e-Commerce Ontology Web

(e-COW)

by

Mohamad Fairus b Zainol Abidin

Dissertation submitted in partial fulfillment of

the requirements for the

Bachelor of Technology (Hons)

(Business Information Systems)

May 2006

Approved by:

Ms. Mazeyanti bt Ariffin

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

e-Commerce Ontology Web
(e-COW)

by

Mohamad Fairus b Zainol Abidin

A project dissertation submitted to the
Business Information System Programme
Universiti Teknologi PETRONAS
in partial fulfillment of the requirements for the
BACHELOR OF TECHNOLOGY (Hons)
(BUSINESS INFORMATION SYSTEMS)

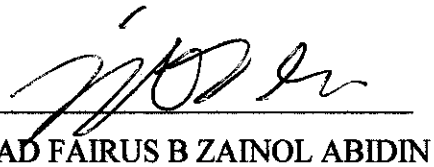
Approved by,

(Ms. Mazeyanti bt Ariffin)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
May 2006

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



MOHAMAD FAIRUS B ZAINOL ABIDIN

ABSTRACT

The objective of "e-Commerce Ontology Web" (e-COW) is to implement the ontology concept and semantic web in knowledge management to improve business operation in organization. Generally, current problem of e-commerce system is searching information is very difficult and sometimes may miss relevant information especially details for certain products. Furthermore, human browsing and reading required to extract relevant information from information sources. The semantic web deals with important application areas such as knowledge management and electronic commerce (both B2C and B2B). It may help to overcome many of the current bottlenecks in these areas. In feasibility study, it helps author in identifying the framework of semantic web and ontology development process. The important feature of the e-COW is retrieving information based on the class and attributes of data in tree views, and advance search using ontology agents. In addition the system will display products details in meaningful and organized which can help users to understand easily. With development of many knowledge management systems today, it shows that the knowledge management is very important especially to organization. As a result, we know ontology has become important concept in knowledge management especially for managing and organizing information and data.

ACKNOWLEDGEMENT

Bissimillah Ar Rahamani Ar Raheem

In the name of Allah, the Most Gracious, the Most Merciful

The success of the existence of the Final Year Project (FYP) is not the achievement of a single individual. Many have contributed their time and energy to ensure the development of the project ran smoothly.

First and foremost I would like to recite my greatest gratitude to the Most Merciful Allah for giving me the opportunity in completing this FYP on time and without much hassle or problem, although there might major problem which can result in my delay of turning the assignment in the time constrain.

I would like to take this opportunity to express my deepest gratitude to these individuals for their assistance and guidance. Their contributions are very valuable and may Allah reward them.

My gratefully acknowledges:

1. To my utmost gratitude Ms. Mazeyanti bt Mohd Ariffin who act as my FYP Supervisor for her guidance in helping me to produce a good output and also fully support me in terms of knowledge and understanding..
2. To Ms. Sharifah Bahiyyah and others KM Supervisors who are willing to give me fully support and comments for me to improve the project.
3. Special thank to my parents and families for giving me the full moral support in completing the BIS degree in general and this report in more specific.
4. To my fellow friends, they have been there for me during the good and bad times as we stay together under the same varsity.
5. To other personnel who did contribute to this research that I have not mentioned their name but did provide their knowledge and experienced also deepen my knowledge of the issues.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

CHAPTER 1
INTRODUCTION

1.1 Background of Study

As we know, there are several billion documents on the World Wide Web (WWW), which are used by more than 300 million users globally, and millions pages on corporate intranets nowadays. The popularity of the Internet and the huge growth of new Internet technologies in recent years have brought about the creation of many e-commerce applications. It continued to rapidly growth in information volumes make it increasingly difficult to find, organize, access, and maintain the information required by users especially those users who are using e-commerce web site.

The concept of a Semantic Web that provides enhance information access based on the exploitation of machine-process-able meta-data has been proposed. Technology is not the only key factor for the development of the current e-applications. But the context of e-commerce especially requires an effective communications between machines. Besides, the practice of knowledge management is about culture not only technology because the technology is just an enabler and knowledge management is a journey.

