جـامـعـة الـشـرق الأوسـط للدراسـات العليا
MIDDLE EAST UNIVERSITY FOR GRADUATE STUDIES

Towards **an O**ntology **f**or **R**equirements **E**ngineering **P**rocess **in E**-**G**overnment **A**pplications

By

Mohammad Nizar Ahmad
Faculty of Information Technology
Middle East University for Graduate Studies

Supervisors

Prof. Mohammad A. Al Fayoumi
Faculty of Information Technology
Middle East University for Graduate Studies

Associate Supervisor

Dr. Ahmad K. A. Kayed
Software Engineering
Fahad Bin Sultan University

A thesis submitted in partial fulfillment
of the requirements for the Master degree in
Computer Information Systems

Amman, Jordan
August, 2009

AUTHORIZATION FORM

إقرار تفويض

أنا محمد نزار نايف أحمد أفوض جامعة الشرق الأوسط للدراسات العليا بتزويد نسخ من رسالتي للمكتبات أو المؤسسات أو الهيئات أو الأفراد عند طلبها.

التوقيع:

التاريخ:

Authorization statement

I, Mohammad Nizar Nayef Ahmad, Authorize the Middle East University for Graduate Studies to supply a copy of my Thesis to libraries, establishments or individuals upon their request.

Signature:

Date:

COMMITTEE DECISION

This Thesis (Towards an Ontology for Requirements Engineering Process in E-Government Applications) was successfully defended and approved on August 24th 2009

Examination Committee signatures:

Prof. Mohammad A. Al Fayoumi

Associate Professor, Department of Computer Information Systems

Middle East University for Graduate Studies

....

Dr. Ahmad K. A. Kayed

Associate Professor, Department of Software Engineering

Fahad Bin Sultan University

....

Dr. Ghassan F. Issa

Associate Professor, Department of Computer Science

University of Petra

....

Dr. Hazim A. Farhan

Assistant Professor, Department of Computer Science

Middle East University for Graduate Studies

....

DECLARATION

I do my hereby declare the present research work has been carried out by me under the supervision of Professor Mohammad Al Fayoumi and Dr. Ahmad Kayed and this work has not been submitted elsewhere for any other degree, fellowship or any other similar title.

Signature:

Date:

Mohammad Nizar Nayef Ahmad
Department of Computer Information System
Faculty of Information Technology
Middle East University for Graduate Studies

DEDICATION

I dedicate this work to my "late Father", dear mother, dear brothers and sisters, for their love and support; they were the light in my academic path and without them nothing of this would have been possible.

I also dedicate this work to my grandfather, grandmother, uncles, and aunts for their kind patience.

And a special dedication to my best friends who give me a hope in my life "Belal and Yousef" and "Hasan" for their love, patience, and continuous support and encouragement.

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to Prof. Mohammad Al-Fayoumi and A/Prof. Ahmad Kayed for their guidance, support and motivation through out my Master's Thesis.

I would further like to acknowledge all of the Information Technology faculty members at the Middle East University for Graduate Studies and I am particularly grateful to Dr. Musbah Aqel and Dr. Hazim Farhan for helping and encouraging my efforts during the thesis work.

Also, I would like to especially thank my dear friends Mr. Ahmad Samhan, and Mr. Hamdi Ghandour, and Mr. Belal Abu Al-Haija, and Mr. Yousef Thouqan for supporting me during the whole time of my studying and research.

Thank you All.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF ACRONYMS

REPEGAs Requirements Engineering Process in E-Government Applications.

RE Requirements Engineering.

EGA E-Government Application.

REP Requirements Engineering Process.

SW Software.

IEEE Institute of Electrical and Electronics Engineers.

NATO North Atlantic Treaty Organization.

KAON KArlsruhe ONtology.

DLP Description Logic Programs.

BRITE Business Register Interoperability Throughout Europe.

BR Business Register.

EU European Union.

HLDO High Level Domain Ontology.

BRPO Business Register Process Ontologies.

MS Microsoft.

SWE Software Engineering.

ABSTRACT

In this work we tried to solve the problem that faces software engineers in understanding and modeling requirements for e-government applications, since they are facing some problems when they intend to define the domain concepts and terms clearly and in a shared manner between them, in this thesis we focus on studying the Requirements Engineering Process in E-Government Applications (REPEGAs) concepts and terminologies that current REPEGAs proposals present, in order to extract a conceptualization for the REPEGAs domain. We collected and studied many documents and reports that discussed REPEGAs in their contents, we extracted, studied, evaluated, and enhanced an ontology domain concepts from the most common concepts used in the semantic of the collected documents. Later we extracted and presented general relationships between the suggested ontology concepts. Those presented concepts along with the extracted relationships are introduced as an ontology that is considered as a first in the specific domain of REPEGAs. We condensed a lot of concepts used to define the most common discussed and studied REPEGAs into a smaller set of concepts consists of some concepts; this ontology can be used by software engineers, researchers, practitioners, and stakeholders as a common agreement of REPEGAs pool of knowledge in order to solve the inconsistency problem in the semantic between them while defining or using any of the definitions of the discussed REPEGAs. In addition, our ontology provides a base to evaluate any related definition semantic for one of the studied attributes.

Keywords:
Conceptualization, Ontology, Software Engineering, Requirement Engineering Process, E-Government Applications, Semantic Inconsistency, Relationship Lattice, Ontology Evaluation.

ملخص

في هذا البحث حاولنا حل المشكلة التي تواجه مهندسي البرمجيات في فهم ونمذجة متطلبات تطبيقات الحكومة الإلكترونية, حيث أنهم يواجهون بعض المشاكل عندما يريدون تعريف مصطلحات ومفاهيم هذا المجال بوضوح وبهيئة مشتركة بينهم. في هذه الرسالة نركز على دراسة المفاهيم والمصطلحات التي تعبر عن مرحلة هندسة المتطلبات في تطبيقات الحكومة الإلكترونية التي تتناولها الأبحاث في هذا المجال لاستخلاص تصوّر دلالي عام منها. لقد قمنا بتجميع ودراسة العديد من التقارير والوثائق المجمعة. بعد ذلك قمنا باستخلاص ودراسة وتقييم وتحسين للمفاهيم وتقديمها كمفاهيم مقترحة للأنتولوجي المقترح من خلال دراسة المفاهيم الشائعة والمستخدمة في التقارير والوثائق المجمعة. بعد ذلك قمنا باستخلاص ودراسة وتقديم علاقات عامة بين مفاهيم الأنتولوجي المقترح. هذه المفاهيم المقدمة بالإضافة للعلاقات العامة المقترحة بينها قدمت كأول أنتولوجي في مجال عملية المتطلبات في تطبيقات الحكومة الإلكترونية. لقد قمنا بتلخيص الكثير من المفاهيم المستخدمة في تعريف المتطلبات الخاصة بتطبيقات الحكومة الإلكترونية المتداولة في الدراسة والنقاش إلى مجموعة أصغر من المفاهيم تتكون من بعض المفاهيم الخاصة بها, هذه الأنتولوجي المقدمة يمكن أن تستخدم من قبل مهندسي البرمجيات, الباحثين, الممارسين, وأي شخص له علاقة بها بحيث يمكن استخدامها كمنبع معرفة متفق عليه بالإجماع, الأمر الذي يحل مشكلة عدم الإجماع على معاني المفاهيم الدلالية المستخدمة في تعريف أي من الصفات المدروسة التي تعبر عن المتطلبات في تطبيقات الحكومة الإلكترونية. بالإضافة لذلك فإن الأنتولوجي المقدم ممكن أن يستخدم كأساس لتقييم أي تعريف تم تقديمه لمرحلة المتطلبات في تطبيقات الحكومة الإلكترونية.

CHAPTER ONE

INTRODUCTION

1.1 OVERVIEW

In this chapter we review the thesis. A brief background about the scope of the thesis is given; Requirements Engineering (RE), E-Government Application (EGA), and the field of ontology. Then we give an idea about our research problem and how it has been addressed. We end the chapter by giving information about tools used in the work, our own contribution, and the outline of the thesis chapters.

1.2 REQUIREMENTS ENGINEERING

Recently, Requirements Engineering Process (REP) concept has been widely developed to be included in many of our life existing fields; financial, industrial, trading, etc. Requirements Engineering Process in E-Government Applications (REPEGAs) have been created as a matter of applying the e-government concept on the results of requirements engineering process, to help the software engineers to get a unified reference of concepts to use in their gathering and analyzing requirements for developing e-government applications. Requirements engineering is composed of main 4 processes which contain the following concepts: feasibility study, requirements elicitation and analysis, requirements specification and requirements validation.

During the last years, many researchers (individuals and groups) discussed and presented requirements engineering process of e-government application in their works which show that till now there is a lack of consensus on the semantic of many of concepts and terminologies used in the field of REPEGAs. According to above and in more specific our research is focusing on studying the most common REPEGAs concepts and terminologies that current EGAs proposals present to extract a conceptualization for the EGAs, after that we will study this conceptualization in order to build an ontology that produce a coherent and consistent semantics for REPEGAs concepts and terminologies that can be used by software engineers, researchers, practitioners, and stakeholders as a common agreement of REPEGAs pool of knowledge. Before defining the research problem, a brief introduction to the related fields of this research is given.

1.3 THE DISCIPLINE OF SOFTWARE ENGINEERING

Since the dawn of computing in the 1940s, the use of computer software has been rising enormously. Nowadays, computer software plays many important roles, and considered as a way for delivering a product as it is the basis of controlling operating systems, networks, and other applications, and is also considered as products themselves [43]. They serve the human kind in almost all of the fields of government, banking and finance, education, transportation, entertainment, medicine, agriculture, and law [48].

Computer software is a general term used to describe a collection of computer programs, procedures, and documentations that perform some tasks on a computer system [54]. The increasing development of science and technology makes the need for software an important issue especially for software products that is typically a single application or suite of applications built by a software individuals/companies to be used by many customers, businesses or consumers [8].
Software products are categorized under two major types, generic products; which are products developed to be used by any customer in the market, and customized products; which are developed especially to a customer or to a group of customers [52].
The evolving of software development makes developers take a more systematic and planned way to develop their software products, Software Engineering revealed in order to help developers to do so. The IEEE Computer Society defines software engineering as:
The application of a systematic, disciplined, quantifiable approach to the development, operation, and maintenance of software; that is, the application of engineering to software [21].
The term software engineering first appeared in late 1950s and early 1960s.
Programmers have always known about civil, electrical, and computer engineering and debated what engineering might mean for software. The NATO Science Committee sponsored two conferences on software engineering in 1968 and 1969 [44], which gave the field its initial boost. Many believe that these conferences marked the official start of the profession of software engineering.
In the early decades of software engineering revealing, it was motivated to face the Software Crisis problem appeared at that time, researchers and practitioners tried every possible way to solve this crisis (Cost and budget overrun, property damage, and software life and death), In 1987, Fred Brooks published the No Silver Bullet [6] article, arguing that no individual technology or practice would ever make a 10-fold improvement in productivity within 10 years.
Software engineering had been widely affected by the appearance of the Internet, programmers and developers were required to deal with many new issues and merge it within their developed software (images, maps, animations, web browsers usage, etc). Simpler and faster methodologies that developed running and inexpensive software products have been introduced to small organizations in order to satisfy their demands, some of these methodologies are: Rapid prototyping, Agile development, Extreme programming, and others [44].

The need for computer software has grown dramatically, thousands of billions of dollars are spent on the development of computer software. Software products provide us with a more productive, safer, and flexible working environment to help us to be more successful, accurate, efficient, and productive [51]. Despite these successes, Computer Software and Software Engineering face many key challenges such as heterogeneity challenge, delivery challenge, trust challenge, cost challenge, timelines challenge, and quality challenge [52]. Researchers and practitioners are continuously searching to solve these challenges, they solved some of them, and searching to solve others. However this is a good characteristic of the evolving Software Engineering discipline[46].

1.4 SCOPE OF E-GOVERNMENT APPLICATIONS

E-government (E-gov) is defined as a way for governments to use the most innovative information and communication technologies, particularly web-based Internet applications, to provide citizens and businesses with more convenient access to government information and services, to improve the quality of the services and to provide greater opportunities to participate in democratic institutions and processes **[57]**.

From the definition of E-Gov, we can realize the importance of building and using E-gov applications in our life, to facilitate the process of getting specific service for citizens.

So the trend of building and implementing applications for E-gov is increased everyday, and the governments adopting this approach of computerize their services incrementally.

Governments worldwide are faced with the challenge of transformation and the need to reinvent government systems in order to deliver efficient and cost effective services, information and knowledge through information and communication technologies. Development of Information and communication technologies catalyzed and led up to E-government, also e-government presents a tremendous impetus to move forward in the 21st century with higher quality, cost-effective, government services and a better relationship between citizens and government.

One of the most important aspects of e-government is how it brings citizens and businesses closer to their governments. We can outline eight different potential types or models in an e-government system that is useful to define scope of E-government studies: Government-to-Citizen (G2C); Citizen-to-Government (C2G); Government-to- Business (G2B); Business-to-Government (B2G); Government-to-Government (G2G); Government-to-Nonprofit (G2N); Nonprofit-to-Government (N2G); and Government-to-Employee (G2E). Also we can examine some examples in E-government practices and presents a generally-applicable framework for analysis of challenges and problems in E-government development.

The waves of e-government are rising through public organizations and public administration across the world. More and more governments are using information and communication technology especially Internet or web-based network, to provide services between government agencies and citizens, businesses, employees and other nongovernmental agencies. More and more attractions appeal researchers and practitioners come to search for a consensus regarding e-government diagrams and initiatives. E-government may be defined as a continuum from information provision when organizations and public agencies publish static information to the Internet to web interactive communication and E-transactions, and to one-stop integrated virtual governmental services[57].

1.5 WHAT IS AN ONTOLOGY

Ontology as seen from philosophical perspective is the science of studying beings (studying of what is, of the kinds and structures of objects, properties, events, processes and relationships in every area of reality), this term which was coined in 1613 included in many philosophical areas from the metaphysics of Aristotle to the object-theory of Alexius Meinong [50].

Philosophical ontology handles the precise utilization of words as descriptors of entities; it gives an account for those words that belong to entities and those that do not [13]. In both Computer Science and Information Science, an ontology is a representation of a set of concepts within a domain and the relationships between those concepts. It is used to reason about the properties of that domain, and used to define the domain [2].

1.6 WHY DEVELOP AN ONTOLOGY

Recently, the term ontology has been widely included in the field of computer and information science. When building frameworks for information representation of data and knowledge base systems, designers use a wide variety of terms and concepts. Studies showed that there is an inconsistency problem in the semantic of the terms that are used, e.g. identical databases labels are used but with different meanings, and also the same meaning expressed using different names. Methods must be found to resolve the terminological and conceptual incompatibilities [49]. An ontology in this context is a dictionary of terms formulated in a canonical syntax and with commonly accepted definitions designed to yield a lexical or taxonomical framework for knowledge-representation which can be shared by different information systems communities [49].

Ontologies are used in a variety of current fields; Artificial intelligence [18], Software engineering [38], the Semantic web [35], Biomedical informatics [1] , Library science [18], Information architecture [38], Ecommerce content standard [19], and other fields, as a form of knowledge representation about the domain or some part of it.

In this era the presence of consistent global information has become an important issue. In every domain researchers and practitioners need to share information to conduct their works in a professional manner. To do that in a correct way inconsistencies between terms and concepts must be reduced. Ontology defines a common vocabulary for them; it contains machine-interpretable definitions of basic concepts in the domain and relationships among them[46].

From studying the role of ontologies in different knowledge domains, many studies showed that creating and developing and also enhancing ontologies has become important to many fields and areas of domain knowledge, because of its approved advantages effects when using them in the field of the studied knowledge domain.

Many reasons support our recommendation of creating, developing, and using ontologies, some of them are:

• Ontologies support applications (especially distributed ones) to exchange information and to process transactions independently [23].

• Ontologies make the reusing of a domain knowledge possible [12].

• Ontologies provide semantic-aware information systems, which can support enterprise, government, and personal activities at the same time [12].

• Ontologies can share different applications [36].

• Ontologies can use other ontologies [36].

• Ontologies can analyze, support, and enhance domain knowledge [36].

• Ontologies are used as a semantic support representation for many areas [23].

• Ontologies are used to capture the domain information independently of any application requirements [24].

Ontology shows enormous potential in making software more efficient, adaptive, and intelligent. It is recognized as one of the areas which will bring the next breakthrough in software development. The idea of ontology has been welcomed by visionaries and early adopters.

Since 1991, the semantic Web initiative, lead by W3C, has changed the ontology landscape completely, through the initiative, researchers and developers join forces to provide standard semantics markup languages based on XML, ontology management systems, and other useful tools. Also, the Web provides interesting applications of ontology that are critical to daily life such as search and navigation. In addition, people rediscover the value of ontology in other important applications such as information and process integration [73, 74].

1.7 COMPONENTS OF AN ONTOLOGY

Different knowledge representation formalisms and corresponding languages exist for the formalization and implementation of ontologies. Each of them provides different components that can be used for these tasks. However, they share the following minimal set of components [7]:

• Classes represent concepts, which are taken in a broad sense. Classes in the ontology are usually organized in taxonomies through which inheritance mechanisms can be applied.

