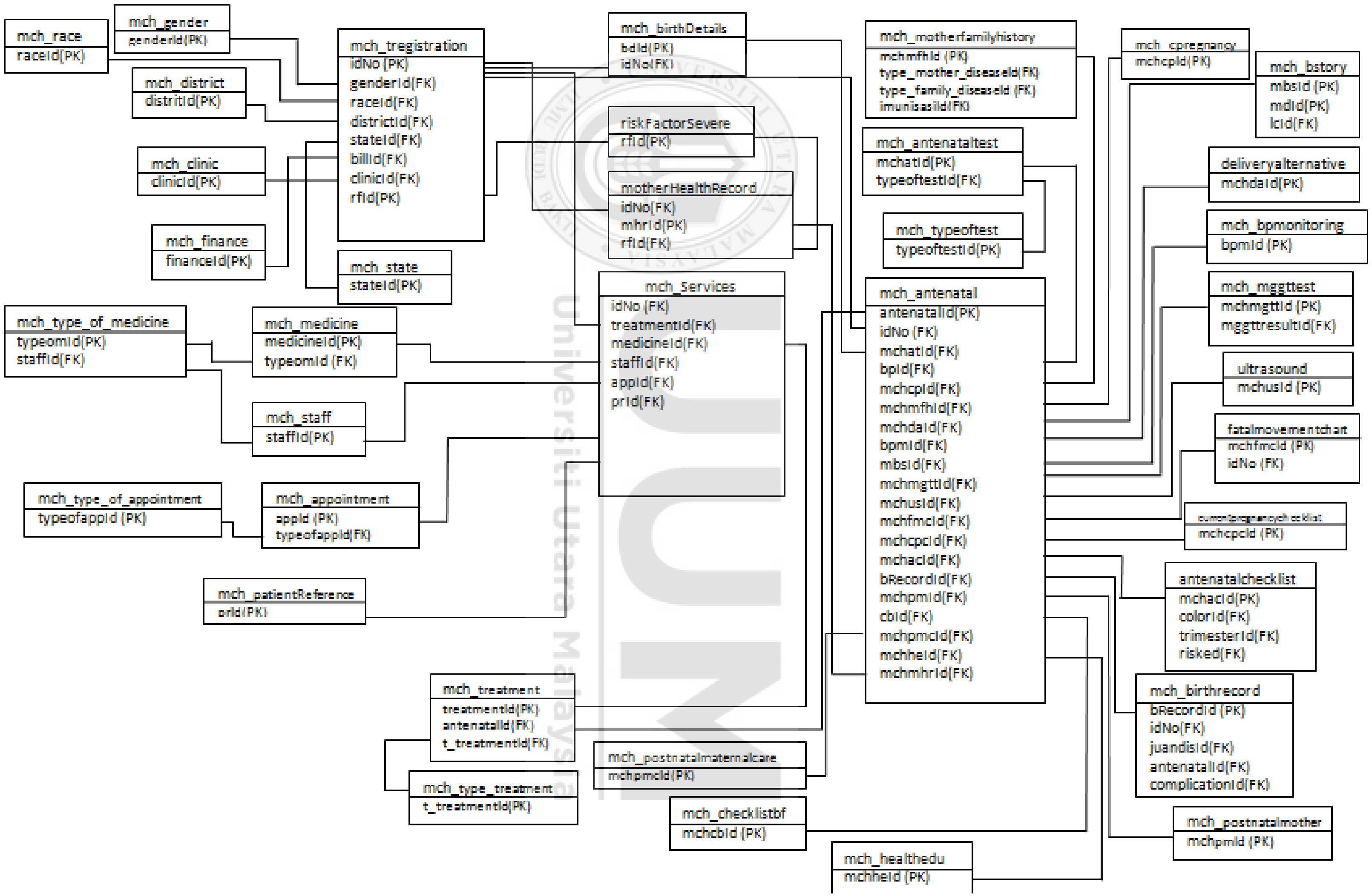


Figure 5.4: Conceptual Design for MCH Rural Health Care from Expert Review 2

Table 5.2: List of Fact, Dimension and Measure from Expert Review 2

Fact	**Dimension**	**Measure**
opd_services	opd_medicine	totalPopulation
opd_registration	opd_type_of_medicine	totalPatient
mch_services	opd_staff	totalBirthLifeMale
mch_registration	opd_treatment	totalBirthLifeFemale
mch_antenatal	opd_type_treatment	totalCommunityClinic
	opd_hopsitalremark	totalHealthClinic
	opd_appointment	totalmncclinic
	opd_type_of_appointment	totalMO
	opd_race	totalDO
	opd_gender	totalPharmacist
	opd_district	totalNNurse
	opd_payment	totalMA
	opd_clinic	totaltreatmentdiabetc
	opd_state	totalappointment
	mch_type_of_medicine	totalIncidensecase
	mch_medicine	
	mch_staff	
	mch_type_of_appointment	
	mch_appointment	
	mch_type_treatment	
	mch_treatment	

mch_patientReference

mch_gender

mch_race

mch_district

mch_clinic

mch_finance

mch_state

riskfactorsevere

mch_birthdetails

motherhealthrecord

mch_motherfamilyhistory

mch_antenataltest

mch_typeoftest

mch_cpregnancy

mch_bstory

deliveryalternative

mch_bpmonitoring

mch_mggtest

ultrasound

atalmovementchart

antenatalchecklist

mch_birthrecord

mch_postnatalmother

mch_healthedu

mch_checklistbf

mch_postnatalmaternalcare

currentpregnancychecklist

5.2.2 Result by using G-UML Modeling Approach for Case Study

Figure 5.5 shows the result by using G-UML modeling approach for designing DW based on case study Rural Health Care. The same result of interview questions and document sampling are used to design this model. From this G-UML model, there are three (3) facts produces after normalizing to reduce redundancy data. Table 5.3 shows the facts, dimension produced. Meanwhile, Table 5.4 shows the facts and measures produced. All tables have been normalized to reduce redundancy and control manipulating data based on goal-oriented approach used. Goal-oriented approach assists to produce details analysis of stakeholders, decision-makers and developer perspective intentions.