As a consequence for this situation, the author has taken the initiative to develop a simulation program (web based) for an e-commerce application that using concept of semantic web and ontology. The application is known as eCommerce Ontology Web (e-COW). This program shows on how the meaningful information can be retrieved for example producer names can be traced in short time by identifying the properties of the products with accuracy searching process and also how the information related between customer needs and wants, and supplier details, through intelligent searching process by providing framework to identify products and services in global markets.

1.2 Problem Statements

Internets have an important role to play in the more effective exploitation of both explicit and tacit knowledge. The changing information from paper to the internet can also benefits in term of speed of update and hence accuracy. The traditional search engines return ranked retrieval list that offer little or no information on the semantic relationships among the documents. Important information is often scattered across web and/or internet resources. So, the issues is how the user can retrieve the information from the internet especially from e-commerce and avoid overlooked the information.

1.2.1 Problem Identification

The current e-commerce application or knowledge management systems have several weaknesses such as:

- **Searching information**. Existing keyword-based search retrieves irrelevant information that the keyword in a context than the one in which the searcher is interested or may miss relevant information that employs word other than the keyword in discussing the desired content.
- **Extracting information**. Human browsing and reading is currently required to extract relevant information from information sources.

Furthermore, HTML also become one of the problems in having an excellent capabilities of e-commerce application, although HTML exercises powerful control over the appearance of a Web page but HTML does not provide a means of presenting rich syntax and semantic of data or does not present the content in a way that another software can easily understand and use

1.2.2 Significance of the Project

This project is significant with the development of current e-commerce application because it gives a great impact in managing and organizing information and data within the e-commerce application or website. Users around the world are always requesting meaningful information to make decisions but with the help of ontology concept, data can be more meaningful and understandable.

Derived from that, the author took the initiative to create an e-commerce application with ontology concept to manipulate data and information into hierarchically form or semantic representation of information for products catalogue. Based on hierarchically representation, user can find better information about related products within less time. Besides, the user can find others products with the same characteristics or even same manufacturer by identifying the class or category of the products.

1.3 Objectives and Scopes

As for this project, the author has defined two main objectives. The objectives derived from the background of study and problem statement as solutions for those particular problems. The general objectives are, first, is to develop a simulation program for better retrieving data and the second ones is to provide a meaningful information for better decision making. This project focuses on the supermarket products such as apparel.

1.3.1 Objective of the Project

The two objectives of the project which have been mentioned above are stated in this sub-section as below:

- To develop an e-commerce application for better retrieving data.

  Retrieving information is very crucial part for this project. So, through process underlying the ontology concept and hierarchically representation, the user can retrieve the information more easily. Furthermore, the user will more understand about the characteristics of the products such as the color, name, weight and other and user also can know the category of the products. When the products have been classified based on category, it will assists users to find other options or products if the products they want are not available for consumption.

- To provide a meaningful information for better decision making

Decision making in buying process is always become a crucial part in business transaction. So, better information or meaningful information is always important in helping individual to make decisions especially in buying a certain products. Through this application, it will help user to make decision making for buying certain via intelligent searching.

As general conclusion for objectives, user can easily get the accurate information with what they have requested before and can find another alternatives if the products are not available in the supermarket. Therefore to make those objectives successfully, the understanding in ontology concepts, hierarchically data representation, components of semantic web, frameworks of semantic web and also understanding the process of e-commerce is highly needed.

1.3.2 Scope of study

The scope of this project is to develop a simulation program of ontology and semantic web application for e-commerce. The aspects which have been considered are stated as below:

- To conduct a research on identifying the component of semantic and also component of ontology and how they works.

In early stage, the research is conducted to identify the main components of semantic web and also main components of ontology. Based on the research, the author analyzes on how components of semantic web and ontology work in improving the process of retrieving information and creating hierarchically representation of information.

- To understand what are semantic web and ontology

Based on the earlier the research, the author comes out with his own understanding about what is semantic web and ontology and how it helps the e-commerce application. Based on the outcome of the research, the author defines the requirements of semantic web and continues with developing the simulation program.

- Easy to use system

Among the platform, author has used web-based application because almost everyone knows about web application, so through that it helps a lot and reduce time consuming for author to teach the default navigation and integration in web based application.

- Focus on apparel products.

The primary focus for the simulation program is apparel products or known as apparel ontology. So the system uses the apparel product as a primary product for the development of e-commerce application using ontology and semantic web.