• Relationships represent a type of association between concepts of the domain. They are formally defined as any subset of a product of n sets, that is:

$R = C1 \times C2 \times \ldots \times Cn$. ontologies usually contain binary relationships. The first argument is known as the domain of the relationship, and the second argument is the range.

• Instances are used to represent elements or individuals in an ontology.

Ontology is an essential data structure for conceptualizing knowledge [53]. It is commonly used as a fundamental structure for capturing knowledge by analyzing relevant concepts and relationships in the area under search [33]. It depends mostly on the analysis of textual data over a collection of text documents by using natural language processing to do that and more such as obtaining semantic graph of a document; visualization of documents; information extraction to find relevant concepts; and visualization of context of named entities in a document collection [53].

1.8 LEVELS OF ONTOLOGY

Different authors like P'erez, Jones, Storre, Robert, Malka, and others have organized ontology in their studies and reports into different levels [63, 90, 94, 95, and 100]:

• Lexical, vocabulary, or data layer. The focus here is on concepts, facts, etc. that ontology included, and the vocabulary used to represent these concepts.

• Hierarchy or taxonomy. An ontology typically includes a hierarchical is-a relationships, or subsumption relationships between concepts.

• Other semantic relationships. The ontology may contain other relationships besides is-a relationship. This typically includes measures such as precision and recall.

• Context level. Ontology may be a part of a larger collection of ontologies. Another form of context is the application where the ontology is to be used

• Syntactic level. The ontology is usually described in a particular formal language and must match the syntactic requirements of that language (use of the correct keywords, etc.). Various other syntactic considerations, such as the presence of natural-language documentation, avoiding loops between definitions, etc., may also be considered.

• Structure, architecture, design. Unlike the first three levels on this list, which focus on the actual sets of concepts, instances, relationships, etc. involved in the ontology, this level focuses on higher-level design decisions that were used during the development of the ontology. This is primarily of interest in manually constructed ontologies. For some applications, it is also important that the formal definitions and statements of ontology are accompanied by appropriate natural-language documentation, which must be meaningful, coherent, up-to-date and consistent with the formal definitions, sufficiently detailed, etc[46].

Let us not forget that ontologies have been applied and played an important role in different areas of Software Engineering fields as they have been playing in other disciplines. They provide a general framework reference of agreed concepts and terminologies among researchers, practitioners, and stakeholders; they enhance collaboration, communication, and knowledge sharing, they represent all assumptions related to the entities and relationships between them that belong to the area under search, and finally they contribute in reducing gabs between researchers, etc created by conceptual confusion [104,106,123].Hence, building an ontology to capture the conceptualization knowledge about Software Quality Attributes domain will achieve a significant successful solution for the semantic conflicts problem the field suffers from[46].

1.9 ONTOLOGY DEVELOPMENT PROCESS AND LIFE CYCLE

The ontology development process refers to the activities that are performed when building ontologies[46]. It identifies three categories of activities as shown in Figure 1.1.

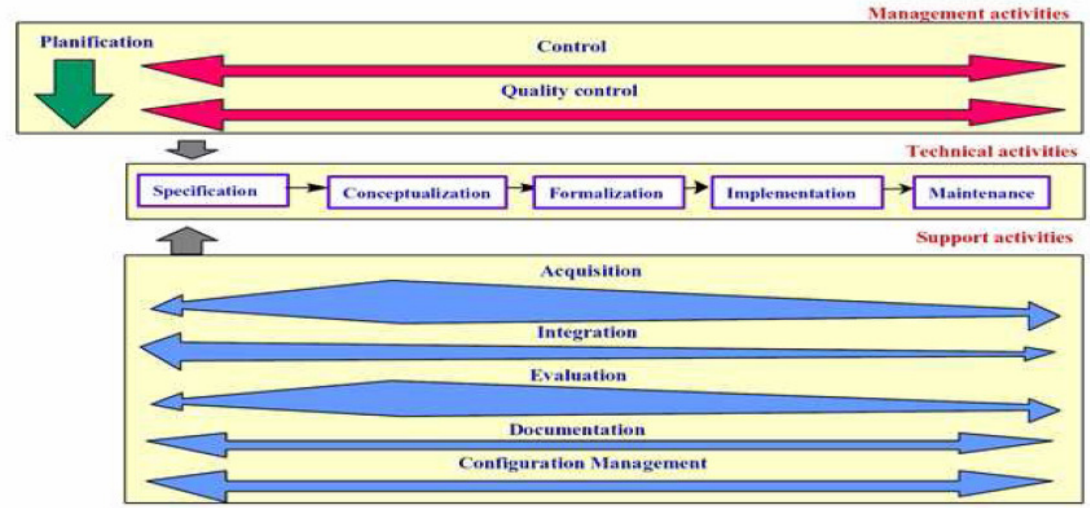


Figure 1.1: Methontology ontology development process life cycle [9].

a) Ontology management activities :

The management process activities are responsible for the project management issues [22, 24, 39]**..**

1. Scheduling is the first activity of the ontology life cycle. The objective is to plan the main tasks to be done, how they will be arranged and the required resources, i.e. people, software and hardware.
2. Control is performed along the whole ontology life cycle in order to survey that they are not undesired deviations from the initial schedule.
3. Quality is responsible for checking that the quality of each methodology output (ontology, software and documentation) is assured.

b) Development Process:

The development process includes all the activities that produce the successive prototype refinement stages towards the desired ontology. The process starts with specification that produces an informal output that then evolves increasing its level of formality, as it passes through the different activities, towards the final computable model, which can be directly understood by the machine [22, 24, 39]. It consists of:

1. Specification:

The specification establishes the ontology purpose and scope. Why the ontology is being built, what are the intended uses and end-users. The specification can be informal, in natural language, or formal, e.g. using a set of competence questions.

2. Conceptualisation:

The objective of this activity is to organize and structure the knowledge acquired during knowledge acquisition using external representations that are independent of the knowledge representation and implementation paradigms in which the ontology will be formalised and implemented next. An informally perceived view of a domain is converted into a semi-formal model using intermediate representations based on tabular and graph notations. These intermediate representations (concept, attribute, relation, axiom and rule) are valuable because they can be understood by domain experts and ontology developers. Therefore, they bridge the gap between people's domain perception and ontology implementation languages.

3. Formalisation: The goal of this activity is to formalise the conceptual model. There are ontology development tools that automatically implement the conceptual model into several ontology languages using translators. Therefore, formalisation is not a mandatory activity.

4. Implementation: This activity builds computable models using ontology implementation languages. There are many ontology languages and they do not have the same expressiveness nor do they reason the same way.

5. Maintenance: This activity updates and corrects the ontology if needed due to the necessities of the current development process or other processes that reuse this ontology in order to build other ontologies or applications.

c) Support Process:

The support activities are performed in parallel with the development-oriented activities [22, 24, 39].

1. Knowledge Acquisition:

First of all, the source knowledge must be captured using knowledge elicitation techniques. The sources of knowledge are listed giving a description and specifying the elicitation techniques used in each case.

The techniques used to extract knowledge from sources can be partially automatic by means of natural language analysis and machine learning techniques.

2. Evaluation:

The evaluation activity judges the developed ontologies, software and documentation against a frame of reference. Ontologies should be evaluated before they are used or reused. There are two kinds of evaluation, the technical one, which is carried out by developers, and users evaluation.

3. Integration, merging and alignment: The integration activity is needed if other ontologies are reused. There are two options when an ontology is integrated in the current ontological framework. First, there is ontology alignment that consists in establishing different kinds of mapping between the ontologies, hence preserving the original ontologies. Second, ontology merging that produces a new ontology from the combination of the input ontologies.

4. Documentation: Documentation details each completed stage and product.

5. Configuration Management: Configuration management records ontologies, software and documentation versions in order to control changes.

1.10 PROBLEM STATEMENT

More recently, In the world of globally dispersed, free/open source software development processes, discovering and understanding the context, structure, activities and contents of requirements engineering process found in practice has been and remains a challenging problem **[31]**.

In the development of E-gov applications which is considered large scale systems, requirements engineering is a very crucial activity. The processes involved in requirements engineering process include domain analysis, elicitation, specification, assessment, negotiation, documentation and evolution. This implies that getting well defined domain for developing an E-gov application is difficult and critical **[56]**. Recent surveys have confirmed the growing recognition of requirements engineering process as a new area of knowledge in software engineering research and practice **[47]**.

Semantics representation techniques, and more precisely using Ontology-based evolution for E-gov applications, each of which has to meet the recognition of requirements engineering activities as of its nature dynamic evolution, conceptual complexity, consensus building and necessity for multiple perspectives, traceability of conceptualization, design scalability, and reusability **[4]** in our E-gov applications.

The need to share understanding between domain specification become a critical issue. The current problem is incapability to cover all domain concepts when we develop E-Gov applications. Therefore, understanding between these domain specification can be improved by building Ontology ideally capture more knowledge about the multiple aspects, concerns, and activities involved in understanding and analyzing E-gov applications **[56]**.

1.11 CONTRIBUTION OF THE THESIS

The main objectives of this thesis are the following:

- Extract concepts used in the semantic of the most common discussed REPEGAs.
- Extract general relationships between the extracted concepts.
- Introduce the extracted concepts and relationships as an ontology for the domain of REPEGAs.
- Introduce ways in order to use the provided ontology to solve the semantic inconsistency problem found in the field of REPEGAs.

The main aim of this research is to develop ontology for ideally capturing the analysis domain knowledge about the multiple aspects, concerns and activities involved in requirements engineering process for E-gov applications. This contribution aims at enabling software engineers to find out shared understanding, common concepts and terms for describing requirements analysis domain practices by different requirements domain models used in software development in E-gov applications.

A number of issues we addressed in this research:

- How could ontology affect software development process, in particular, software requirements process for E-gov applications?
- How ontology is built for Software analysis domain?
- How is this ontology evaluated in a specific domain involved in requirements engineering process for E-gov applications?
- How will ontology help software engineers in capturing knowledge about analysis domain process?
- What is the overall role of ontology in software engineering process, in particular software analysis process?

This thesis aims to address the needs of two main kinds of interested audiences:
• The first kind are the e-government applications researchers and standard developers (e.g., international standardization institutes and committees), who is responsible for producing concepts, terms, and standards in the field.
• The second kind are the e-government applications practitioners, who may be confused by the terminology differences and conflicts in the existing standards and proposals when they would use them in their works.

REPEGAs discipline is considered in the emerging phase, and it suffers from the typical symptoms of any relatively evolving disciplines. REPEGAs are currently in the phase in which terminologies, principles, and methods are still being defined, consolidated, and agreed. In particular, there is a lack of consensus on the concepts and terminologies used in the semantic of this field. Studies showed that inconsistencies in the semantic used different research proposals often occur [24, 39].
In our research we focused on studying REPEGAs concepts and terminologies that current SWE of EGAs proposals, documents, and reports present. We prepared text corpora from them to be used in tools to extract the most discussed and used concepts from it. After that experts (doctors and professors in the field of SWE and EGAs) were asked to study and filter the resulted concepts and provided them to us.
An evaluation phase depended on a coverage technique was done to the resulted concepts, followed by an enhancing step to the evaluated ontology domain concepts which leaded us to increase the number of the suggested concepts in the ontology domain, after that a coverage evaluation is done again to the new suggested ontology domain concepts.
In order to extract general relationships among the suggested ontology domain concepts, we returned to the prepared text corpus again and ran out two tools on it. We studied them, filtered them, listed them and represented part of them using a lattice representation.
After we have finished our research steps, and depending on the results we had, we claim that we have presented the conceptualization of the common discussed REPEGAs by an ontology. According to the results of the suggested ontology, we also claim that we condensed the semantic of thousands of concepts used to define any of the discussed REPEGAs into a smaller set of concepts , and that will help experts, software engineers, researchers, practitioners, and stakeholders in the field

of REPEGAs to share and use a common and agreed semantic of concepts when defining any of the studied requirements of EGAs, and that will lead us to resolve the inconsistencies of the semantic appeared among documents and reports that define any of the studied requirements of EGAs.
In addition to this, our ontology provides a base to evaluate any related presented definition semantic for one of the studied requirements. The way of doing that is if a high percentage of the concepts used in the semantic of the presented definition are covered by our ontology domain, the presented definition semantic can be accepted, but if not we claim that it is a weak semantic to be used to define such requirements for EGAs.

1.12 THESIS ORGANIZATION

- This thesis is organized into the following chapters:

Chapter 2:
Chapter two will give a brief idea about the most relevant work in the literature that is related to our study, in both general and specific related work.

Chapter 3:
Chapter three talked about the preparation of the text corpora for the REPEGAs knowledge domain. This preparation is done by collecting and studying a large number of documents and reports related to the field of E-Government Applications like as (E-Health, E-Tax, E-Law, E-Tourism....etc). The chapter also is discussing how the prepared text corpora were used to extract and create our primary ontology domain concepts using TextToOnto tool with support of an MS Access tool and then with support of human experts. In this chapter, we also focused on evaluating the suggested ontology domain concepts using a coverage methodology. After preparing the needed corpus, and by using a tool created by Kayed [26], we counted the covered concepts and calculated a coverage percentage for them.

Chapter 4:
Chapter four will extract general relationships between the new concepts of the suggested ontology domain after studying and filtering the results of two tools. Then it presented the relationships as groups. then, a general lattice representation for part of the resulted relationships was done.

Chapter 5:
Chapter five presented and discussed the conclusions of our research; the final results and how we used them to contribute to the studied domain are presented among the conclusions. Future work is suggested at the end of this chapter.

CHAPTER TWO

RELATED WORK AND METHODOLOGY

2.1 OVERVIEW

This chapter gives a brief idea about the most relevant work in the literature that is related to our study.

2.2 GENERAL RELATED WORK

With the appearance of knowledge representation techniques, semantics evolution and more precisely using ontologies in specific domains such as requirements engineering, the need for compromises between domain models expressiveness and precision has become more and more challenging for software engineers.

Ontology has widely been used during the last years. Researches related to software development and requirements engineering issues, in particular, building ontologies for requirements engineering using potential elements (such as goals, viewpoints, data, operations, agents, scenarios and resources) has been carried out. These studies and contributions illustrated below:

[Dardenne 91]: On Formal Requirements Modeling Languages: RML Revisited
Has proposed a formal framework for integrating goals and goals refinement in requirements models. This framework gave raise to KAOS methodology for eliciting, specifying and analyzing goals, requirements scenarios, and responsibility assignments.

[Rumbaugh 91] : An Introduction To Multi-Paradigm Modelling And Simulation
Has proposed multi-paradigm frameworks to combine multiple languages in a semantically meaningful way so that different facets can be captured by languages that fit them best.

[Mylopoulos, 92]: Intentions and Agents, From Entities and Relationships to Goals and Agents
Has proposed a qualitative framework. This framework is introduced to relate soft goals; these goals are rarely to be satisfied in a clear cut sense in requirements models. This framework gave rise to NFR methodology for capturing and evaluating alternative goal decompositions.

[Manna 92, Koymans 92]: The Temporal Logic of Reactive and Concurrent Systems
Have proposed an optional formal assertions layer to support various forms of formal reasoning. Goals and requirements on objects are formalized in a real-time temporal logic.

[Potts 94, Rumbaugh 91, Rubin 92, Sutcliffe 97, Dubios 93, Fickas 92, Hsia 94]: Object-Oriented Modeling and Design
They all puts much effort on using Scenario-based technique for elicitation and validation requirements in hypothetical situations. This technique helps identifying exceptional cases and also populates more abstract conceptual models.

[Ross and Schoman, 97]: Structured Analysis (SA): A Language for Communicating Ideas
They have introduced SADT as a specific modeling technique. This technique was a precursor in many aspects. It supports multiple models linked through consistency rules- a model for data, in which data is defined by producing/consuming operations; a model for operations, in which operations are defined by input/output data, and a data/operation duality principle.

[Weidenhaupt, 98] Scenario Usage in System Development: A Report on Current Practice
He has confirmed in a recent study that scenario-based reasoning technique is an important artifact used for a variety of purposes, in particular, in cases when abstract requirements modeling fails.

[Devedzic, 2002]: Understanding ontological engineering
Has explored ontologies which are needed in all phases of software engineering lifecycle, each of which must have knowledge, whether to data structure, methods or domain. This makes ontologies everywhere and they make it possible to smoothly integrate artificial intelligence with other software disciplines.

[Zlot, 2002]: Modeling task knowledge to support software development
Has defined a structure to represent the task knowledge with support to software engineers in understanding business problems starting from the understanding of the task, which comprises these problems. This structure combines task ontologies and problem solving methods to support capturing knowledge about specific domain throughout the development process.

[Obrst, 2003]: Ontologies for semantically interoperable systems
Has discussed the use of ontologies for semantic interoperability in homogeneous environments.

2.3 SPECIFIC RELATED WORK

In addition to previously presented related work, we separated and illustrated more specific related work to software quality in a dependent subsection, because of its important role in our work.