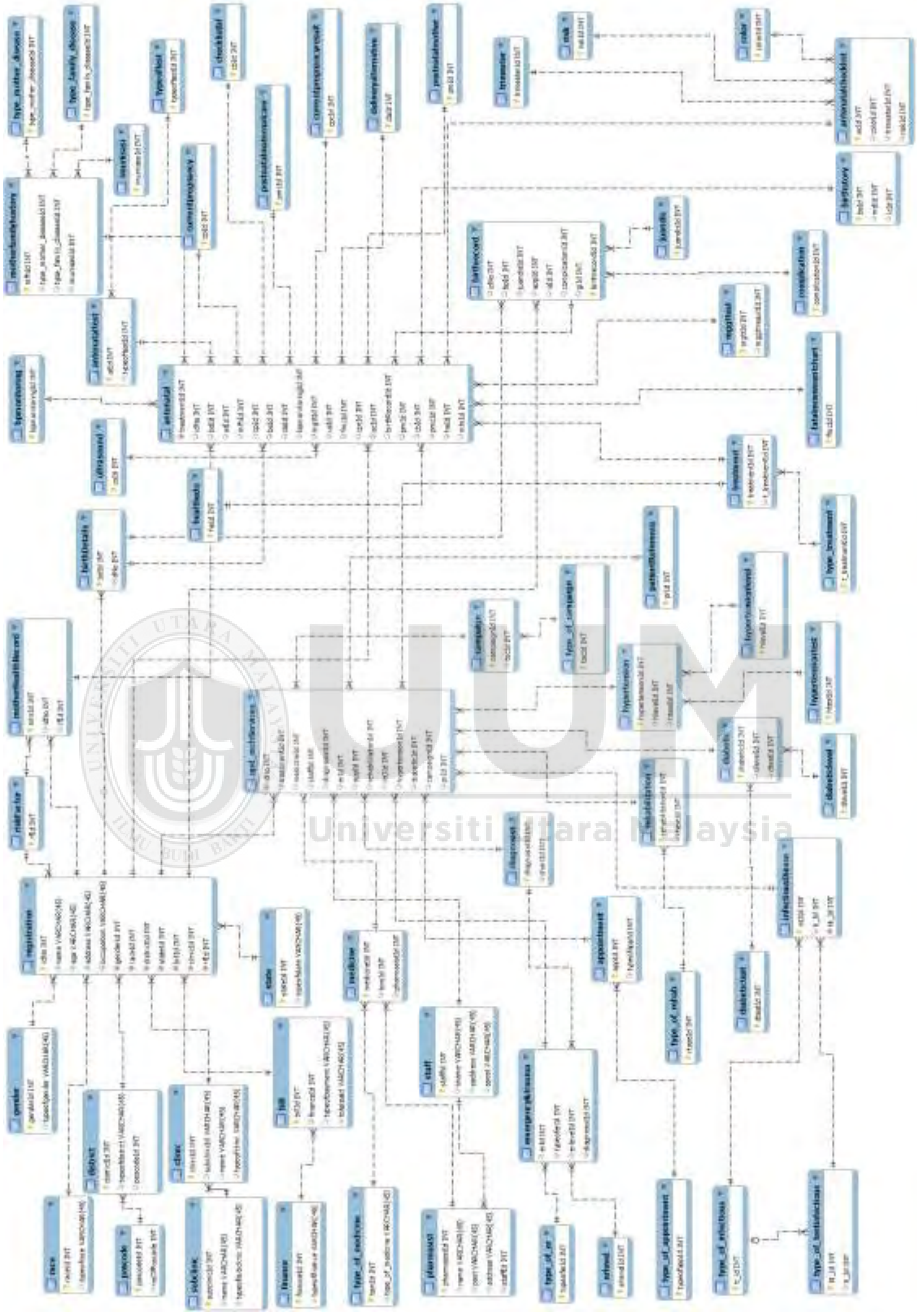


Figure 5.5: G-UML Class Diagram for OPD_MCHServices for Case Study

Table 5.3: Fact and dimension tables produced by using G-UML Model for Rural Health Care

Fact Table	**Dimension Tables**	**Dimension Attribute**
opd_mchServices	registration	gender
antenatal		race
registration		district
		poscode
		state
		finance
		bill
		clinic
		subclinic
	infectiousDisease	type_of_infectious
		type_of_testinfectious
	rehabilitation	type_of_rehab
	emergencyTrauma	type_of_er
		erlevel
	diabetic	diabeticlevel
		diabetictest
	hypertension	hypertensionlevel
		hypertensiontest
	dignosist	Chart
	appointment	type_of_appointment

treatment	type_of_treatment
medicine	pharmasist
Staff	
campaign	type_of_campaign
patientReference	
antenatal	registration
	birthDetails
	antenataltest
	motherfamilyhistory
	currentpregnancy
	antenatalchecklist
	deliveryalternative
	bpmonitoring
	mggttest
	ultrasound
	fatalmovementchart
	currentpregnancychecklist
	birthstory
	birthrecord
	postnatalmaternalcare
	postnatalmother
	checklistbf
	healthedu
	riskFactor
	motherHealthRecord

Table 5.4: Fact and Measures tables produced by using G-UML Model for Rural Health Care

Fact	***Measure***
opdmch_service	totalNoOfTreatment
	totalNoOfAppointment
antenatal	totalNoOfPatient
registration	totalNoOfCharge
	totalNoOfCampaign
	totalNoOfBirthLife
	totalNoOfBirth
	totalNoOfPopulation
	totalNoOfClinics
	totalNoOfStaff
	totalNoOfPharmasist
	totalNoOfPatientAntenatalgetHypertension
	totalNoOfEmergencyCareReported
	totalNoOfPatientAntenatalgetDiabetic
	totalNoOfCommunityClinics
	totalNoOfMaternalandChildClinics
	totalNoOfMedicalOfficer
	totalNoOfNurse
	totalNoOfMedicalAssistant
	totalNoOfIncidenceCase
	totalNoOfWrongPerscription

5.3 Analysis Requirements for Verification Process

First part of verification process is by using fact testing. The process started with testing on the design produced by experts review and G-UML model. The fact test can be achieved by checking, for each question reviews provided, how many questions can be answered based on the design produced to the required measures that have been included in the fact schema and to the required aggregation level that can be expressed as valid grouping set on the fact schema (Golfarelli & Rizzi, 2009).

The verification process is done by checking the design produced according to the questions that frequently asked by top management to medical practitioners in clinics at all the time. This is important to show that the DW developed is fully covered the user and decision-maker requirements. It was also conducted with regard to the number of artefacts produced by the two designs. The artefacts produced are very important to show the correctness of the late requirements produced to develop a DW. The connection between the artefacts shows that the data is normalized to reduce redundant data and space of storage. It also gives an advantage to end users to get fast access the DW.