Basically, the program shows the data in hierarchically information and it demonstrates on how the meaningful information can be traced and retrieved from the apparel ontology.

CHAPTER 2
LITERATURE REVIEW AND THEORY

2.1 Literature Review

In literature review section, the author will analyze the details and findings of experts and journalist which related to the projects. It based on the research and the process of extracting information from reading the journals and paperwork of others expert people.

2.1.1 Semantic Web

As described by **Berners-Lee (2001)**, the semantic web is '*an extension of the current Web, in which information is given well defined meanings, better enabling computers and people to work in cooperation*'.

The semantic web is meant to enable an environment, in which independent, where inter-connected information system can exchange the knowledge and action specifications, and resulting in execution of an activity acceptable to all the system involved. In practice, the semantic Web is a collaborative effort led by the World Wide Web Consortium (W3C) with the participations of a large number of researchers and industrial partners. W3C defines the semantic web as the representations of data on the World Wide Web. It is based on the Resource Description Framework (RDF), which integrates a variety of applications using XML (extensible markup language) for syntax and URIs for naming (www.w3c.org).

2.1.2 Ontology

The ontology has been described by Neches and colleagues (1991), who defined ontology as " *An ontology defines the basic terms and relations compromising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary*".
The descriptive definition tells what to do to build and ontology, and gives some fuzzy guidelines. This definition identifies the basic terms and relations between terms, identifies rules to combine terms, and provide definitions of such terms and relations.

Besides that, Kai Zhang and colleagues (2004) has defined ontology "as an explicit specification of a conceptualization".
Ontology unifies all the concepts and relations of domain. In many practice applications, we often need to integrate some existed ontologies to be a unified ontology. The target ontology commonly constructed for special ontology application. So we should consider the application requirements in ontology integration procedure.

Peter Carp (2000), defined ontology as "*ontology is a specification of a conceptualization that is designed for reuse across multiple applications. By conceptualization, we mean a set of concepts, relations, objects, and constraints that define some domain of interest.*"
Based on that, we know the ontology is an explicit description of a domain that describe concept, properties and attributes of concepts, constraints on properties and attributes and individuals *(often, but not always)*. Furthermore, we know that ontology can be reused across multiple applications.

2.2 Theory

Theory is mainly to elaborate and explain related field topic that helps author in the development process of ontology. The theory is derived from the research and analyzing the information via books, internets and journals.

2.2.1 Semantic Network

Organizing the information and data is very crucial in providing an excellent retrieval system. So the author has conducted analyzing process on how data should be managed appropriately for better performance. The author makes a conclusion that data can be represented in a semantic network for easy retrieval. Besides, the data that represent in semantic diagram can be structured so that it easier and faster to retrieve. The data can be organized as a five-level taxonomy of products, each level a two-character numerical values (as identifier) and a textual description. These levels are defined as follows:

- *Segment*. The logical aggregation of families for analytical purposes.
- *Family*. A commonly recognized group of inter-related commodity categories.
- *Class*. A group of commodities sharing a common use or function.
- *Commodity*. A group of products or services that can be substituted.
- *Business Function*. The function performed by an organization (e-commerce organization) in support of the commodity. (See figure 2.1).