[Kayed, 2005]: Building e-laws Ontology: New Approach

Has discussed a new approach of building Ontology for e-Laws.

First, the author described the Ontology Life Cycle, as illustrated in Fig 2.1:

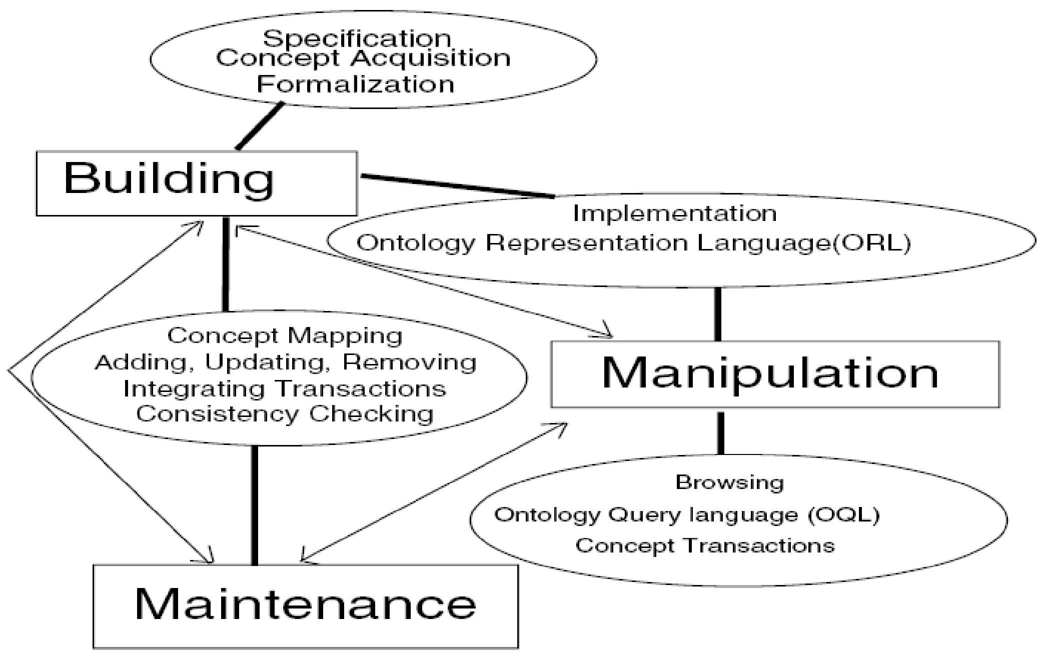


Figure 2.1: Ontology Life Cycle.

The author summarizes the methodologies for building ontologies around three major stages of the ontology life cycle, as follow:

- Building Stage: In this stage, four steps are needed: specification, conceptualization, formulization, and implementation.
- Manipulation Stage: In this stage, an ontology query language should be provided for browsing and searching; efficient lattice operation; and domain specific operations.
- Maintenance Stage: In this stage, developers should be able to syntactically and lexically analyze the ontology, adding, removing modifying definitions, and also translating from one language into another.

Ontology serves as an abstract data type for concepts in domain, therefore an experiment has been conducted to extract concepts for e-law ontology. Text-mining tools have been used to extract concepts in the domain of e-commerce laws. The following summarizes the steps to build this ontology[25]:

- Collect many law cases for e-commerce.
- Extract top concepts.
- Refine the results.
- Categorize the concepts.
- Define the relationships among concepts.
- Build the ontological hierarchy.
- Formalize the concepts.

[Dritsas, Gymnopoulos, Karyda, Balopoulos, Kokolakis, Lambrinoudakis and Katsikas, 2006]: A knowledge-based approach to security requirements for e-health applications

They introduced a knowledge-based approach to security requirements analysis and design for e-health applications.

The design of the security patterns follows the development of a security ontology; this is shown in Figure 2.2. The development of the security ontology was carried out in the following phases:
During the first phase, a set of questions (called competency questions) was determined. These are loosely structured questions, indicating the type of answers and information we would like to receive when using the ontology. For the purposes of this work, the focus when building the ontology has been the area of e-health applications. This process enabled us to identify the important concepts within the e-health domain and the corresponding terms. The latter served as the basis for the formation of ontology classes (second phase). During the first phase, a large number of relevant terms was identified and recorded. Based on their relevancy to the e-health domain a subset of them was selected to form the ontology classes; other terms formed the properties of the classes; some terms were excluded as irrelevant. The terms used as ontology classes are the following: Stakeholder, Objective, Threat, Countermeasure, Asset, Vulnerability, Deliberate attack, Security Pattern and Security Pattern Context. The third phase involved drawing the relations among the ontology classes and deciding upon their hierarchy. To reach this decision two approaches could be used: a top-down approach, where general concepts are included first and are later specialized; a bottom-up approach which suggests that specific classes are defined first and are then grouped into general concepts. We used a combination of the two approaches for designing the ontology. Finally, in the fourth phase we provided value types and allowed values or cardinality for the class properties (called slots) and the slot properties (called facets); this process is called instantiation of the ontology. These four phases are repeated several times until we get the final ontology.

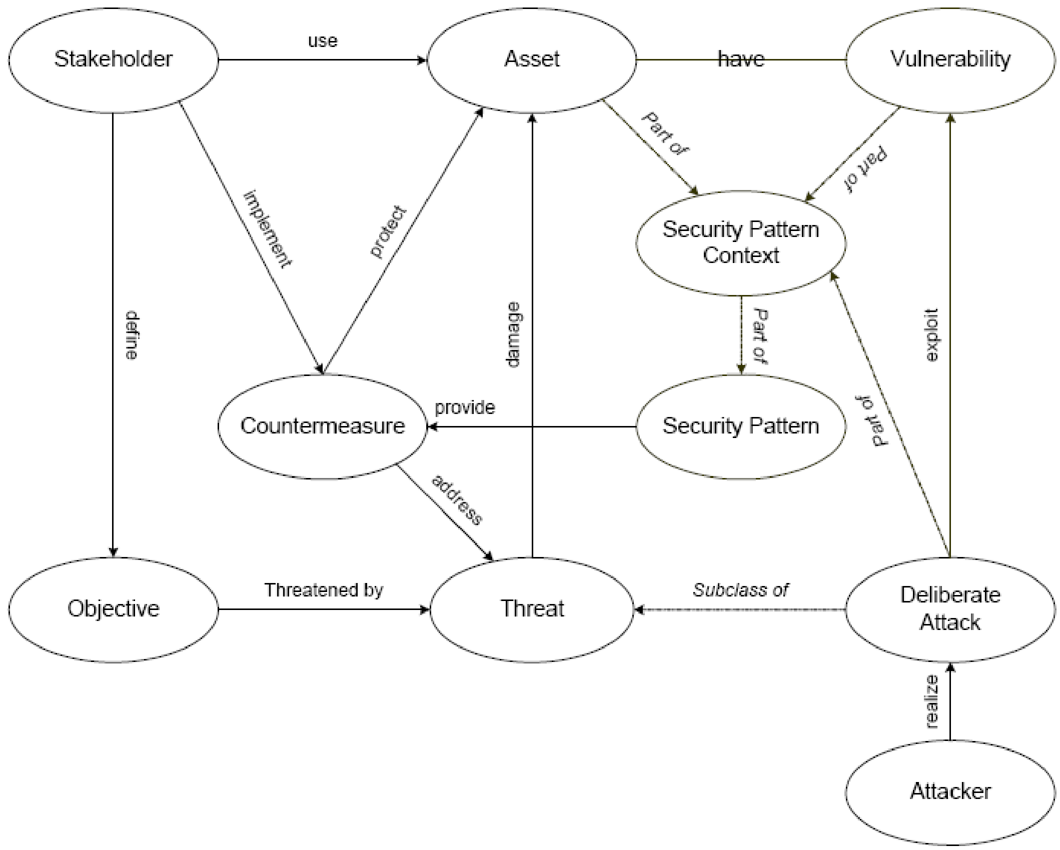


Figure 2.2: The Security Ontology.

[Herborn and Wimmer, 2006]: Process Ontologies Facilitating Interoperability in eGovernment A Methodological Framework

They introduced the approach to develop a common BRITE (Business Register Interoperability Throughout Europe) domain ontology which links up national domain ontologies and BR processes.
The project BRITE aims to build interoperability between the BRs in order to facilitate EU (European Union)-wide transactional services for companies to e.g. register their branch in another country. These goals will be achieved by the application of ontology-driven semantics.

They introduced a methodological framework to secure real interoperability between different institutions, even with language barriers and massive process diversities. The approach has the advantage of not having to change legacy systems, but to link them up via standardized overall, domain and process ontologies. It consists of the following steps:

– Defining a High Level Domain Ontology (HLDO): as shown in Fig. 2.3

- Identifying Domain ontologies of interest
- Selecting subset of vocabulary of Domain Ontology
- Merging subsets
- Mapping merged subsets to HLDO
- Adding semantics to the HLDO
- Provide interoperability layer through the use of the HLDO

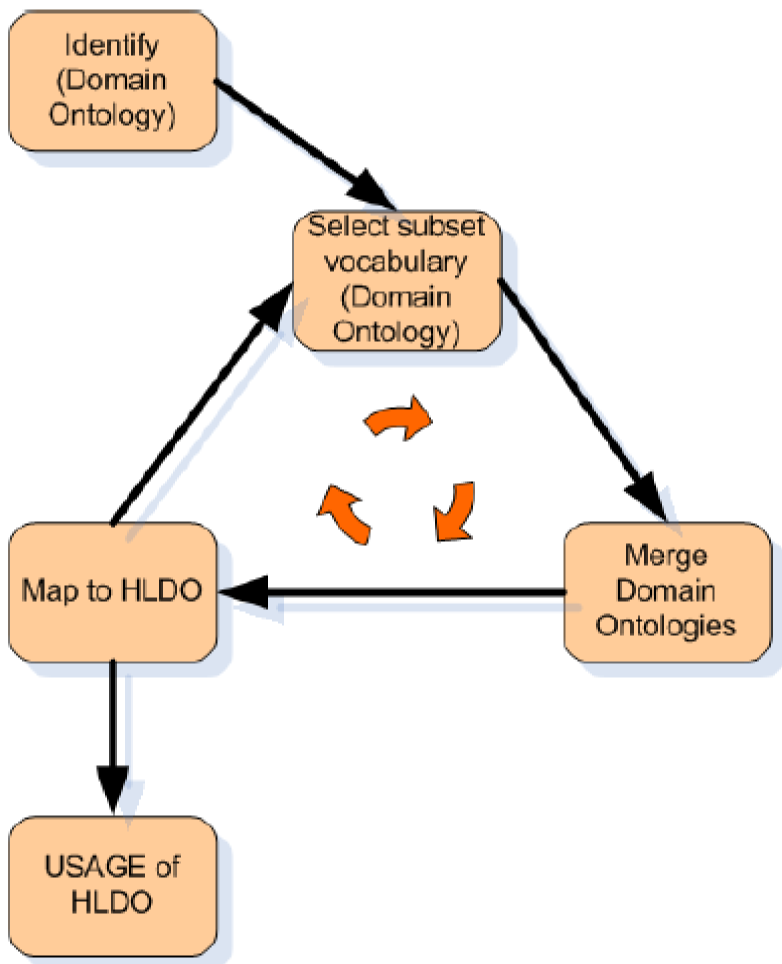


Figure 2.3: Methodology to define HLDO

– Identifying national BR process ontologies (BRPO)
– Developing high level BR process ontologies
– Integrating national processes

This research is based on the idea that a general eGovernment ontology is considered as top level Domain Ontology, as illustrated in the following figure 2.4.

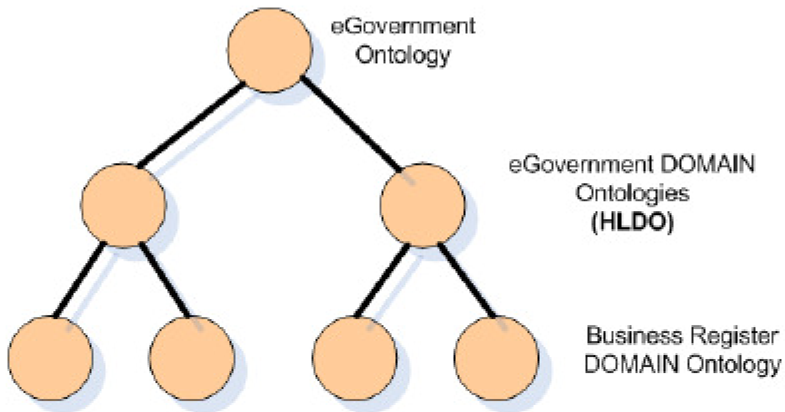


Figure 2.4: Hierarchy of Ontologies within eGovernment

[Moulin, Bettahar, Barthes and Sbodio, 2007]: Ontology Based Categorization in eGovernment Application

They discussed the Ontology Based Categorization in eGovernment Application.

They proposed a method for automatically classifying instances of concepts in knowledge bases. In this categorization, instances may themselves belong to a category defined by a rule or may be associated to specific instances or concepts defined in an ontology.

The solution they proposed is based on the definition of an ontology of the social care domain. The ontology is designed to be inserted and used by different modules of a platform allowing civil servants to deliver services to citizens. The ontology was built by several groups of experts following a classic methodology. In this sense we applied the principle that terms must be accessible to future users. We added a mechanism allowing to build indexes from all important information found in the ontology.

An ontology representation structure may be very expressive, but that may lead to difficulties when reasoning. Conversely, a restriction of the knowledge representation allows better control of the reasoning task, but may prevent building some concepts. It is thus necessary to follow a method which preserves the expressive nature of the ontology and insures the control of the reasoning task. They chose to extend the ontology with rules in order to create new assertions in knowledge bases.

Rules allow engines to create the assertions that classify instances in knowledge bases without modifying the structure of the ontology.

This is very important in the e-Government domain when laws change.

2.4 METHODOLOGY

This research will be carried out through a theoretical and an empirical study. Our approach of the problem is divided into six steps as shown in Figure 1.2:

- The first step of this research is a literature review on almost all existing proposals and ontologies in requirements engineering process, with the focus on a specific domain concerning with the software requirements of e-government applications domain. This review presented, discussed, and analyzed different sources for requirements engineering process in general and for requirements engineering process of e-government applications in particular, such as researches, reports, documents, and proposals produced by various individuals, institutes, and committees in the field.
- The second step of our work focused on paving the way to capture and extract the ontology domain concepts from the knowledge domain prepared in the first step using some tools. Later a support from experts in the field to study and filter the results was asked.
- The third step of our work handled the evaluation of the resulted ontology domain concepts, by following a technique categorized as a coverage approach in the domain.
- In the fourth step, enhanced results were reached depending on the results of the evaluation step.
- In the fifth step, we captured and extracted general relationships between the suggested ontology domain concepts, by providing the prepared knowledge domain to two tools, after that we studied and filtered the resulted relationships into groups, a general lattice representation to a part of the resulted relationships was constructed.
- In the sixth and the final step, we showed how the results contribute to the domain, and suggested many ways to use them in order to reach to a common, shared, and agreed semantic when defining any of the studied requirements of e-governments applications.

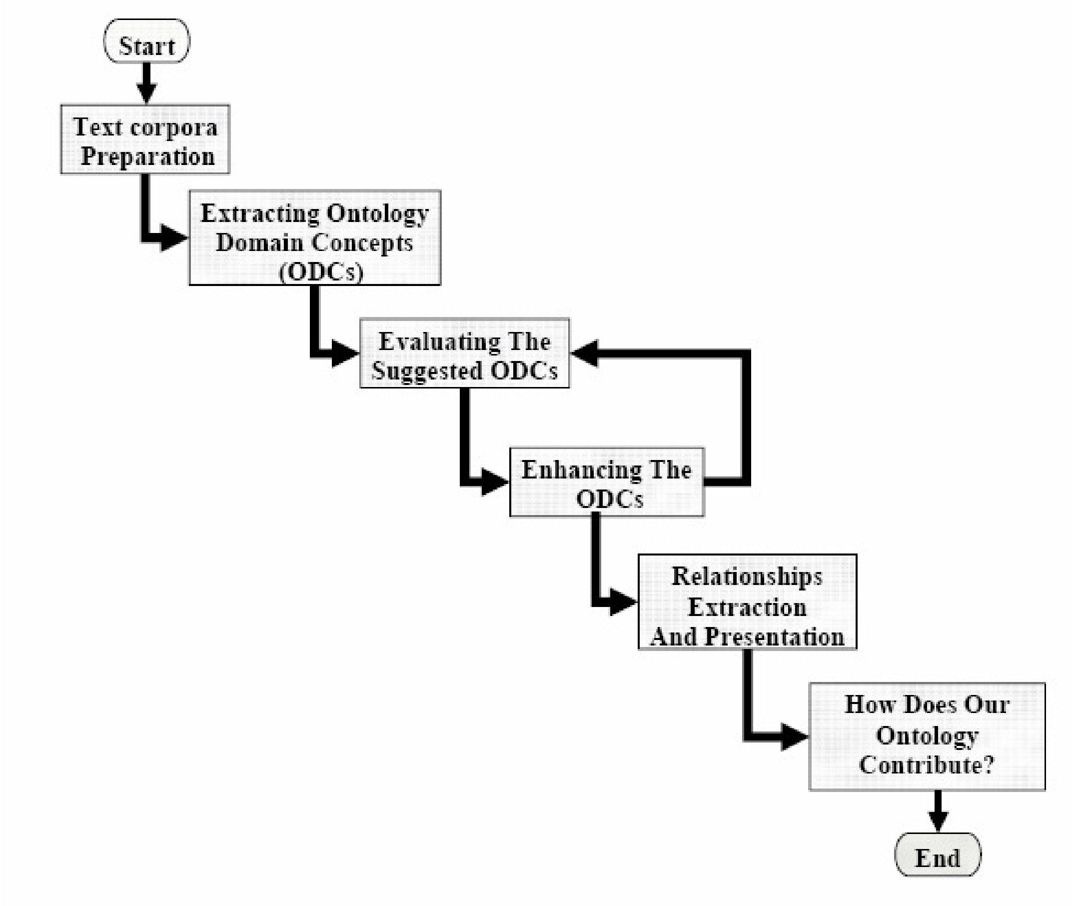


Figure 2.5: Steps of the methodology of the study

2.5 SOFTWARE USED IN THE WORK

In our work we used many tools in order to reach to some necessary results, below is a brief description of those tools used in this work:

2.5.1 KAON

KAON consists of a number of different modules providing a broad bandwidth of functionalities centered around creation, storage, retrieval, maintenance and application of ontologies. It was and currently is being further developed in a joint effort mainly by members of the Institute of Applied Informatics and Formal Description methods (AIFB) at University of Karlsruhe and the Forschungzentrum Informatik (FZI) – Research Center for Information Technologies, Karlsruhe [29].