Table 5.5: Result of verification process

Verification Process	G-UML Model	Conceptual Model by Expert Review 1	Conceptual Model by Expert Review 2
Questions Review for Case Study Rural Health Care			
1. Population Structure	✓	✓	✓
2. Life Expectancy at Birth	✓	✓	✓
3. Morbidity	✓	○	○
4. Health Facilities	✓	✓	✓
5. Post Information	✓	✓	✓
6. Promotion Matters	✓	○	○
7. Treatment Services	✓	✓	✓
8. Pharmacy Services	✓	○	✓
Artefacts produced for Case Study Rural Health Care			
i. Fact	3	2	5
ii. Dimension	33	52	48
iii. Measure	21	15	15

Table 5.5 shows the result between conceptual models design by expert reviews and G-UML design. Based on verification process, result from expert review 1 (Figure 5.2) and expert review 2 (Figure 5.3 and Figure 5.4), there are three questions that conceptual design from expert reviews cannot answer. The questions are number three, number six and number eight. This is because there is no table for diabetes, hypertension, campaign and medicine provided to answer the questions given. There are no measure allocate for this point for the design produced. The tables produced

are not completely normalized. For example, table registrations for both facts are not normalized. Meanwhile, based on G-UML model produced, the G-UML model can answer all the questions given. This is because the requirements produced are related to the user requirements. By using goal approach, it helps analyst to identify and analyse early and late requirements based on stakeholder and decision maker intentions, problems in decision-making process on organizational setting, exploring system solution and alternatives in early requirement phase and late requirement phase.

Besides that, the total number of artefacts produced by G-UML model shows the relationship for each table produced fulfils the user requirements need. There is some table produce are not completely normalized such as OPD services and MCH services for expert review 1 and expert review 2 designs. The artefacts produced must follow the DW concept where the DW proposed support the design. It must have fact, dimensions, dimension attribute, measure and relationship to automate the DW model. The artefact used to provide efficient support for communication between the designer and the end user to build the user's query specification requirements when DW is implementing.

Second verification process is based on expert verification instrument. According to Pederson and Jesen (1998), proposed DW model must meet the listed metrics for validity. Based on the expert verification metrics answered by expert review shows that all the experts agreed that the proposed G-UML modelling for DW is approximately ‘Very satisfactory’. This is because the proposed DW modeling consider the early requirements which some of the analyst and DW developer ignored and focused on the development DW itself.

Table 5.6: Analysis of expert verification instrument

Metrics	**Expert 1**	**Expert 2**	**Mean**
Explicit Hierarchies in dimensions	5	4	4.50
Symmetric treatment of dimensions and measures	4	4	4.00
Multiple Hierarchy in each dimension	5	5	5.00
Support for correct aggregation	4	4	4.00
Non-strict hierarchy	4	4	4.00
Many-to-many relationships between facts and dimensions	4	4	4.00
Handling uncertainty	5	5	5.00
Mean	4.43	4.29	4.36

Based on the result in Table 5.6, the evaluating of proposed DW modeling attracted not less than 4 (Satisfactory). There are two metrics has high average (5.00) which means that in one dimension, there can be more than one path along the dimensions to be aggregate dimensions and the model can identifies the uncertainties fact, dimension and attribute. The analysis shows that the proposed DW modeling support and expend the explicit hierarchies in dimensions. Meanwhile, the proposed DW modeling support and fulfil the others metrics with ‘Satisfaction’. Overall, the cumulative mean of the metrics which is 4.36 suggests the proposed DW modeling is appropriate and satisfactory.

Third verification process completed by nurse from Klinik Kesihatan Mempaga, Pahang with verification of requirements produced from proposed DW requirement modeling. Based on question review (Table 3.3) and the requirements produced, nurse is asked to answer the questions. The result of verification shows that all requirements produced can answer all the questions given. Details explanation is given based on the DW requirement modeling produced and how the modeling can produce the information and give the answers required based on the question reviews in Table 3.3. The respondent show the "Very Satisfactory" result from the verification process with regard to the result from Table 5.5 and Table 5.6.