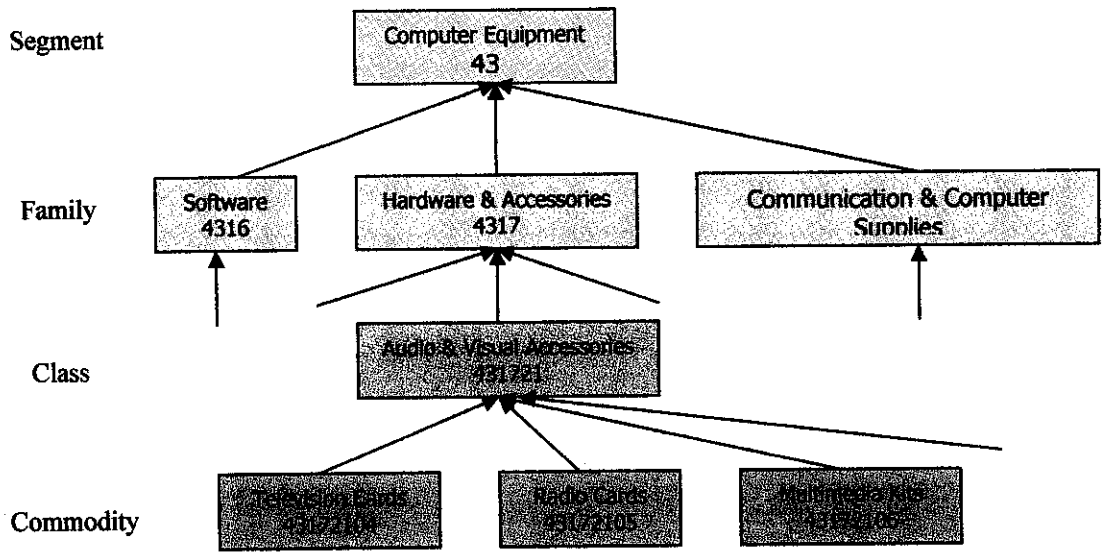


Figure 2.1: Semantic Network

2.2.2 Semantic Web Knowledge Based Architecture

By reading and analyzing the materials, the author has figured out the architecture for semantic web-based knowledge management. Figure 2.2 (see appendix) shows the architecture for knowledge management based on the semantic web. The architecture addresses all the key stages of the knowledge management life cycle which are knowledge acquisition, knowledge representation, knowledge maintenance and knowledge use. The explanation for each of the knowledge management life cycle as below:

- Knowledge Acquisition

Given the large amount of unstructured and semi-structured information held on organizational intranets, automatic knowledge extraction from unstructured and semi-structured data in external data repositories is required and this is show in the bottom layer of the figure 2.2. Support for knowledge acquisition is also needed and the knowledge engineer needs to be supported by ontology editing tools which support the creation, maintenance and population of ontologies.

- Knowledge Representation

Once knowledge has been acquired from human sources or automatically extracted, it is then required to represent the knowledge in an ontology language (and of course to provide a query language to provide access to the knowledge so stored). This is the function of the ontology repository.

- Knowledge Representation

Ontology middleware is required with support for development, management, maintenance, and use of knowledge bases.

- Knowledge Use

Finally, and perhaps most importantly, information access tools are required to allow end users to exploit the knowledge represented in the system. Such tools include facilities for finding, sharing, summarizing, visualizing, browsing, and organizing knowledge.

2.2.3 Evaluating Current System

There are many e-commerce ontologies systems throughout this world. The author evaluates most of the system in order to understand the concepts rely on the current systems. There are five e-commerce ontologies has been evaluated by the author which are:

- United Nations Standard Products and Services Codes (UNSPSC)
- North America Industry Classification (NAICS)
- Standard Classification of Transported Goods (SCTG)
- E-cl@ss
- RosettaNet

For the **United Nations Standard Products and Services Codes (UNSPSC)** has been created by the United Nations Development Programme (UNDP) and Dun & Bradstreet. UNSPSC is a global commodity code standard that classifies general products and services and is designed to facilitate electronic commerce through the exchange of product descriptions.

Initially the UNDP managed the code of the Electronic Commerce Code Management Association (ECCMA). This partnership finished, ans as a result there are now two different version of the UNSPSC: the United Nations Standard Products Codes owned by UNDP, and the Universal Standard Product and Services Classification managed by ECCMA. In October 2002, both organizations signed an agreement in which they proposed to have on single version of the classification, which has marked the beginning of the UNSPSC unification project.

The UNSPSC coding is organized by five-level taxonomy of products, each level containing a two-character numerical value and a textual description. These levels can be defined as follows:

- *Segment*. The logical aggregation of families for analytical purposes.
- *Family*. A commonly recognized group of inter-related commodity categories.
- *Class*. A group of commodities sharing a common use or function.
- *Commodity*. A group of products or services that can be substituted.

- *Business Function.* The Function performed by an organization in support of the commodity.

For UNSPSC version 6.0315 version, it contains about 20,000 products organized in 55 segments. Segments 43 For instance which deals with computer equipment, peripherals and components, contains about 300 kinds of products.

Thee **NAICS** was created by the Census Office of USA in cooperation with Economic National Classification Committee of USA, Statistics Canada and Mexico's Instituto Nacional de Estadistica, Geografica e Informatica (INEGI). It classifies products and services on general, and is used in USA, Canada and Mexico. NAICS was developed after the Standard Industrial Classification (SIC) was revised. SIC was created in the 1930s to classify establishments according to the type of activity they were primarily engaged in and to promote the comparison of their data describing various facets of the US economy.