The KArlsruhe Ontology [29] and Semantic Web tool suite a.k.a. KAON Tool Suite is an open source ontology management infrastructure. However, there exist also external components which support functionalities such as e.g. ontology learning from texts. An overview of the KAON Tool Suite and its main components; KAON, KAON Extensions and TextToOnto, is presented by Figure 2.6.

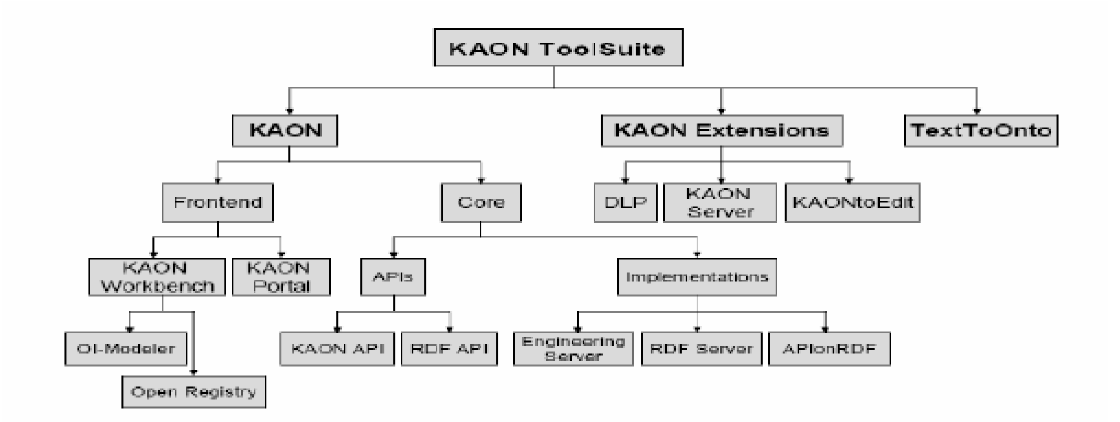


Figure 2.6: An overview of the KAON Tool Suite and its main components; KAON, KAON Extensions and TextToOnto [29].

KAON (consisting of KAON Frontend and KAON Core) includes a variety of different modules for ontology creation and management. The Frontend is represented by two applications developed in order to be used particularly by human users:

• KAON Workbench: provides a graphical environment for ontology based applications. It includes the OI-Modeler – a graphical ontology editor - and the Open Registry (a.k.a. ontology Registry), which provides mechanisms for registering and searching ontologies in a distributed context.
• KAON Portal: is a simple tool for multi-lingual, ontology-based Web portals.
The Core of KAON supports programmatic access to ontologies by including both APIs and implementations for managing local and remote ontology repositories [14].

KAON Extensions are a collection of optional components not included in the standard distribution of KAON [14].
• DLP (Description Logic Programs) support efficient ontology reasoning by mapping Description Logic into Logic Programs.
• KAON Server can be considered as Application Server for the Semantic Web, which provides a generic infrastructure to facilitate plug'n'play engineering of ontology-based applications.

• KAONtoEdit is a plug-in for OntoEdit [39], which allows working directly on implementations of the KAON API in order to load, modify and store KAON ontology models.

• TextToOnto is a KAON-based tool suite supporting the ontology engineering process by providing a collection of independent tools for ontology learning and maintenance.

In Our work we focused on using the TextToOnto Extension because of its capability to help users to learn about ontologies from a provided text.

2.5.2 TextToOnto

TextToOnto [32] is a tool suite built upon KAON in order to support the ontology engineering process by text mining techniques. Providing a collection of independent tools for both automatic and semi-automatic ontology extraction. it assists the user in creating and extending OI-Models. Moreover, efficient support for ontology maintenance is given by modules for ontology pruning and comparison. In particular, the current distribution of TextToOnto comprises the following tools:

- TaxoBuilder: for building concept hierarchies
- TermExtraction: for adding concepts to an ontology
- InstanceExtraction: for adding instances to an ontology
- RelationExtraction: for semi-automatic learning of conceptual relationships
- RelationLearning: for automatic and semi-automatic relationship learning
- OntologyComparison: for comparing two ontologies
- OntologyPruner: for adapting an ontology to a domain-specific corpus

Figure 2.7 shows the front-end of the TextToOnto tool as an extension of KAON tool.

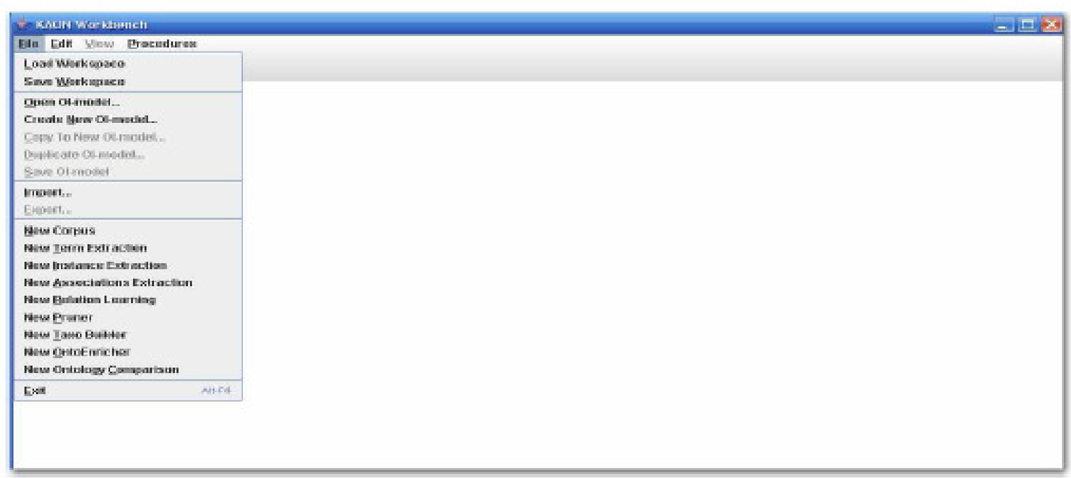


Figure 2.7: The Front-end of the TextToOnto tool as an extension of KAON tool.

2.5.3 MS ACCESS AND MS VISUAL BASIC TOOLS

MS Access and MS Visual Basic have been used to implement algorithms. Screen shots of the program are provided in Figure 2.8, for further reading about it you may refer to [71, 72].

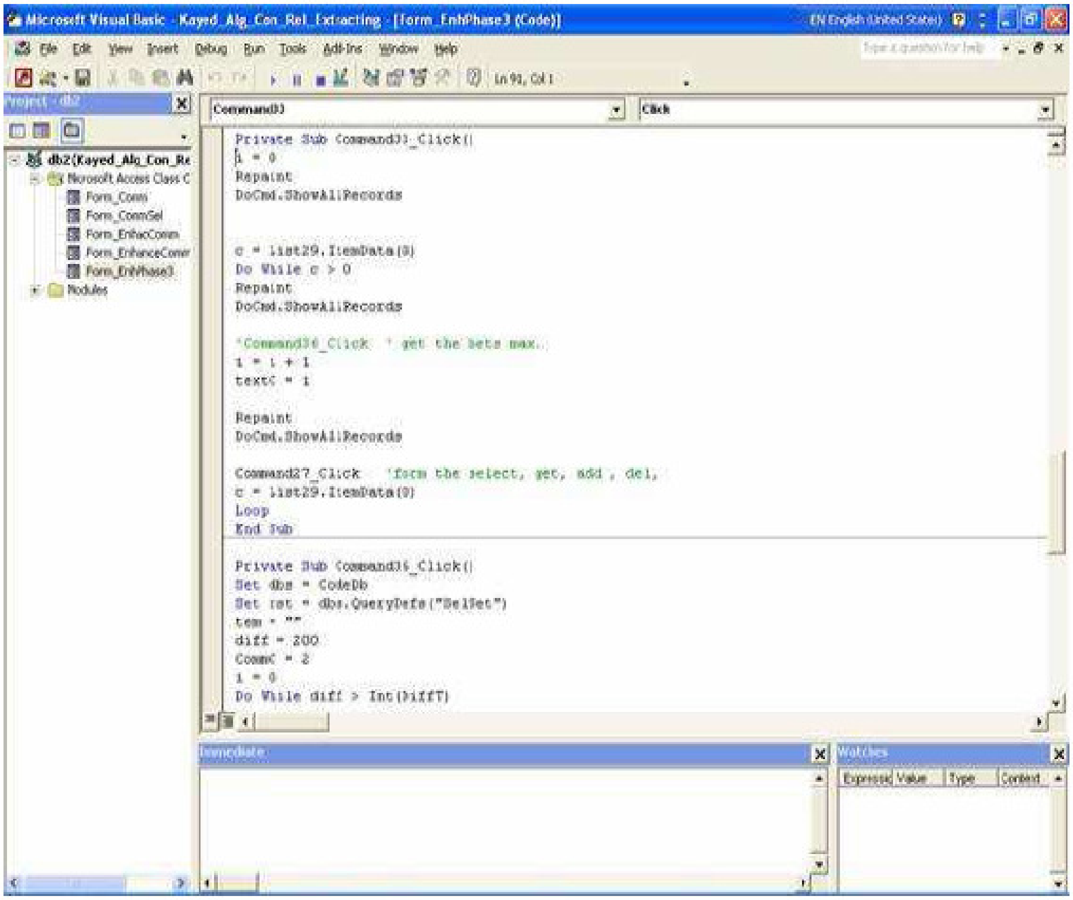


Figure 2.8: Part of the algorithm used in the MS Visual Tool.

CHAPTER THREE

BUILDING ONTOLOGY DOMAIN CONCEPTS

3.1 OVERVIEW

In this chapter we are talking about the preparation of the text corpora for the REPEGAs knowledge domain. This preparation is done by collecting and studying a large number of documents and reports related to the field of e-government applications. The chapter also is discussing how the prepared text corpora were used to extract and create our primary ontology domain concepts using TextToOnto tool with support of an MS Access tool and then with support of human experts. In this chapter, we also focused on evaluating the suggested ontology domain concepts using a coverage methodology. After preparing the needed corpus, and by using a tool created by Kayed [26], we counted the covered concepts and calculated their coverage percentage.

3.2 PREPARING TEXT CORPORA FOR REQUIREMENTS ENGINEERING PROCESS IN E-GOVERNMENT APPLICATIONS DOMAIN

As mentioned earlier, TextToOnto is a tool provided for ontology engineering process depending on text mining techniques and natural language processing algorithms [32]. To use this tool we needed to prepare text corpora, in linguistics, text corpora consists of large set of electronically processed and stored texts. They are needed when doing statistical analysis, checking occurrences, or validating linguistic rules on a specific domain. TextToOnto tool deals with corpora of text or html type.

For our research text corpora was prepared to be used within the TextToOnto tool and later within an Access tool, e-government applications relevant domain documents, reports, and publications were collected. In our case, we collected as much as possible of what we could reach to of publications, documents, and reports that were related to the field, almost about 68 different related documents to requirements engineering, and e-government applications were collected. We believed that in such a large collected domain, heterogeneous and homogenous text collection, concepts, and terms can be found. Upon the discussion of REPEGAs and their definitions, a more deep study was conducted to these collected documents and they were filtered into many related documents, reports, and publications. After that from these resulting files we created a document containing a summary from their semantic. Later, we converted them into text files. By that our text corpora for the REPEGAs domain were ready, the corpus which consisted of the documents were entered into the TextToOnto tool and the corpus which consisted of the summary was entered into the Access tool later on. The following discussion shows the details of how they had been used.

3.3 EXTRACTING ONTOLOGY DOMAIN CONCEPTS

Ontology domain concepts extraction is considered the most important part in building an ontology. In order to extract ontology domain concepts we must study the semantic of the prepared text corpora. To do so, at first we used the TextToOnto tool [32]. We added the prepared text corpus (from related documents) to the tool by using the New Corpus function. Figure 3.1 shows the creation and addition of the prepared corpus to the tool.

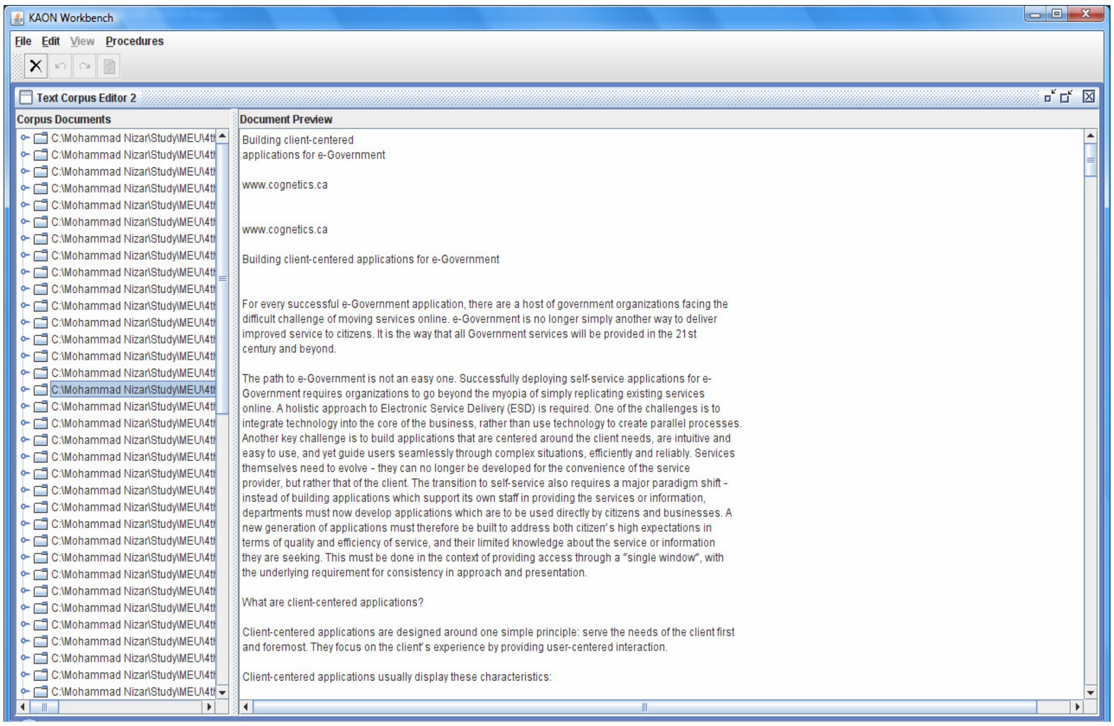


Figure 3.1: Creating a Corpus using TextToOnto Tool.

Later we used the New Term Extraction function in order to extract concepts from provided text corpus. This tool depends on natural language processing algorithms in addition to semantic lexicon filtering techniques. When decided to declare parameters to be used in the tool, at first the used frequency threshold was set to be 5 and above but the result included more than 2000 concepts and that was very large to be considered as an initial result for the ontology domain concepts; it was difficult to be handled, so the declared was 10, 15, and 20 as frequency thresholds to be taken; from the results that came; the chosen option is to stick with retrieving concepts that their frequency in the given text corpus were 10 frequencies or above, also the choice was to retrieve concepts that consist on one unique word as a term; to have a suitable number of concepts (not too large and also not too small) to be collected and studied in our work. Figure 3.2 shows this step and some of the resulted concepts.