Conclusion, based on the result produced by expert reviews and the involvement of expert from Health Care domain itself shows that G-UML model is an appropriate modeling to design DW and satisfactory. It was verified with regard to the number of questions answered, artefacts produced by the G-UML model, cumulative mean of the metrics by expert reviews and verification of late requirements produced. Besides that, the G-UML model has explicit hierarchies in dimensions, has symmetric treatment of dimensions and measures, and contains multiple hierarchies in each dimension, support for correct aggregation, non-strict hierarchy, many-to-many relationships between facts and dimensions and identifies the uncertainties fact, dimension and attribute.

5.4 Summary

This chapter explains the process of verification DW requirement modeling. The main objective is to verify the correctness of proposed modeling design by G-UML model. Three expert reviews are involved in this process. Two of them are from DW industry and one of them is medical practitioner from Health Care domain and they are expert in their career field with 4 to 9 years of experience.

There are five steps involves in verification process for expert reviews from DW industry. First, all the details information are given including Case study Rural Health Care as Section 3.3, interview result (Table 3.1a) and document sampling (Table 3.1b). The process started with expert reviews designs the DW model using the current model which is conceptual schema or UML class diagram. Then, expert reviews are asked to check the DW model design with a questions review (Table 3.13) whether the design can answer the question given or not. Based on the result from questions review, expert reviews are required to answer the Expert Verification Instrument provided (Appendix A).

Analysis requirement of verification process has been conducted to compare with G-UML model design. Based on the result produced by expert reviews shows that there are a few questions are not answers. This is because the conceptual design produced not allocated the table required to answer the questions. The normalization process is not complete. The result produced by G-UML model shows the important of early requirements and late requirements for design DW model. The verification process shows that the G-UML model developed is fully covered the user and decision-maker requirements. It was verified with regard to the number of questions answered and artefacts produced by the G-UML model. The artefacts produced are very important to show that the data is normalized to verify the correctness of the late requirements produced to develop a DW. The connection between the artefacts assists to reduce redundant data and give more space of storage and the cumulative mean of the metrics suggests the proposed DW modeling is appropriate and satisfactory. The cumulative mean of the metrics is 4.36 suggests the proposed DW modeling is appropriate and satisfactory.

The third expert review from Health Care domain required answering the questions provided in Appendix D. Based on the result shows that the expert review is satisfy with the information and record provided, regarding to the late requirements produced by G-UML model.

In summary, G-UML model proposed shows that the important of early and late requirements for designing DW. It proves that DW development processes can be designed from an early requirement phase to designing phase. The early and late requirements produced shows that the DW model was meets user requirements and can produced stable a DW model.

CHAPTER SIX
DISCUSSION AND CONCLUSION

6.1 Introduction

This chapter reviews all the finding and concludes the research works by giving a comprehensive view according to the research objectives. The main contribution of this research is provided. However, the research needs continuous effort to design efficient and effective DW modelling due to the complexity of the DW domain. Therefore, the limitations in designing DW modelling are discussed and overviews of the solutions are highlighted as well as recommendations for future work.

6.2 Discussion

This study proposed G-UML modeling approach to identify and analyze the early and late requirements for design stable DW model that comprises all the requirements needs by the stakeholders and decision-makers. The specific objectives are derived as below:

Objective 1: To identify the early and late requirements for designing DW systems. This research objective has been discussed in Chapter 3 and Chapter 4.

Objective 2: To develop DW requirement model by using G-UML modeling approach. This research objective have discussed in Chapter 3 and Chapter 4.

Objective 3: To verify the correctness of DW requirement model developed. This research objective have discussed and applied for case studies in Chapter 5.

The summary of the research works related to the research objectives are discussed in the following section based on Table 5.1.

6.2.1 Discussion Objective 1: To identify the early and late requirements for designing DW systems

DW requirement modeling started with requirement elicitation and requirement modeling. Requirement elicitation starts with qualitative research method to captured early requirements. Meanwhile, requirement modeling focused on requirements analysis for organizational modeling and decisional modeling using goal-oriented approach based on early requirements captured (Section 4.2). Both modeling is used to mapping the organizational goals and decision-maker perspective.

Early requirements are analyzed by goal analysis, fact analysis, dimension analysis, attribute analysis and measure analysis. The output produced help analyst to understand the goals of stakeholder that underlie in DW design and problem in decision-making process on organizational setting. This help analyst to explored the system and find out the best solution and alternatives to produce late requirements.