NAICS products are identified by means of a six-digit code, in contrast to the four-digit SIC code. The NAICS code includes a greater number of sectors and permits more flexibility to design subsectors. It also includes a greater number of sectors and permits more flexibility to design subsectors. It also provides additional details not necessarily appropriate for all three NAICS countries. The international NAICS agreement fixes only the first five digits of the code. The sixth digit, when used, identifies subdivisions of NAICS industries that consider the user's needs in individual countries. Thus, six digit US codes may differ from counterparts in Canada or Mexico, but up to the five-digit level they are standardized. The general structure is:

XX	Industry Sector (20 broad sectors up from 10 SIC)
XXX	Industry Subsector
XXXX	Industry Group
XXXXX	Industry
XXXXXX	US, Canadian, or Mexican National specific.

For **SCTG** was sponsored by the Bureau of Transportation Statistics (BTS). It is a product classification for collecting and reporting Commodity Flow Survey (CFS) data. SCTG was developed by US Department of Transportation's (DOT), Volpe National Transportation Systems Center (Volpe Center), Standards and Transportation Divisions of Statistics Canada, US Bureau of Census (BOC), and the US Bureau of Economic Analysis (BEA).

This classification has four levels, each of which follow two important principles. First, each level covers the universe of transportable goods, and second, each category in each level is mutually exclusive. The general structure is:

XX	Product Category
XXX	Commodities or Commodity Groups (different is US and Canada)
XXXX	Domestic Freight Transportation Analyses
XXXXX	Freight Movement Data

The first level of the SCTG (two digits) consists of 43 product categories. These categories were designed to emphasize the link between industries and their outputs. The second level (three digits) is designed to provide data for making comparisons between the Canadian goods and US goods. Categories specified at this level consists of commodities or commodity groups for which very significant product movement levels have been recorded in both the US and Canada. The third level (four digits) is designed to provide data for domestic freight transportation analyses. Four-digits categories may be of major data significance to either the US or Canada, but not necessarily for collecting (and potentially reporting) freight movement data. Product codes at this level have been designed to create statistically significant categories for transportation analysis.

E-cl@ss is a German initiative to create a standard classification of material and services for information exchange between suppliers and their customers and companies such as BASFG, Bayer< Volkswagen-Audi, SAP, etc., make use of it.

The e-cl@ss classification consists of four levels of concepts (called material classes), with a numbering code similar to the one used in UNSPSC (each level adds two digits to its previous level). These four levels are: segments, main group, and group and commodity class. Inside the same commodity class we can gave several products (in this sense, several products can share the same code).

E-cl@ss contains about 12,000 products organized in 21 segments. Segment number 27, which deals with Electrical Engineering, contains about 2,000 products. The main group 27-23, which deals with process control systems and with other computer devices contains about 400 concepts.

E-cl@ss provides a set of attributes for every product that is a leaf in the classification. The set of attributes is an addition of individual characteristics describing the related commodity. This set distinguishes e-cl@ss from UNSPSC and offers a solution to the shallowness of that. For example, PC system (with code 23-01-99-03) has attributes like product type, product names, etc., in e-cl@ss.

The e-cl@ss search tool, which is available on-line, allows finding terms with an interface in different languages (German, Spanish, English, and Czech). In fact, the terms found are presented in any of these languages. The e-cl@ss classification can also be downloaded from same URL.

The **RosettaNet** classification has been created by RosettaNet, which is a self-funded, non-profit consortium of about 400 companies of Electronic Components, Information Technology, Semiconductor Manufacturing and Solution Provider companies. It started within the IT industry, then RosettaNet is currently being expanded to other vertical areas, notably the automotive, consumer electronics and telecommunications industries.

The RosettaNet classification does not use a numbering system, as UNSPSC does, but is based on the names of the products it defines. This classification is related to the UNSPCS classification and provides the UNSPSC code for each product defined in RosettaNet. This classification has only two levels in its product taxonomy:

- RN Category. A group of products, such as Video Products.
- RN Product. A specific product, such as Television Card, Radio Card, etc.