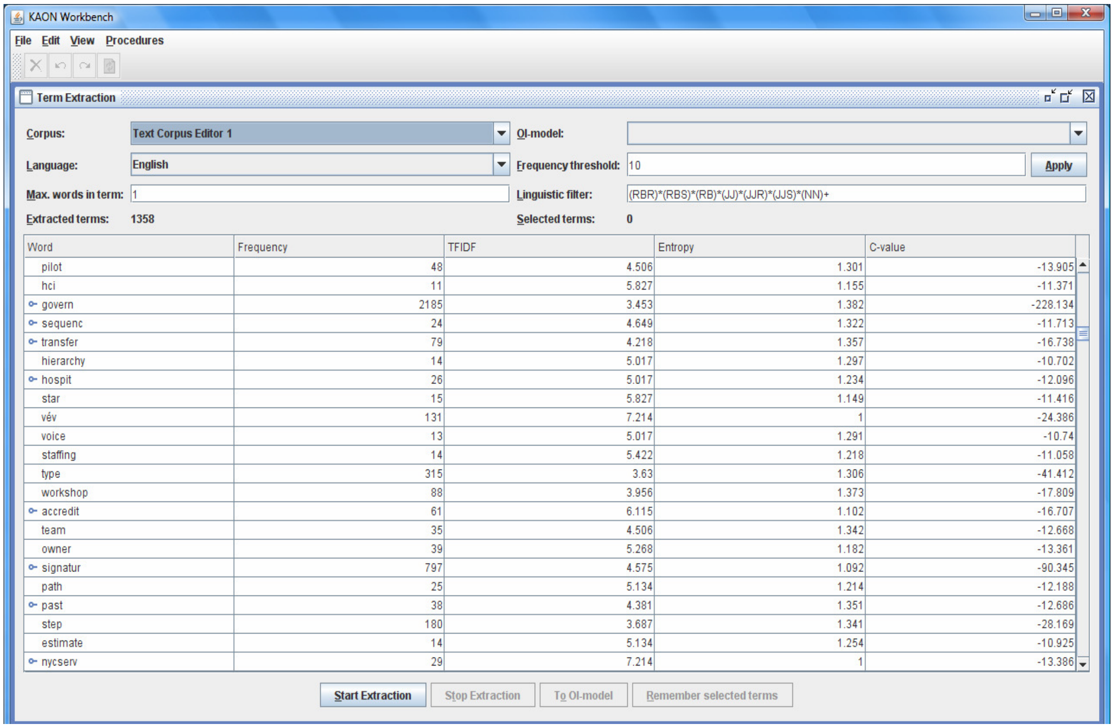


Figure 3.2: The Term Extraction process using the TextToOnto tool.

This step; using TextToOnto tool to extract concepts, provided us with about 1358 single concepts having a 10 as frequency of appearance in the given text corpus.

After that and in order to refine these resulting concepts we used a tool created by Kayed [26], it is a combination of MS Access tool and MS Visual Basic language. We provided it with the resulted concepts (1358) and with the other previously prepared text corpus (text corpus from the abstract file), It depends on a semantic counting algorithm that counts the unique frequencies of the concepts in a given set of texts, so by using this algorithm it studied which concepts from the provided 1358 concepts were found in the semantic of the provided corpus and how many times? This tool provided us with almost 90 single concepts. Table 3.1 below lists the resulting concepts from the tool.

Table 3.1: the resulted 90 concepts (out of 1358).

Concept	Frequency	Concept	Frequency	Concept	Frequency
adoption	44	etax	52	scope	59
broker	47	principle	52	official	17
treatment	14	tax	855	function	178
democracy	43	vote	655	confidentiality	67
client	132	categorization	18	channel	48
penetration	29	elicitation	13	consistency	18
feature	23	policy	234	accuracy	32
web	899	standard	119	awareness	50

vision	39	inspection	13	abuse	35
intrusion	53	competence	23	purpose	91
esignature	84	clarity	10	stakehold	58
infrastructure	153	approach	401	relationship	93
attack	51	terminology	10	eprocurement	37
court	29	safety	17	survey	245
stock	11	draft	26	election	122
citizen	174	confirmation	11	deployment	65
decryption	14	enhancement	12	efficiency	65
enotification	12	intelligence	30	semantics	46
interact	124	reliability	42	trust	154
risk	85	classification	32	strategy	140
sector	325	metrics	20	ehealth	21
scalability	10	standardisation	49	elearning	14
pattern	73	author	321	threat	34
research	398	feasibility	12	evaluation	100
security	671	investment	68	plan	94
copyright	26	database	172	law	172
solution	207	aim	31	knowledge	358
etourism	15	factor	39	identification	182
improvement	30	govern	2185	tender	10
diagram	30	middleware	11	scope	59

After we had the resulting 90 concepts from the used Access tool, we reached to the final part of extracting the ontology domain concepts. We took those concepts, and applied an elimination process for the stopping words (extremely common words like use, can, the, of, etc) from them. The concepts set resulting from the elimination process was sent to human experts (professors, doctors, and practitioners) in the field of RE and EGAs, and we asked them to help us in condensing the set of concepts into a smaller one. After a while the results were sent back to us, we collected them, studied them upon the agreement of all experts on the sent concepts; they all considered them related and important to the studied field, and merged them into one set of concepts. The result of this part was 25 concepts, which we suggested as an ontology domain concept.
Table 3.2 below shows the suggested ontology domain concepts.

Table 3.2: The suggested 25 ontology domain concepts.

Concept	Concept	Concept	Concept	Concept
adoption	security	classification	confidentiality	trust
client	elicitation	investment	accuracy	strategy
esignature	policy	database	awareness	evaluation
citizen	standard	govern	purpose	knowledge
sector	reliability	scope	stakehold	identification

An evaluation process for the suggested ontology domain concepts must be done in order to see how much the obtained concepts belong to the knowledge domain of e-government applications, and to know if the resulting concepts are enough to build a good ontology from them or not. That what we are discussing in the next chapter.

3.4 ONTOLOGY EVALUATION PROCESS

Various methodologies to evaluate ontologies have been presented in the last decade, most of them belong to one of the following categories:

• Evaluations based on using the ontology in a context of an application or project, to evaluate how effective it is. The use of the system may reveal weakness or strength points in the ontology [3]. For our research it is hard to build an application in order to be used considering the time we have.

• Evaluations based on the effort done by human experts, who try to assess how well the ontology meets a set of predefined criteria, standards, and requirements [42]. To reduce the role of human intervention in our work especially after we depended on human experts when extracting the suggested ontology domain concepts, we did not use this approach to evaluate the suggested ontology domain concepts.

• Evaluations based on comparing the ontology with other ontologies in the same domain [5]. As we declared earlier, our ontology is presented as a first in the specific domain of REPEGAs, so we could not use this approach for evaluation.

• Evaluations based on studying ontology relationships considering some criteria [5]. For our ontology we extracted and presented general and basic relationships between the extracted concepts from the domain, it is not adequate to be evaluated using this approach.

• Evaluations based on studying and comparing the formal representation of the ontology with other ontologies formal representations, criterias, or measures [53]. As mentioned earlier, our ontology is presented as a first in the specific domain of REPEGAs, so we could not use this approach for evaluation.

• Evaluations based on fitting or covering techniques between an ontology and a domain of knowledge that the ontology is created for [16, 25].

The last methodology; the coverage methodology, can be decomposed into two different coverage approaches:
The first is done by comparing the new ontology domain concepts with a considered existing gold standard domain concepts, to see how much does the studied domain fit in the resulted ontology.

The second approach is done by comparing the ontology domain concepts with concepts of a prepared knowledge domain to see how does the suggested ontology concepts cover from the studied domain concepts.

For our research we used a coverage technique; the requirements for this approach are available (text corpora, tools, etc). We combined the two last approaches of the coverage methodology. We prepared a corpus that combined the semantic of golden standards and the semantic of EGAs knowledge domain. From many related documents, reports, and publications we extracted the semantic of the most common discussed REPEGAs and their various discussed definitions in those files. We reached to almost 25 REPEGAs and a wide range of definitions for them.
The complete extracted 25 REPEGAs and their definitions in addition to the sources they are taken from are presented in Appendix A.

As shown in the table of Appendix A, the most common and discussed 25 concepts of e-government applications were extracted from the studied knowledge domain found in various documents and reports we collected earlier. The table shows that every term has many definitions which come from many sources. If we take a deep look at them, we will see inconsistencies on the semantic of the used concepts. That really makes the researchers in the field confused about those concept semantic. After completing the ontology that we aim to build through our work, we will show how to use it in order to solve the problem which appears from using various semantics in the definitions of any of the studied concepts.

3.5 THE COVERAGE PROCESS AND THE EVALUATION RESULTS

After we prepared the text corpus for the evaluation process (as seen in the table before), we used two tools to help us in conducting the coverage technique. First, for each e-government applications requirement definition(s) we extracted its single and unique concepts using the TextToOnto tool as used before. After that we eliminated the stopping words from the resulting concepts.

Later, in order to know how much our suggested ontology concepts cover from each requirement definition concepts which were extracted earlier, we used an access tool created by Kayed [26]; we provided the program with two groups of concepts, the first group consisted of the single concepts of each requirement definition(s), and the second group consisted of our suggested ontology domain concepts; it is the same tool used in the refinement process for the extracted ontology concepts. The results after that appeared with which concepts from our ontology domain covered concepts from each requirement definition(s).

Depending on the results provided by the software we used, and for each e-government applications requirement definition(s), we counted how many concepts our ontology domain concepts covered, and calculated the average coverage for each one.

Finally we calculated the average for all the resulting coverage averages for all of the requirements definition(s). The result of the coverage process showed that an average of 75% of the definitions concepts was covered by our ontology domain concepts. That was a very high percentage for the studied domain text corpus.

The evaluation process revealed that our ontology domain concepts covered almost 75% from the given knowledge domain. This result supports our claim; that we can condense the thousands of concepts used to define the most common and discussed e-government applications requirements into a smaller set of concepts (25 concepts).

The full results of our coverage technique are presented in Appendix B.

CHAPTER FOUR

EXTRACTING ONTOLOGY DOMAIN RELATIONSHIPS

4.1 OVERVIEW

In this chapter we extracted general relationships between the new concepts of the suggested ontology domain after studying and filtering the results of two tools. We presented the resulting relationships as groups. After that, a general lattice representation for part of the resulting relationships was done.

4.2 EXTRACTING RELATIONSHIPS BETWEEN CONCEPTS

Extracting structured information and relationships from text between concepts has been widely studied lately and became a rich subject of research. Many works and documents and even theses have been published about this subject exclusively. When the decision was to proceed in this step; extracting and creating relationships between concepts of our ontology domain, we found that if we want to create a detailed and complete ontology relation taxonomy, then this work will be large enough to be a thesis by itself.

Such details go beyond our work scope, and we suggested it to be done in the future. So, we went for extracting and presenting a basic and general representation of the relationships that we could extract between our ontology domain concepts.

In this step, and in order to extract relationships between our suggested ontology domain concepts, we used and studied the results of two tools. Firstly, we used the TextToOnto tool in order to extract relationships (associations) between concepts. We provided the tool with the text corpus we prepared previously to extract concepts from, and also we provided it with the concepts we want to study the relationships between them.

When we ran this step the used tool provided us with about 28854 relationships. Figure 4.1 shows part of the results of this step.

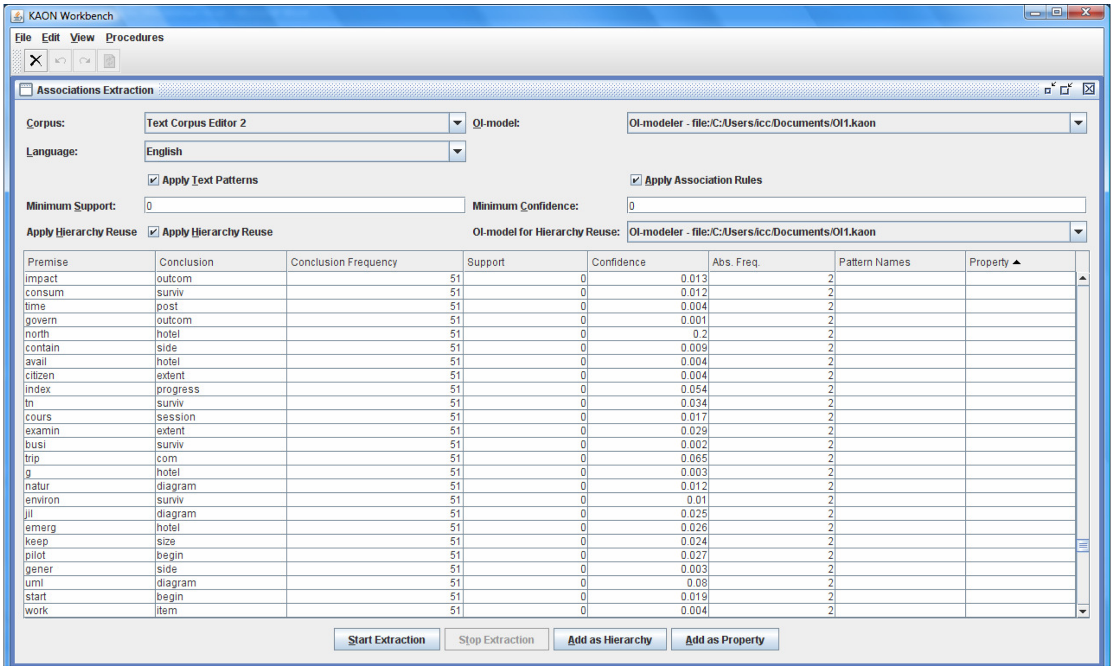
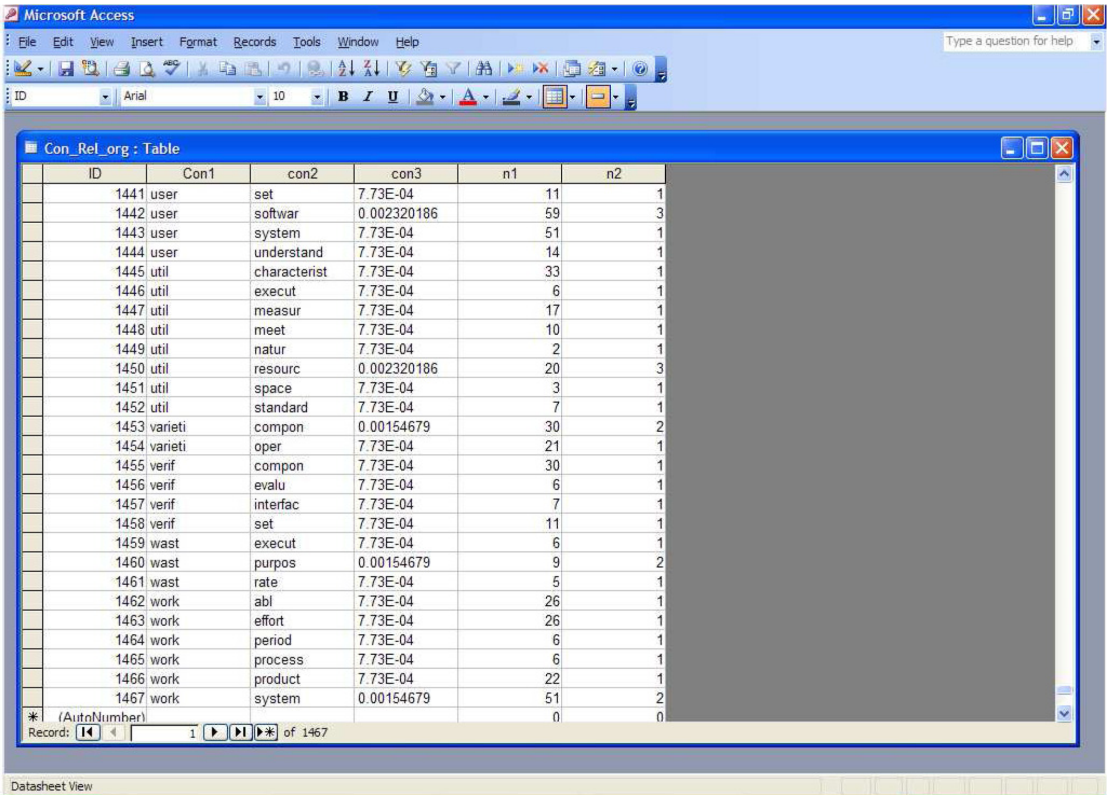


Figure 4.1: Part of the resulting relationships using TextToOnto tool.

To get benefits from the resulting relationships from the TextToOnto tool, we used them as an input for another tool; a tool created by Kayed et al [28]. Such a tool accepts the relationships resulting from the TextToOnto tool as an input, and implements a counting and relevancy algorithm on them. This tool provided us with about 656 relationships categorized in groups of concepts. Figures 4.2 and 4.3 shows part of using this tool and part of its results. Later, we took the 656 resulting groups from this tool, studied them, filtered them upon containing our ontology domain concepts or not (because the text corpus we provided to the tool contained much more concepts than our ontology domain did).