The late requirements produced shows the requirement specifications including functional and non-functional requirements for the DW design which focuses on completeness, consistency, and automated verification of requirements. Late requirements are produced at the end of G-UML tasks (Table 5.3 and Table 5.4). All the processes involved used Rural Health Care to achieved objective one. The output from this process is actor diagram, rationale diagram and extended rationale diagram is support objective one. All the diagrams produced uses in the next process of development requirement modeling which is support objective two.

6.2.2 Discussion Objective 2: To develop DW requirement model by using G-UML modeling approach

The process development requirement modeling continues as the result from requirement modeling produced. The process starts with architectural design where the extended rationale diagram mapped from organizational modeling as a bridge to decisional modeling (Figure 4.11) to produce conceptual schema (Figure 4.13 and Figure 4.14). There are 10 steps used in this mapping process. The process mapping continues from conceptual schema to UML Class Diagram to produce stable DW modeling based on G-UML modeling approach (Figure 4.17). There are 7 steps used in this mapping process. Stable DW design helps stakeholder and decision-maker in decision making process.

Next, the details design process continues to introduce additional detail for each architectural component of the G-UML modeling such as determines multiplicity constraint which is how goals assigned to each other are fulfilled by association, aggregation to show the relationship between fact, dimension and measure. The table constraint is produced as Table 4.12. The processes involved are support to achieve objective two.

6.2.3 Discussion Objective 3: To verify the correctness of DW requirement model developed

Objective three has been achieved by process verification the correctness of DW model developed. Two expert reviews used to verify the modeling from DW industry. They are from private company and they are expert in their career field. Meanwhile, there is one expert review from Health Care domain to verify the requirements produced. There are five step involves in verification process.

The process started with expert reviews from DW industry design the DW modeling using the current model which is conceptual schema or UML class diagram. Then, expert reviews are asked to check the DW model design with a questions review (Table 3.3) whether the design can answer the question given or not. Based on the result, expert review is required to answer the Expert Verification Instrument provided (Appendix A). Analysis requirement of verification process has been conducted to compare with G-UML model design. Expert review from Health Care domain is required to verify the requirements produced. The details explanation about the requirements produced is provided to better understand and one set of question is provided to the expert to answer the questions (Appendix D).

The result produced by G-UML model shows the important of early requirements and late requirements for design DW model. The verification process shows that the DW design developed is fully covered the user and decision-maker requirements and the cumulative mean of the metrics suggests the proposed DW modeling is appropriate and satisfactory. It was verified with regard to the number of artefacts produced by the two designs. The artefacts produced are very important to show the correctness of the late requirements produced to develop a DW. The connection between the artefacts shows that the data is normalized to reduce redundant data and give space of storage. It also gives an advantage to end users to get fast access the DW. Based on the result in second part of verification process, the cumulative mean of the metrics suggests the proposed DW modeling is appropriate and satisfactory. The cumulative mean of the metrics is 4.36 suggests the proposed DW modeling is appropriate and satisfactory. The result from the respondent in Health Care domain shows the "Very Satisfactory" result from the verification process with regard to the result from Table 5.5 and Table 5.6.

Table 6.1 The G-UML Model tasks

Tropos Methodology	Research Method	Activities	Output (Page No.)
Objective 1: To identify the early and late requirements for designing DW systems			
Phase 1: Early Requirement Phase	Requirement Elicitation	Identify and analysis all the requirements gathered	• List of Early Requirements for DW for Health Care

Phase 2: Late Requirement Phase	Requirement Modeling	Organizational Modeling:- -Goal Analysis -Fact Analysis -Attribute Analysis	DW Requirement Modeling based on Goal-oriented Approach:- -Actor diagram -Relationale diagram -Extended Relationale diagram
		Decisional Modeling:- -Goal Analysis -Fact Analysis -Dimension Analysis -Measure Analysis	

Objective 2: To develop DW requirement model by using G-UML modeling approach.

Phase 3: Architecture Design Phase Phase 4: Details Design Phase	Develop DW Requirement Model using Goal –UML Approach	Design DW Conceptual Schema and mapping to UML Class Diagram	• Conceptual Schema of Health Care • UML Class Diagram of Health Care

Objective 3: To verify the correctness of DW requirement model developed.