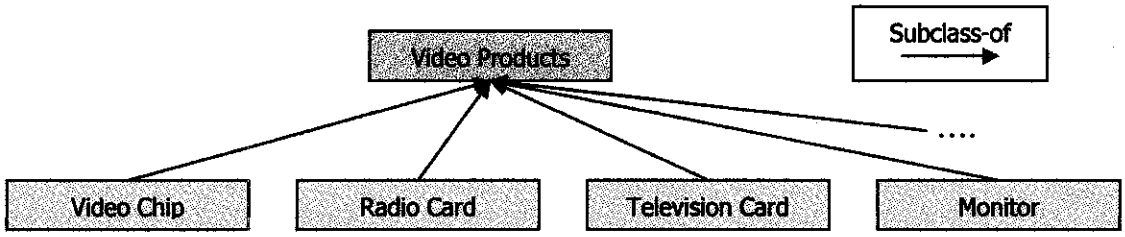


Figure 2.3 Partial view of RosettaNet Classification

The RosettaNet classification consists of 14 categories and about 150 products. It should be added that RosettaNet is more specific than the UNSPSC classification. Figure 2.3 shows a small section of the RosettaNet classification related to video products for computer equipment. Unlike in the previous formats, the order of contents here is of great importance, since the relationship between products and the category they belong to are given by the order in which they appear. Hence Monitor, Radio Card, Television Card, and Video Chip are products from the category Video Product in the RosettaNet classification.

CHAPTER 3
METHODOLOGY

3.1 Procedure Identification

For developing ontology for this project, the author has analyzed Ontology Development Process, mainly for developing ontology for products. The ontology development process was identified on the framework of METHONTOLOGY methodology for ontology construction. The ontology development process refers to which activities are performed when building the ontologies. It consists of three development phases and every development has its own stage. The three phases and its stages are:

- Pre-development
    - Environment study
    - An environment study it carried out to know the platforms where the ontology will be used, the applications where the ontology will be integrated, etc.
    - Feasibility Study
    - The feasibility study is to determine whether it is possible to build the ontology and also to check whether suitable to build the ontology
- Development
    - Specification
    - The specification activity states why the ontology is being built, what is intended uses are and who the end-users are.
    - Conceptualization
    - The conceptualization activity structures the domain knowledge as meaningful models at the knowledge level.
    - Formalization

- The formalization activity transforms the conceptual model into a formal or semi-computable model.
- Implementation
- The implementation activity builds computable models in an ontology language

- Post development
  - Maintenance
  - The maintenance activity updates and corrects the ontology if needed.
  - Use
  - The ontology is (re) used by other ontologies or application

The ontology development process identifies which activities are to be performed. However, it does not identify in which activities should be performed. The ontology life cycle identifies when the activities should be carried out, that is, it identifies the set of stages through which the ontology moves during its life time, describes that activities are to be performed in each stage and how the stages are related .

Based on ontology development process, the author has come out with appropriate methodology which is Prototype Model to conduct the overall progress of this project in better way. The reason is because Prototype Model can reduce the amount of time until user begins to see a working prototype. It consists of seven phases the overall development of this project has gone through which are:

- Feasibility Study
- Requirement Definition
- System Design
- Coding, Testing
- Implementation
- Maintenance

As shown, in Figure 3.1: Prototype Model, first stage is the initial investigation. This is where the author did an analysis and initial investigation of this project. Then in the requirement definition the author needs to clarify the methods that be used in this project and also need to identify the approach that used for this project. After that the coding and testing stage will be implemented. Then the system

design will be applied. The requirement stage, coding and testing stage, and system design stage is a round stage to make the project well. After satisfied with the system, the system can be implemented and maintenance stage needs to be done frequently in order to make sure it does not have any problem.