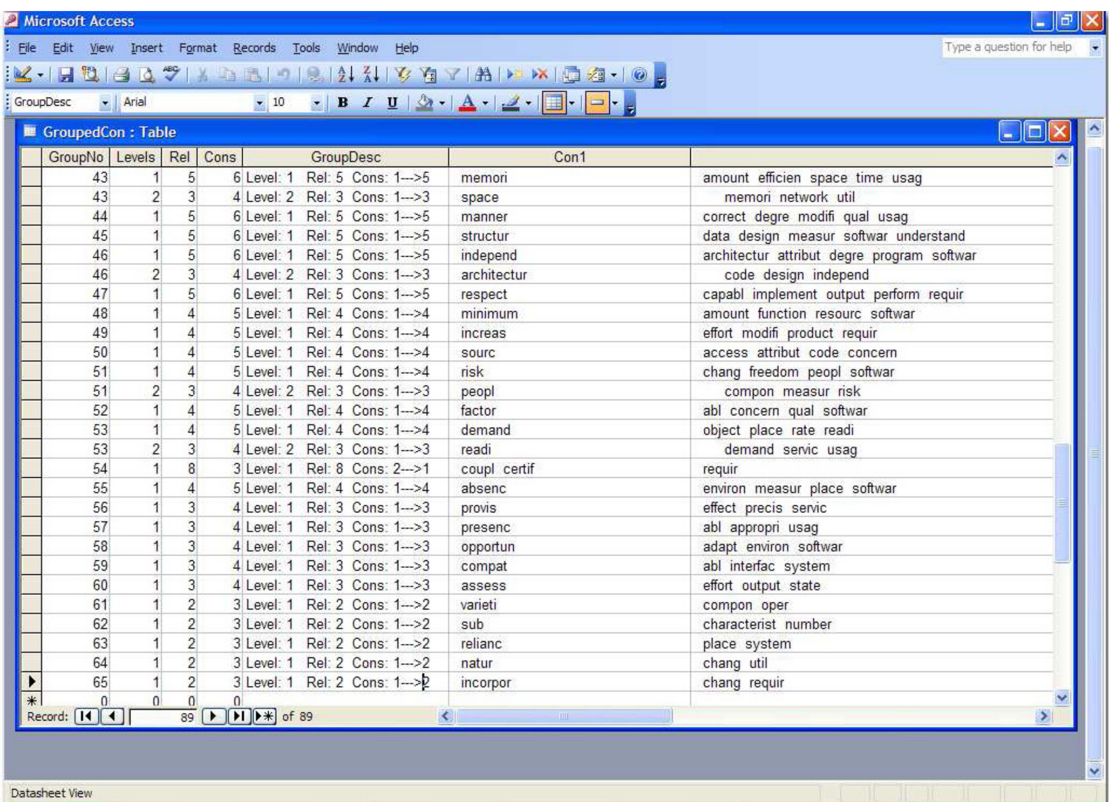


Con_Rel_org : Table

ID	Con1	con2	con3	n1	n2
1441	user	set	7.73E-04	11	1
1442	user	softwar	0.002320186	59	3
1443	user	system	7.73E-04	51	1
1444	user	understand	7.73E-04	14	1
1445	util	characterist	7.73E-04	33	1
1446	util	execut	7.73E-04	6	1
1447	util	measur	7.73E-04	17	1
1448	util	meet	7.73E-04	10	1
1449	util	natur	7.73E-04	2	1
1450	util	resourc	0.002320186	20	3
1451	util	space	7.73E-04	3	1
1452	util	standard	7.73E-04	7	1
1453	varieti	compon	0.00154679	30	2
1454	varieti	oper	7.73E-04	21	1
1455	verif	compon	7.73E-04	30	1
1456	verif	evalu	7.73E-04	6	1
1457	verif	interfac	7.73E-04	7	1
1458	verif	set	7.73E-04	11	1
1459	wast	execut	7.73E-04	6	1
1460	wast	purpos	0.00154679	9	2
1461	wast	rate	7.73E-04	5	1
1462	work	abl	7.73E-04	26	1
1463	work	effort	7.73E-04	26	1
1464	work	period	7.73E-04	6	1
1465	work	process	7.73E-04	6	1
1466	work	product	7.73E-04	22	1
1467	work	system	0.00154679	51	2
(AutoNumber)				0	0

Record: 1 of 1467

Figure 4.2: Using TextToOnto results as an input for the second MS Access tool.



GroupedCon : Table

GroupNo	Levels	Rel	Cons	GroupDesc	Con1	
43	1	5	6	Level: 1 Rel: 5 Cons: 1--->5	memori	amount efficien space time usag
43	2	3	4	Level: 2 Rel: 3 Cons: 1--->3	space	memori network util
44	1	5	6	Level: 1 Rel: 5 Cons: 1--->5	manner	correct degre modifi qual usag
45	1	5	6	Level: 1 Rel: 5 Cons: 1--->5	structur	data design measur softwar understand
46	1	5	6	Level: 1 Rel: 5 Cons: 1--->5	independ	architectur attribut degre program softwar
46	2	3	4	Level: 2 Rel: 3 Cons: 1--->3	architectur	code design independ
47	1	5	6	Level: 1 Rel: 5 Cons: 1--->5	respect	capabl implement output perform requir
48	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	minimum	amount function resourc softwar
49	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	increas	effort modifi product requir
50	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	sourc	access attribut code concern
51	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	risk	chang freedom peopl softwar
51	2	3	4	Level: 2 Rel: 3 Cons: 1--->3	peopl	compon measur risk
52	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	factor	abl concern qual softwar
53	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	demand	object place rate readi
53	2	3	4	Level: 2 Rel: 3 Cons: 1--->3	readi	demand servic usag
54	1	8	3	Level: 1 Rel: 8 Cons: 2--->1	coupl certif	requir
55	1	4	5	Level: 1 Rel: 4 Cons: 1--->4	absenc	environ measur place softwar
56	1	3	4	Level: 1 Rel: 3 Cons: 1--->3	provis	effect precis servic
57	1	3	4	Level: 1 Rel: 3 Cons: 1--->3	presenc	abl appropri usag
58	1	3	4	Level: 1 Rel: 3 Cons: 1--->3	opportun	adapt environ softwar
59	1	3	4	Level: 1 Rel: 3 Cons: 1--->3	compat	abl interfac system
60	1	3	4	Level: 1 Rel: 3 Cons: 1--->3	assess	effort output state
61	1	2	3	Level: 1 Rel: 2 Cons: 1--->2	varieti	compon oper
62	1	2	3	Level: 1 Rel: 2 Cons: 1--->2	sub	characterist number
63	1	2	3	Level: 1 Rel: 2 Cons: 1--->2	relianc	place system
64	1	2	3	Level: 1 Rel: 2 Cons: 1--->2	natur	chang util
65	1	2	3	Level: 1 Rel: 2 Cons: 1--->2	incorpor	chang requir
0	0	0	0			

Record: 89 of 89

Figure 4.3: Part of the resulted relationships groups from using the second tool.

The resulting relationships between groups of our Ontology concepts after filtering are mentioned in Appendix C.

If we took a look at the table of Appendix C, which shows the relationships between some groups of our ontology concepts, we would see a column of the group number; which refers to the ID of the group of concepts that had a relationship between them. The second column shows the level number which refers to the number of levels of the relationships when concepts from the same group have relationships with other concepts. Before filtering, every group was consisted of two levels: level one indicated that there is a relationship between a group of concepts; call it g1, and another group of concepts, say g2, while in level 2, the reverse of the relationship is given, that is the relationship that g2 has with g1. So we filtered and eliminated them from the table above and said that g1 had a relationship with g2 and vice versa instead. Also the second and the third level from a group may show that a concept or (concepts), which considered as a part of a group of concepts, had a relationship with other concepts from the domain. We did not eliminate this type of relationships and we used it later in our representation.

Also, we would see the group 1 column; which refers to a side of the group of concepts that had a relationship with another group of concepts shown in group 2 column.

4.3 A LATTICE REPRESENTATION OF THE RELATIONSHIPS

After we studied and filtered the resulting relationships from the tools we used, as shown in Table 4.1 above, we considered representing them in a general form of lattice representation, but if we did it to all relationships groups, it will be a large and complex representation in addition to time consuming. So we took part of those groups and represented them as shown in the Figures below:

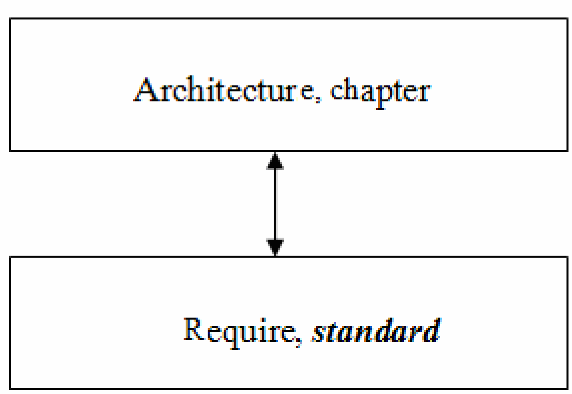


Figure 4.4: Group 4 relationship Lattice representation: One Level Relationship.

Figure 4.4 above shows one level relationship, which indicates that a group of concepts which consists of (Architecture, chapter) has a relationship with another group of concepts which consists of (Require, ***standard***) and vice versa.
When we looked at the semantic used to define the studied requirements, we found that concepts from group one came in the semantic along with concepts from the second group.
This is another example of one level relationship that results from the extracted relationships, which is shown in Figure 4.5.

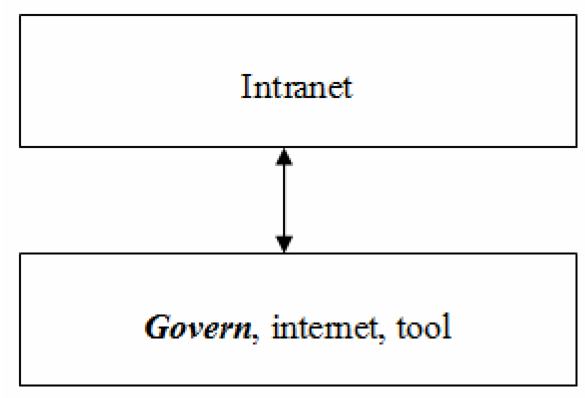


Figure 4.5: Group 5 relationship Lattice representation: One Level Relationship.

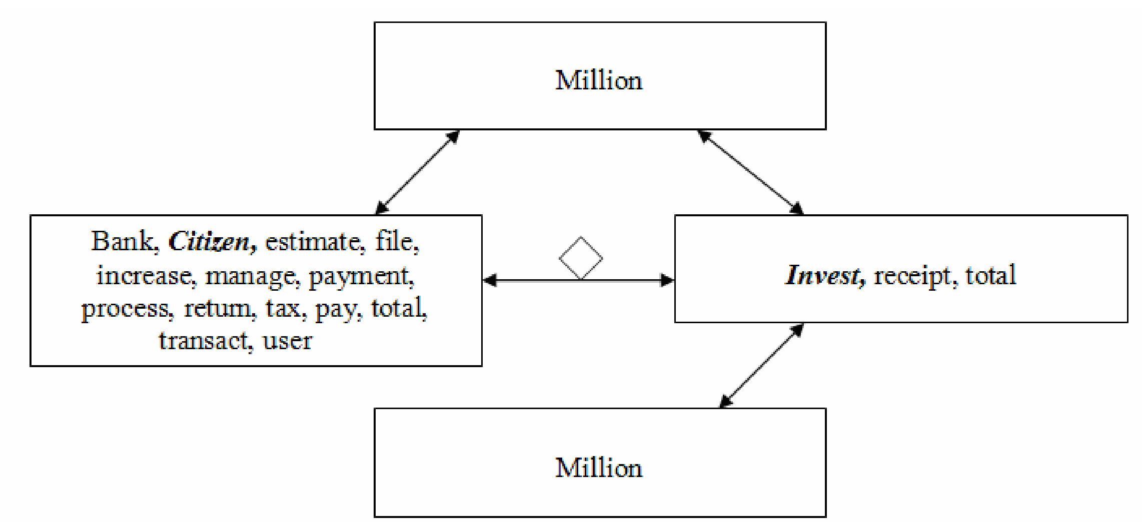


Figure 4.6: Group 6 relationship Lattice representation: Two Levels Relationship.

Figure 4.6 above shows a two levels relationship. We separated the second group concepts from each other because we needed to connect a part from it (***Invest***, receipt, total) with a second level group of concepts consists of (Million). These groups are used together in the semantic when defining some of the studied requirements. The diamond shape on the arrow between some concepts means that those concepts together considered as a group, but separated for a good reason.

This is another example of two level relationship, which is shown in figure 4.6

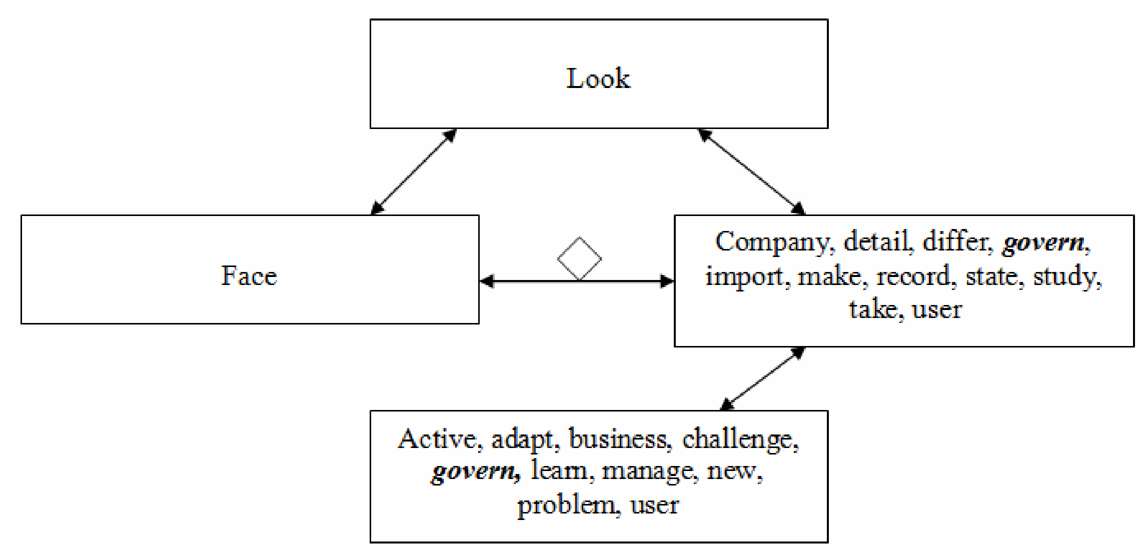


Figure 4.7: Group 9 relationship Lattice representation: Two Level Relationship.

The following figures 4.8 and 4.9 shows a three levels relationship.

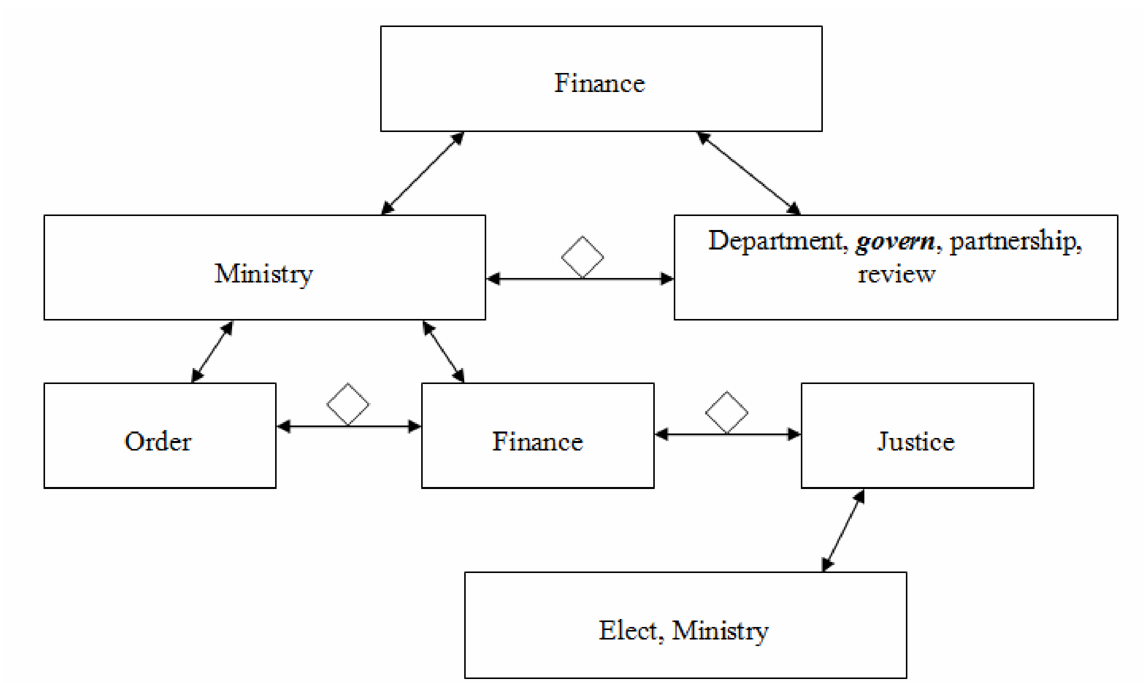


Figure 4.8: Group 14 relationship Lattice representation: Three Level Relationship.