Phase 5: Implementation Phase	Requirement Model Verification	Expert Review Verification	• List of Late Requirements for DW for Health Care

6.3 Limitation and Recommendations for Future Work

There is very limited tools can support the modeling of goal-oriented diagrams. The current tools contains the basic functions, do not include the in-depth modeling particular tasks to represent actor, goal, resource, dependency and others. These tools are not supported the current modeling tools such as UML and ER diagram. Therefore, this research had subject to this limitation and constraints research to for supporting the proposed solutions. Indeed, the development of this tool is not included in this research work.

The process mapping goal-oriented to UML was carried out manually because of unavailable appropriate tools. The current tools such as DW-Tools, OME and MySQL Workbench were impossible to apply because of insufficient reasoning in supporting the mapping processes. This is because the current tool was built to specific domain, unsuitable for this research. The manual process slows the mapping processes and high possibility to get errors.

In respect to modeling and design DW modeling, this research suggests the following issues to tackle the current limitations towards future work. This suggestion due to the requirement needs in designing DW modeling.

6.4 Conclusion

This study had proposed a new method for designing DW requirement modeling, called G-UML modeling approach. The G-UML modeling approach focused on requirement elicitation, requirement modeling, develops requirement modeling for

DW and requirement modeling verification. The G-UML approach helps analyst and DW developer to clarify requirements needs, to analyze the intentions of the stakeholders and decision-maker, to meet business goals, to understand the function of requirements before develops DW systems, the flow of relationship and association between each requirement, and possibly decrease the failure of the DW in early requirement phase. The G-UML modeling approach also used to help analyst to represent late requirements in simple way to make sure the stakeholders, and decision making as the end user understand the flow of requirements functionality. This approach has covered both early and late requirements which are none of the current approaches have done. In summary, G-UML modeling approach proposed shows that the important of early and late requirements for designing DW. It proves that DW development processes can b designed from an early requirement phase to designing phase. The early and late requirements produced shows that the DW model was meets user requirements and can produced stable a DW model.

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

Appendices

Appendix A

Experts' Verification Instrument

Universiti Utara Malaysia

EXPERT VERIFICATION INSTRUMENT FOR REQUIREMENT MODELING FOR DATA WAREHOUSE

This instrument is for the verification of the proposed requirement modeling for data warehouse. The instrument has two (2) sections. Section A is to elicit your profile data while Section B contains questions used to assess the adequacy of the proposed model in terms of correctness of requirement capturing and independence of design levels in the model. This is done using seven (7) different criteria adapted from Pedersen and Jesen's (1998) "Multidimensional Data Modeling of Complex Data".

All information supplied will be treated with utmost confidentiality, for the purpose of this research only, and will anonymous reportage in academic publications. Your attention is kindly appreciated.

Section A: Expert Profile

Name	
Position	
Research Interest	
Experience(in Years)	

Section B: Having assessed the proposed data warehouse model, please, kindly rate the model through the following items using 5-point Likert Scale of 1 (Not Satisfactory), 2 (Fairly Satisfactory), 3 (Neutral), 4 (Satisfactory) and 5 (Very Satisfactory).

No.	**Metrics**	**Description**	**1**	**2**	**3**	**4**	**5**
1	Explicit Hierarchies in dimensions	The relation of hierarchies in the dimensions should be captured explicitly in the data model, so then can identify the different level of hierarchy in dimension					
2	Symmetric treatment of dimensions and measures	The data model should allow the fact measures to be treated as dimensions.					

3	Multiple Hierarchy in each dimension	In one dimension, there can be more than one path along which to aggregate data					
4	Support for correct aggregation	The data model can produce the result according to the condition of requirements need when aggregating data. One aspect of this is to avoid double-counting of data.					
5	Non-strict hierarchy	The data model shows the hierarchies in a dimension are not always strict; it may have many-to-many relationships between the different levels in a dimension.					
6	Many-to-many relationships between facts	The data model show the relationship between fact and dimension is					

	and dimensions	not always the classical many-to-one mapping.					
7	Handling uncertainty	The data model should allow grouping a few requirements in a result.					

Appendix B

INTERVIEW GUIDE

This guide presents the questions to be asked during interview session

Position Name:____

POSITION	TASK & GOAL	ACTION TAKEN	DATA

Appendix C