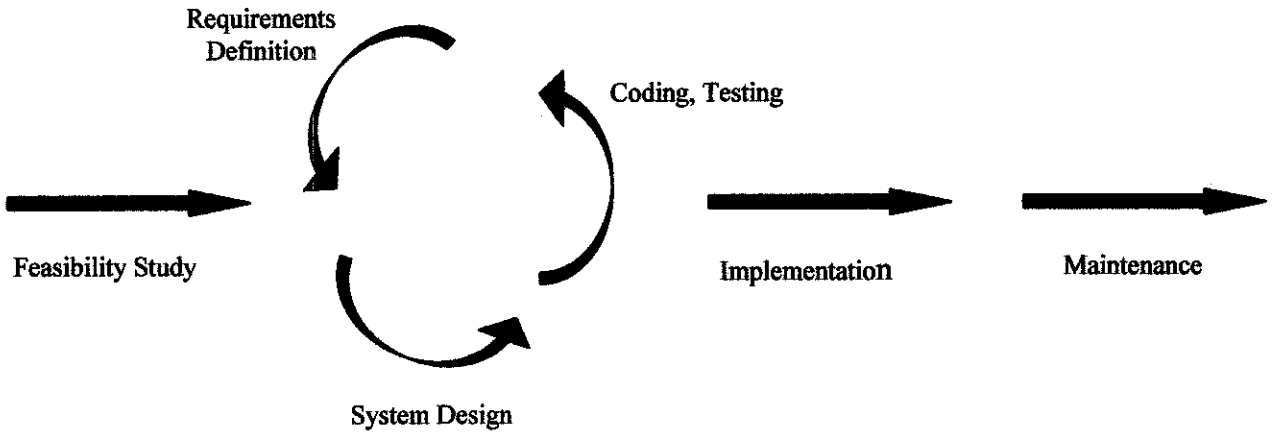


Figure 3.1: Prototype model

The reason the author uses this model because, although the objective for this project is clear where the author wants to has a web based application which demonstrate on how the ontology and semantic web help the users in retrieving the exact information and making decision on buying certain products but the requirements for this project are uncertainty and not so unclear. So, when using this methodology, the requirements for this project can always been reviewed in order to make the system in a right track and also meet its objective.

There are several tasks need to be performed in order to develop ontology by understanding the ontology components. The ontology components include concept, attributes, relations, constants, formal axioms, rules, and instances which attached to each task, and illustrate the order proposed to create such components during conceptualization activity or system design. The tasks that been carried out to create and ontology as stated as below (see figure 3.2):

- Build glossary of terms

To build the glossary of terms that identifies the set of terms to be included on the ontology, their natural language definition, description and type.

- Build concept of taxonomies

To build concept taxonomies to classify concepts. The output of this task could be one or more taxonomies where concepts are classified.

- Build concept of dictionaries

To build concept the concept dictionary, which mainly includes the concept instances for each concept, their instances and class attributes.

- Describe instance attributes

To describe details each instance attribute that appears on the concept dictionary. The result of this task is the table where class attributes are described.

- Describe class attributes

To describe in details attribute that appear on the concept dictionary. The result of this task is the table where class attributes are described.

- Describe instances

To describe relevant instance that appear in the concept dictionary inside an instance table.

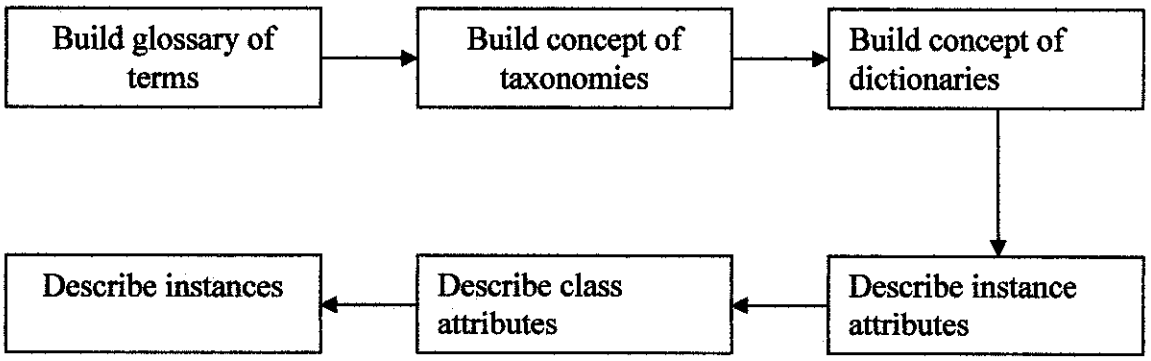


Figure 3.2: Tasks for developing ontology

This task has been carried out to create ontology for product based on their categories. For example the author these steps to develop apparel ontology by implementing every step. These steps will be discussed in more details in sub section of System Design Phase within this chapter.