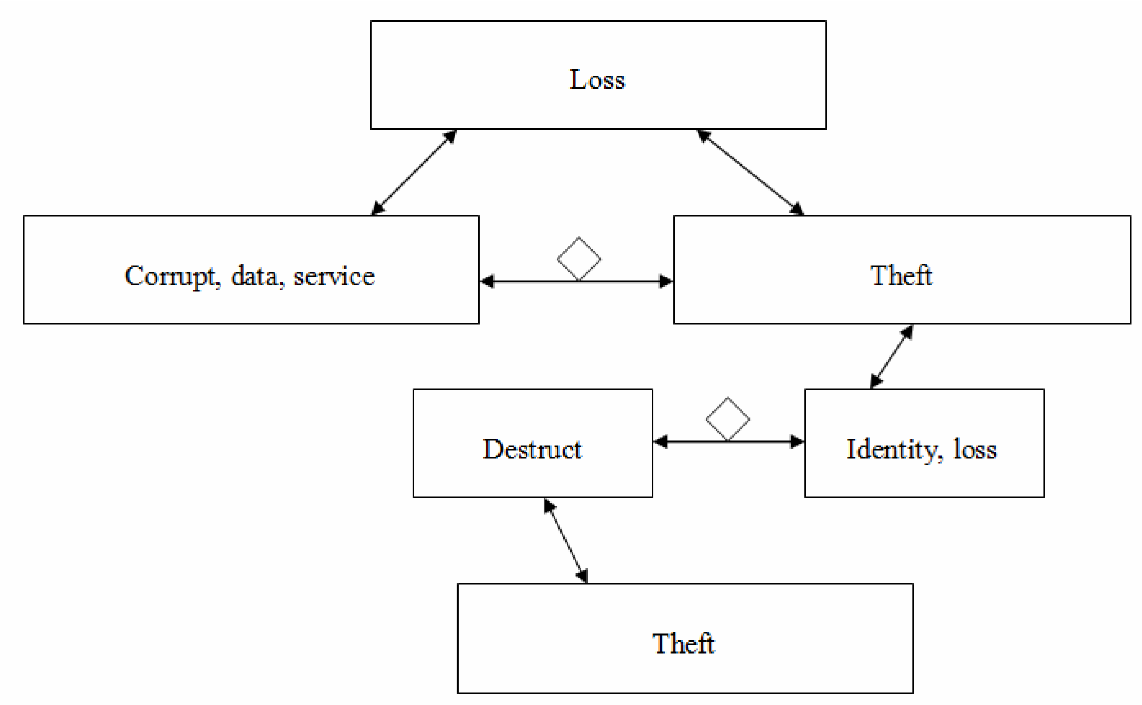


Figure 4.9: Group 15 relationship Lattice representation: Three Level Relationship.

So, for all the groups that appears in the table, we can make a lattice representation as shown, either for one or two or three levels relationship.

CHAPTER FIVE

CONCLUSIONS AND FUTURE WORK

5.1 OVERVIEW

In this chapter we present and discuss the conclusions of our research; our final results and how we used them to contribute to the studied domain are presented among the conclusions. Future work is suggested at the end of this chapter.

5.2 CONCLUSIONS

REPEGAs discipline is considered in the emerging phase, and it suffers from the typical symptoms of any relatively evolving disciplines. REPEGAs are currently in the phase in which terminologies, principles, and methods are still being defined, consolidated, and agreed. In particular, there is a lack of consensus on the concepts and terminologies used in the semantic of this field. Studies showed that inconsistencies in the semantic used different research attributes proposals often occur.

In our research we focused on studying REPEGAs concepts and terminologies that current e-gov. applications proposals, documents, and reports present. We prepared text corpora from them to be used in a tool to extract the most discussed and used concepts from it. After that experts were asked to study and filter the resulting concepts and provided them to us.

An evaluation phase depended on a coverage technique was done to the resulting concepts, followed by an enhancing step to the evaluated ontology domain concepts which led us to increase the number of the suggested concepts in the ontology domain, after that a coverage evaluation was done again to the new suggested ontology domain concepts.

In order to extract general relationships among the suggested ontology domain concepts, we returned to the prepared text corpus again and ran out two tools on it. We studied them, filtered them, listed them and represented part of them using a lattice representation.

Through completing the steps of our work, which have been previously summarized, we reached to many important results. These results are studied filtered and used to support our claim. The sections below present the final results and how we used them to support our contribution to the field.

5.3 PRESENTING FINAL RESULTS

Through completing the steps of our work we found out many important results, which could be summarized as follows:

• Ontology domain concepts: resulted from preparation, studying, and filtering a text corpus related to the domain of requirements engineering process in e-government applications. First, we reached to a result that we can condense the semantic of thousands of concepts used to define the discussed 25 concepts into a smaller set of concepts consisted of 25 concepts with a coverage percentage for the studied knowledge domain of 75%.

• Relationships between groups of concepts for the suggested ontology domain: resulted from studying and filtering the results of two tools; the associations resulted from using the TextToOnto tool after we provided it with a related knowledge domain text corpus and concepts, After that we took those associations and provided another tool created by Kayed et al [28] with them. This process provided us with relationships between groups of concepts from our suggested ontology domain. Again we studied them, filtered them, and finally presented them.

• Each requirement of e-gov applications concepts that belong to our ontology domain: from the evaluation phase for the ontology domain concepts. We reached to every e-gov application definition concept that belongs to our ontology domain concepts.

• Finally, Relationships between each concept in the ontology domain and other concepts also in the same domain: resulted from studying the groups of relationships in addition to the domain itself.

5.4 OUR CONTIBUTION

By reaching and providing those final results discussed in the previous section, let us not forget that our main focus in this work is to provide experts mainly, researchers, and practitioners in the field of SWE with an ontology to be considered as a base and a common agreement knowledge. This supports them in defining the common discussed REPEGAs (25 concept) that we extracted from the fields documents and reports, and reaching to a common, shared, and consistent semantic for them. This solves the inconsistencies of the semantic evident in the definitions of those attributes among many documents and reports.

We have presented the conceptualization of the common REPEGAs by an ontology, which serves in this specific domain.
We also have condensed the semantic of thousands of concepts used to define any of the discussed REPEGAs into a smaller set of concepts consisting of 25 concepts with a high percentage of coverage average for the studied domain reached to 75% of coverage.

We also provide the experts and practitioners in the field of SWE for EGAs who want to define any of the discussed 25 words in the domain with an ontology which contains a set of common used and agreed concepts (for each attribute definition, and for general studied domain), and also with relationships between them (as groups, or relationships between concepts which belong to the same requirements that can be inferred from the presented relationships, or relationships between concepts in the studied domain). So when an expert decides to define a requirement from the discussed domain, we suggest two ways to use our ontology to have a consistent semantic with other definitions in the field. First after an expert writes down his own definition, he can compare the concepts he used in the semantic of his definition with our ontology domain concepts and try to map from his used concepts to our concepts from the ontology domain if needed, and try to use the provided relationships between them to connect the semantic of the concepts together in a strong, meaningful, and consistent manner. The second way that we suggest to reach to an agreed semantic is that before the expert write down his own definition we recommend to take a look on the ontology domain and use its concepts and relationships along with his experience as a base knowledge to consist the definition he wants.

If experts in the field follow one of these suggested ways when defining one of the discussed REPEGAs, eventually they will reach a common, agreed, and consistent semantic between them, and this will be a successful way to solve the presented problem.

In addition to this, our ontology provides a base to evaluate any related presented definition semantic for one of the 25 studied words. The way of doing this is if a high percentage of the concepts used in the semantic of the presented definition is covered by our ontology domain, the presented definition semantic can be accepted, but if not we claim that it is a weak semantic to be used defining such an concept.

A table mentioned in Appendix D, shows a comparison between my thesis work and related work, that check the common and different points with the previous studies in the same domain of research, also the results shows my contribution more clearly.

5.5 FUTURE WORK

Through conducting this research, many ideas and issues were unfolded but not accomplished yet because of time, resources, and other constraints. We would like to suggest a few ideas for future study:

- Providing a description for each concept used in the provided ontology domain in order to help experts and practitioners who want to use them while defining one of the discussed e-government applications.

- A formal representation for the ontology:
  Some suggestions for the tools we used to be more user friendly, as the possibility of copying records as all in all not a record in a time, and putting some notes in the interface that help the users to use the tool easily.

- Extracting and presenting the detailed types of relationships between our ontology domain concepts.

- Completing the lattice representation for the ontology domain relationships.

- Using another approach to evaluate and enhance the ontology domain, and comparing the results with the results we already had. A critique on evaluating the ontology may be done by conducting set of experiments and trying to deploy the ontology in some applications to show how effective, useful, and expressive is the proposed ontology to the audience in a context of software engineering domain and especially to the audience in the context of e-government applications domain.

- Using our ontology domain and convert it into an Arabic ontology for the same studied domain but in the Arabic language.

- Providing full coverage for all ontology domains, to let the results be more accurate and reliable.

REFERENCES

[1] Baker G.; Brass A.; Bechhofer S.; Goble C.; Paton N.; Tambis R., “Transparent access to multiple biological information sources”, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology , ISMB-98, Menlow Park, California, AAAI Press, 25-34,June 28-July 1 1998.

[2] Boehm B.; Brown J., Kaspar H.; Lipow M.; McLeod G.; Merritt M., “Characteristics of Software Quality”, Elsevier, North Holland, 1st Edition, ISBN: 0444851054, 1978.

[3] Brank J.; Grobelnik M.; Mladenić D., “A Survey of Ontology Evaluation Techniques”, in proc. Of the 8th international multi conference information society, SIKDD, 2005.

[4] Brazier, F.M.T., Jonker, C.M., and Treur, J. (1998), Principles of Compositional Multi-agent System Development. In: J. Cuena (ed.), Proceedings of the 15th IFIP World Computer Congress, WCC'98, Conference on Information Technology and Knowledge Systems, IT&KNOWS'98, pp. 347-360.

[5] Brewster C.; Alani H.; Dasmahapatra S.; Wilks Y., “Data-driven ontology evaluation”. InProc. of the 4th International Conference on Language Resources and Evaluation, Lisbon,2004.

[6] Brooks F., “No Silver Bullet - essence and accidents of software engineering”, Compute*r* magazine, Vol. 20, Issue 4, PP: 10-19, Apr. 1987.

[7] Calero C.; Ruiz F.; Piattini M., “Ontologies for Software Engineering and Software Technology”, Springer Berlin Heidelberg, New York, ISBN-10 3-540-34517-5, 1998.

[8] Conde D., “Software Product Management: Managing Software Development from Idea to Product to Marketing to Sales(Execenablers)”, Aspatore Books, 1st Edition, ISBN 1-58762-202-5, Sep. 2002.

[9] Corcho O.; Fernandez M.; et al., "Building Legal Ontologies with METHONTOLOGY and WebODE", In Benjamins, R.; Casanovas, P.; Breuker, J. & Gangemi, A. (ed.): "Law and the Semantic Web".

[10] De los Angeles Martin M.; Olsina L., “Toward an Ontology for SW metrics and indicators as the Foundation For Catalog Web system”. Web Congress, First Latin American, Santiago, Chile, Vol. 30 , Issue 3, 10-12,PP: 103 -113, ,2003.

[11] Dellschaft K.; Staab S., “On How to Perform a Gold Standard Based Evaluation of Ontology Learning”. International Semantic Web Conference, PP: 228-241, 2006.

[12] Ernst M., “Ontology building: A survey of editing tools”, Online, accessed 2008 April,Available from URL: http://www.xml.com/pub/a/2002/11/06/ontologies.html

[13] Floridi L., “Blackwell Guide to the Philosophy of Computing and Information”, Preprint version of chapter “Ontology”, Oxford: Blackwell, PP: 155– 166, 2003.

[14] FZI Karlsruhe; AIFB Karlsruhe, “KAON: the Karlsruhe ontology and semantic web framework developer’s guide for KAON 1.2.7”, University of Karlsruhe, Germany, 2004.

[15] Garcia F.; Bertoa M.; Calero C., “Towards a Consistent Terminology for Software measurements”, Information of Software Technology, Vol. 48, Issue 8, PP: 631-644, 27 June 2005.

[16] Girardi R., De faria C, "An ontology-based technique for the specification of domain and user models in multi-agent domain engineering, ACM.

[17] Greenspan S., Mylopoulos J., Bordiga A, "On formal requirements modeling languages: RML revisited", Sixth international conference on software engineering (ICSE-6), ACM.

[18] Guarino N.; “Formal ontology and information systems”, Proceedings of the international conference on Formal Ontology in Information Systems, Trento, Italy, Amsterdam: IOS Press,
Netherlands, PP: 3-15,1998.

[19] Guarino N.; Persidis A., "Evaluation framework for content standards", Technical Report, OntoWeb Deliverable 3.5, Padova; 2003.

[20] Hamza H., "Developing business object models with patterns and Ontologies", Companion to the 20th annual ACM SIGPLAN conference on programming, systems, languages, and applications OOPSLA ,2005.

[21] IEEE, "IEEE Standard Glossary of Software Engineering Terminology/IEEE Std 610.12-1990", (Revision and redesignation of IEEE Std 729-1983), Sponsor Standards Coordinating Committee of the IEEE Computer Society, Approved September 28, 1990 IEEE Standards Board, ISBN**:** 978-1559370677, 1990.

[22] ISO/IEC, "Software and Systems Engineering - Guidelines for the Application of ISO/IEC9001:2000 to Computer Software". International Standards Organization, Geneva, witzerland,2004.

[23] Jarrar M., "Towards methodological principles for ontology engineering". PhD Dissertation, Vrije Universities Brussels, 2005.

[24] Jarrar M.; Demy J.; Meersman R., "On reusing conceptual data modeling for ontology engineering". In: Aberer K, March S, Spaccapietra S (eds.), Journal on Data, Vol. 2800, Issue1,PP: 185-207, 2003.

[25] Kayed A., "Building e-laws ontology: New approach", Springer, Lecture Notes in Computer Science ,Springer-Verlag Vol. 3762 , Germany, ISBN/ISSN: 0302-9743, pp 826-830,2006.

[26] Kayed A.; Colomb R., "Extracting Ontological Concepts for Tendering Conceptual Structures", Data & Knowledge Engineering, Vol. 40, Issue 1, Pages: 71 - 89, 2002.

[27] Kayed A.; Colomb R.., "Using BWW Model to Evaluate Building Ontologies in CGs Formalism". Information Systems. Vol. 30, Issue 5, PP: 379 – 398, ISSN: 0306-4379, July2005.

[28] Kayed A.; Hirzallah N.; Al Shalabi L. A.; and Najjar M., "Building Ontological Relationships: A new approach", Journal of the American Society for Information Science and Technology, John Wiley & Sons Inc, 111 River ST , Hoboken , USA , NJ, 07030, 2008.

[29] Knowledge Management Group (WBS); Research Group Knowledge Management (WIM), "Extensions to the Karlsruhe Ontology and Semantic Web Framework, KAON Extensions, Developer's Guide for KAON Extensions 0.6", University of Karlsruhe, Germany, Aug. 2003.

[30] Lamsweerde A., "Requirements engineering in the year 00: A research perspective", ACM, 2000.

[31] Liao L., Yuzhong Qu., Hareton, Leung K.N. "A software process ontology and its application", 2003.

[32] Maedche A.; Volz R.,"The ontology extraction and maintenance framework text-to-onto".In Proceedings of the ICDM'01 Workshop on Integrating Data Mining and Knowledge Management, 2001.

[33] Marko G.; Mladeni D., "Automated Knowledge Discovery in Advanced Knowledge Management", Journal of Knowledge Management, Vol. 9, Issue 5, PP: 132–149, 2005.

[34] Marnik M., Heering J., Sloane A.," When and how to develop domain Specific languages", ACM computing surveys (CSUR), volume 37, Issue 4, 2005.

[35] Mika P., "Social networks and the semantic web: the next challenge", IEEE Intelligent Systems, Vol. 20, Issue 1, PP: 80-93, 2005.

[36] Noy N.; McGuinness D., "Ontology development 101: A guide to creating your first ontology", Technical Report KSL-01-05, Stanford Knowledge Systems Laboratory, and Technical Report SMI-2001-0880, Stanford Medical Informatics, California, USA, 2001.

[37] Nuseibeh B., Easterbrook S, "Requirements engineering: A Roadmap" ACM.

[38] Ontology, computer science, Online, accessed on 2008 April, Available from URL http://en.wikipedia.org/wiki/Ontology (computer_science)#_note-0.

[39] Ontoprise GmbH, "How to work with OntoEdit — user's guide for OntoEdit version 2.6"Online, accessed 2008 April, available from URL http://www.ontoprise.de/documents/tutorialontoedit.pdf.

[40] P´erez G., "Some ideas and examples to evaluate ontologies". Technical Report KSL-94-65, Knowledge Systems Laboratory, Stanford, 1994.

[41] P´erez G., "Towards a framework to verify knowledge sharing technology", Expert Systems with applications, Vol. 11, Issue 4, PP: 519–529, 1996.

[42] Parekh V.; Gwo J.; Finin T.; "Mining Domain Specific Texts and Glossaries to Evaluate and Enrich Domain Ontologies", International Conference on Information and Knowledge Engineering, PP: 533-540, USA, CSREA, ISBN 1-932415-27-0, 2004.

[43] Pressman R., "Software Engineering, a practitioner's approach", sixth Edition, Mc Graw Hill Companies, ISBN: 978-0073019338, April 2005.

[44] Randell B.; Naur P., "Software Engineering", Report on a conference sponsored by the NATO Science Committee, Garmisch, Germany, 7th to 11th Oct.1968.

[45] Robert P.; Malaka R., "A task-based approach for ontology evaluation". In Proc. ECAI 2004 Workshop on ontology Learning and Population, PP: 9–16, 2004.

[46] Samhan A., "Master Thesis: Towards An Ontology for Software Product Quality Attributes", Middle East University for Graduate Studies, Amman Jordan, May, 2008.

[47] Scacchi W.; Jensen C.; Noll J.; Elliott M.," Multi-Modal Modeling, Analysis and Validation of Open Source Software Requirements Processes", Intern. J. Information Technology and Web Engineering, Vol. 1, Issue 3, PP: 49-63, 2006.

[48] SEI, "Software Engineering 2004", Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering, A Volume of the Computing Curricula Series, Aug.2004.

[49] Smith B., “Ontology”, Draft version of chapter published in Luciano Floridi (ed.),Blackwell Guide to the Philosophy of Computing and Information, Oxford: Blackwell, PP:155-166, 2003.

[50] Smith B.; Welty C., “Ontology: Towards a New Synthesis”, Ogunquit, Maine, USA.,ACM1-58113-377-4/01/0010, FOIS’01, Oct. 17-19, 2001.

[51] Software Engineering Institute, CMMI® Web Site, Carnegie Mellon, online, accessed on 2008 April, available from URL http://www.sei.cmu.edu/cmmi/cmmi.html, 2004.

[52] Sommerville I., “Software Engineering", Pearson Addison Wesely, Seventh Edition, ISBN: 020139815, 2007.

[53] V¨olker J.; Vrande D.; Sure Y., “Automatic evaluation of ontologies (AEON)”.In Proc. Of the 4th International Semantic Web Conference (ISWC’05), Springer Verlag Berlin-Heidelberg, Vol. 3729, PP: 716-731, Nov. 2005.

[54] Wordreference.com: WordNet® 2.0, Princeton University, Princeton, NJ. , online, Accessed on 2008 April, available from URL www.wordreference.com.

[55] Zieliński K.; Szmuc T., “Software Engineering: Evolution and Emerging Technologies”,IOS Press, ISBN 1-58603-559-2, 2005.

[56] Zhao G., Gao Y., Meersman R, "An ontology-based approach to business modeling" ACM.

[57] Zhiyuan F., "E-Government in Digital Era: Concept, Practice, and Development", International Journal of The Computer, The Internet and Management, Vol. 10, No.2, 2002, p 1-22.

APPENDICES

APPENDIX A

The complete common extracted REPEGAs from various websites and dictionaries related to the field of study, and their different definitions found in them.

Att ID	Quality Attribute	Definition(s)	Source(s) Reference(s)
1	Accuracy	Conformity to fact.	www.answers.com
		Conformity to truth or to a standard or model	www.merriam-webster.com
		The degree of closeness of a measured or calculated quantity to its actual (true) value	www.wikipedia.org
2	Policy	A plan or course of action, as of a government, political party, or business, intended to influence and determine decisions, actions, and other matters.	www.answers.com
		A high-level overall plan embracing the general goals and acceptable procedures especially of a governmental body.	www.merriam-webster.com
		Is typically described as a deliberate plan of action to guide decisions and achieve rational outcome(s). However, the term may also be used to denote what is actually done, even though it is unplanned	www.wikipedia.org
3	Database	A collection of data arranged for ease and speed of search and retrieval. Also called *data bank*.	www.answers.com
		Usually large collection of data organized especially for rapid search and retrieval (as by a computer).	www.merriam-webster.com
		is a structured collection of records or data that is stored in a computer system.	www.wikipedia.org
4	Awareness	Having or showing realization, perception, or knowledge.	www.merriam-webster.com
		Is a term referring to the ability to perceive, to feel, or to be conscious of events, objects or patterns, which does not necessarily imply understanding.	www.wikipedia.org

5	Strategy	Management plan or method for completing objectives; plan of procedures to be implemented, to do something	www.answers.com
		Is a plan of action designed to achieve a particular goal	www.wikipedia.org
6	Adoption	Deals with the transfer (conversion) between an old system to a target system in an organization. So if a company works with an old software system, it may want to use a new system which is more efficient, has more work capacity etc. So then a new system needs to be adopted, where after it can be used.	www.answers.com
7	Client	Person, company, or organization who uses the professional services of another.	www.answers.com
		Someone that purchases something from someone else or hires a consultant or service.	www.wikipedia.org
8	e-signature	Digital signing, timestamping and asymmetric encryption components.	www.answers.com
		An electronic signature is an "a signature that consists of one or more letters, characters, numbers or other symbols in digital form incorporated in, attached to or associated with an electronic document".	www.wikipedia.org
9	Citizen	A resident of a city or town, especially one entitled to vote and enjoy other privileges there.	www.answers.com
		Mean a person owing allegiance to and entitled to the protection of a sovereign state.	www.merriam-webster.com
		Refers to a person's membership in a political community such as a country or city.	www.wikipedia.org
10	Sector	A part or division, as of a city or a national economy: the manufacturing sector; the private sector.	www.answers.com
		A geometric figure bounded by two radii and the included arc of a circle.	www.merriam-webster.com
11	Security	The existence and enforcement of techniques which restrict access to data, and the conditions under which data may be obtained.	www.answers.com

		Is the degree of protection against danger, loss, and criminals. Individuals or actions that encroach upon the condition of protection are responsible for a "breach of security".	www.wikipedia.org
12	Elicitation	Working with customers to find out about the application domain, the services that the system should provide and the system's operational constraints.	www.answers.com
13	Standard	An acknowledged measure of comparison for quantitative or qualitative value; a criterion. An object that under specified conditions defines, represents, or records the magnitude of a unit.	www.answers.com
		An organization flag carried by a mounted or motorized military unit.	www.merriam-webster.com
		Is an established norm or requirement. It is usually a formal document that establishes uniform engineering or technical criteria, methods, processes and practices.	www.wikipedia.org
14	Reliability	The probability that a component part, equipment, or system will satisfactorily perform its intended function under given circumstances, such as environmental conditions, limitations as to operating time, and frequency and thoroughness of maintenance for a specified period of time.	www.answers.com
		The extent to which an experiment, test, or measuring procedure yields the same results on repeated trials.	www.merriam-webster.com
		Is the ability of a person or system to perform and maintain its functions in routine circumstances, as well as hostile or unexpected circumstances.	www.wikipedia.org
15	Classification	The act, process, or result of classifying. A category or class.	www.answers.com
		Systematic arrangement in groups or categories according to established criteria.	www.merriam-webster.com
16	Investment	Property or another possession acquired for future financial return or benefit.	www.answers.com
		Is a term with several closely-related meanings in business management,	www.wikipedia.org

		finance and economics, related to saving or deferring consumption. Investing is the active redirecting resources from being consumed today so that they may create benefits in the future; the use of assets to earn income or profit.	
17	Govern	To make and administer the public policy and affairs of; exercise sovereign authority in.	www.answers.com
		To control, direct, or strongly influence the actions and conduct of.	www.merriam-webster.com
		Is defined as the body within an organization that has authority and function to make and the power to enforce laws, regulations, or rules.	www.wikipedia.org
18	Scope	To examine or investigate. Often used with out.	www.answers.com
		The sum of all projects products and their features.	www.wikipedia.org
19	Confidentiality	Private or secret; something treated with trust, resulting in a feeling of security that information will not be disclosed to other parties. An example is the confidentiality of conversations and records between attorney and client.	www.answers.com
20	Purpose	An aim or a goal.	www.answers.com
21	Stakehold	One who has a share or an interest in an enterprise.	www.answers.com
		One who is involved in or affected by a course of action.	www.merriam-webster.com
		A person, group, organization, or system who affects or can be affected by an organization's actions.	www.wikipedia.org
22	Trust	To have or place confidence in; depend on.	www.answers.com
		Assured reliance on the character, ability, strength, or truth of someone or something.	www.merriam-webster.com
23	Evaluation	To determine the significance, worth, or condition of usually by careful appraisal and study.	www.merriam-webster.com
		Is systematic determination of merit, worth, and significance of something or someone using criteria against a set of standards.	www.wikipedia.org
24	Knowledge	The state or fact of knowing.	www.answers.com

		The fact or condition of knowing something with familiarity gained through experience or association.	www.merriam-webster.com
		Expertise, and skills acquired by a person through experience or education; the theoretical or practical understanding of a subject.	www.wikipedia.org
25	Identification	The act of identifying.	www.answers.com

APPENDIX B

The following table shows the complete results from the evaluation step of our ontology; that demonstrate each REPEGA extracted concepts and the covered concepts from our ontology in addition to the coverage percentage.

Table B.1: The complete results from the ontology evaluation step.

Att. ID	Attribute	Def. Concepts	Onto. Concepts that cover	Count and Average
1	Accuracy	Accuracy	Accuracy	8 from 11
		Actual	Actual	0.73
		Calculated	Conformity	
		Closeness	Degree	
		Conformity	Fact	
		Degree	Quantity	
		Fact	Truth	
		Model	Value	
		Quantity		
		Truth		
		Value		
2	Policy	Action	Action	8 from 10
		Business	Business	0.80
		Government	Government	
		Guide	Guide	
		Influence	Outcome	
		Outcome	Plan	
		Overall	Policy	
		Plan	Term	
		Policy		
		Term		
3	Database	Collection	Collection	5 from 9
		Computer	Computer	0.56
		Data	Data	
		Database	Database	
		Large collection	Retrieval	
		Rapid search		
		Retrieval		
		Search		
		Speed		
4	Awareness	Ability	Ability	5 from 6
		Awareness	Awareness	0.83
		Knowledge	Knowledge	
		Perception	Realization	

		Realization	Term	
		Term		
5	Strategy	Adoption	Adoption	10 from 12
		Capacity	Capacity	0.83
		Goal	Management plan	
		Management plan	Method	
		Method	Organization	
		Organization	Plan	
		Plan	Software	
		Software	Strategy	
		Strategy	Target	
		System	Transfer	
		Target		
		Transfer		
6	Adoption	Adoption	Adoption	4 from 7
		Capacity	Capacity	0.71
		Organization	Organization	
		Software	System	
		System		
		Target		
		Use		
7	Client	Client	Client	3 from 6
		Company	Organization	0.50
		Consultant	Service	
		Organization		
		Person		
		Service		
8	E-Signature	Digital signing	Digital signing	4 from 5
		Document	E-Signature	0.80
		E-Signature	Electronic Document	
		Electronic Document	Encryption	
		Encryption		
9	Citizen	Citizen	Citizen	4 from 6
		Community	Membership	0.67
		Membership	Resident	
		Protection	Town	
		Resident		
		Town		
10	Sector	City	Division	4 from 5
		Division	Economy	0.80
		Economy	Manufacturing	
		Manufacturing	Sector	
		Sector		

11	Security	Access	Access	5 from 6
		Breach	Breach	0.83
		Data	Data	
		Degree	Protection	
		Protection	Security	
		Security		
12	Elicitation	Application	Application	4 from 5
		Domain	Domain	0.80
		Find	System	
		System	Elicitation	
		Elicitation		
13	Standard	Comparison	Comparison	8 from 10
		Criterion	Criterion	0.80
		Document	Document	
		Engineering	Measure	
		Measure	Organization	
		Organization	Requirement	
		Requirement	Uniform	
		Uniform	Value	
		Unit		
		Value		
14	Reliability	Ability	Ability	6 from 9
		Component	Experiment	0.67
		Experiment	Maintenance	
		Function	Probability	
		Maintenance	Reliability	
		Probability	System	
		Reliability		
		System		
		Test		
15	Classification	Category	Category	5 from 5
		Class	Class	1.00
		Classification	Classification	
		Process	Process	
		Result	Result	
16	Investment	Benefit	Benefit	4 from 6
		Business	Business	0.67
		Economics	Investment	
		Investment	Management	
		Management		
		Profit		
17	Govern	Authority	Authority	4 from 5
		Control	Control	0.80
		Function	Organization	
		Organization	Policy	

		Policy		
18	Scope	Feature	Product	3 from 4
		Product	Project	0.75
		Project	Scope	
		Scope		
19	Confidentiality	Client	Confidentiality	4 from 5
		Confidentiality	Information	0.80
		Information	Security	
		Security	Trust	
		Trust		
20	Purpose	Aim	Aim	3 from 3
		Goal	Goal	1.00
		Purpose	Purpose	
21	Stakehold	Enterprise	Enterprise	4 from 6
		Interest	Organization	0.67
		Organization	Person	
		Person	Share	
		Share		
		System		
22	Trust	Ability	Ability	4 from 5
		Confidence	Confidence	0.80
		Reliance	Strength	
		Strength	Trust	
		Trust		
23	Evaluation	Appraisal	Appraisal	3 from 5
		Condition	Evaluation	0.60
		Evaluation	Study	
		Significance		
		Study		
24	Knowledge	Education	Education	4 from 6
		Experience	Experience	0.67
		Fact	Knowledge	
		Knowledge	Person	
		Person		
		Subject		
25	Identification	Act	Act	3 from 4
		Identification	Identification	0.75
		Process	Process	
		Specific		
The Average of Coverage Averages is : 0.748 ≈ 0.75				

APPENDIX C

Relationships between groups of concepts in the suggested ontology domain:

Table C.1: Relationships between groups of concepts in the ontology domain.

Group No	Level	Group 1	Group 2
Examples of one level relationship			
1	1	System inform service process ***govern*** data user require application payment	Access base case function ***identify*** make model need provide relate support time
2	1	Question ***database*** associate	System
3	1	Impact detail	Application be ***govern*** inform
4	1	Architecture chapter	Require ***standard***
5	1	Intranet	***Govern*** internet tool
Examples of two level relationship			
6	1	Million	Bank ***citizen*** estimate file increase ***invest*** manage payment process receipt return tax pay total transact user
6	2	***Invest*** receipt total	Million
7	1	Centre	Agent application business call ***citizen*** ***govern*** manage order org research science service study support technology tic
7	2	Org science tic	Centre
8	1	Clear	Authenticate automate check create develop example ***govern*** house need problem public represent settlement
8	2	House settlement	Clear
9	1	Look	Company detail differ face ***govern*** import make record state study take user

9	2	Face	Active adapt business challenge ***govern*** learn manag new problem user
10	1	Visual	Complex environment group ***identify*** inform provide represent sophist technology tourism
10	2	Sophist	Automate available technology visual
Examples of three level relationship			
11	1	Supply	Available chain control ***govern*** inform measure power process public side vendor
11	2	Chain	Adopt certify example hash indicate see supply ***trust*** valid value
11	3	Hash	Chain check compute function message usage value
12	1	Workflow	Agent base change diagram engine exchange execute language process transact
12	2	Diagram	Active case flow jail nature process relationship uml workflow
12	3	Uml	Base diagram language result semantic
13	1	Item	Inform institute interest list new return unlink work
13	2	Unlink	Case identity item pseudonym
13	3	Pseudonym	Anonym be data generate refer system technology unlink user
14	1	Finance	Department ***govern*** ministry partnership review
14	2	Ministry	Finance justice order
14	3	Justice	Elect ministry
15	1	Loss	Corrupt data service theft
15	2	Theft	Destruction identity loss
15	3	Destruction	Theft

APPENDIX D

Comparison of my thesis work with related work, that support my thesis contributions:

Researcher	Kayed	Dritsas, Gymnopoulos, Karyda, Balopoulos, Kokolakis, Lambrinoudakis and Katsikas	Herborn and Wimmer	Moulin, Bettahar, Barthes and Sbodio	Mohammad Nizar
Year	2005	2006	2006	2007	2009
Describing Requirements Engineering Process	Yes	Yes	Yes	No	Yes
Describing E-Gov. Applications	Yes	Yes	Yes	Yes	Yes
Describing Ontology Life Cycle	Yes	Yes	Yes	No	Yes
Building Ontology Concepts	Yes	Yes	Yes	Yes	Yes
Making Refinement of Concepts	No	No	No	Yes	Yes
Extracting Relationships between Concepts	Yes	Yes	No	Yes	Yes
Represent Relationships via Lattice Representation	No	Yes	No	No	Yes
Make Evaluation Using Covering Technique	No	No	No	No	Yes
Using Tools to Extract Concepts and Relationships	Yes	No	No	Yes	Yes