3.1.1 Feasibility study

In this phase, the author conducts initial study especially related with ontology and semantic web. As same with pre-development phase in Ontology Development Process model for environment study, the author has analyzed the suitable platform where the ontology will be used and in what environment the ontology will be put in. The proper analysis has been done and at the end of the preliminary research, the author concludes that the ontology for this e-commerce application is apparel products in supermarket. The main domain is of course e-commerce but the specific domain for this ontology and semantic web is apparel products.

The research data have been collected especially related with components, concepts, framework and methodology of semantic web and ontology. The information and data are being collected through journals and findings that have been done by previous researchers. The reason of getting the information through journals is not only because it is readily documented but also the results of the research are proven to be correct and success.

This is a very crucial phase in this Prototype Model, where by, this stage will direct the project direction and become the base for the whole project. In order to have a successful result of this project, all the elements in this stage need to be done carefully. At the early phase of this life cycle, the author focused on the fundamental of this project and the principal of this project in order to bring this project proper direction. Among the element that being consider as the fundamental elements in this project are the project background for this project, the background project that lead to the initiation of this project, and then the author come out with the problems statement that will become a backbone for this project. The reason why the problems statement becomes the backbone for this project is because the existing of this project is to solve those problems that been identified earlier

The end output for this phase is theory and concept of ontology and semantic web, the preliminary identifying of framework for semantic web and ontology, the preliminary identifying for methodology, and the project timeline. The tasks and procedure are part of scope definition the scale of the project..

The next phase is sequence from this phase because the requirement for this project is not very clear and need to be review time by time so the author have used the method prototype model as methodology for this project after manipulating and integrating with ontology development process

3.1.2 Requirement Definition

The requirements for this project are not clear but the objective is well known. So the author has reviewed the journal and understanding the needs of technology today. Today, people are concerned with technology. They need a system that can assist them in making decision and they need an accurate information and meaningful information. From that author concludes that the system should have an intelligent search to assists users and also to have an accurate search engine.

From the requirements that been identified through analyzing time by time, the author has determined the tools such as hardware and software that will be used for project development. All the tools will be discussed in details in sub section Tools within this chapter. The end outputs for this phase are objectives and scope of studies and also the fully methodology has been defined.

3.1.3 System Design

In this phase, the author designs the architecture based on the framework that been identify in early phase for e-commerce application. This includes analyzing the available techniques to implement the best e-commerce ontology web system. This design should be a plan for implementing the identified requirement. When and only when the design is fully completed, then the coding and testing will take place and after that, the implementation phase can take place. This shows how important the designing phase for the development of the system and also the implementation of the system itself. Besides that the designing phase is really rely on the requirement definition and initial investigation phase that has been completed earlier.

The purpose of the design phase is to transform the requirement definition into design specifications for system coding and testing. As the purpose of this project several elements need to be design in order to have a good view of this system before the development phase take part. Among the elements that are needed to be included in designing the phase are system architecture and the system design document is necessary to ensure the development can be sustained during the project's progress. These design requirements are representation of system flow in terms of entities, attributes and relationships and rules. For development of ontology, author has follow the task that been mentioned at the early chapter.

3.1.3.1 Ontology design

Ontology design is the main focused for this project. Without the ontology, there are no points to continue developing of this ontology. The design for the ontology has been carried out based on the tasks for ontology development. There are six steps need to be performed for development of ontology which are build glossary of terms, build concept of taxonomies, build concept of dictionaries, describe instance attributes, describe class attributes, and describe instances (see figures 3.2).

- To build the glossary of terms.

  First, the author builds a glossary of terms that includes all the relevant terms of the domain (concept, instances, attributes that represent properties, relation), their natural language descriptions.. It is important to mention that on the initial stage of designing the ontology. The glossary might contain several terms that give a meaningful understanding also help the author to identify the terms to be used in that particular ontology.

- To build concept taxonomies.

  When the glossary of terms contains a sizable number of terms, the author builds concept taxonomies to define the concept hierarchy. The three approaches can be used to show hierarchy diagram which are top down, bottom up, and middle out. To build concept taxonomies, the author select terms those are concepts from the glossary of terms. For this, it is really

important to identify in the concept taxonomy sets of disjoint concept, that is, concepts that cannot have common instances. We can also identify the subclass of several terms by knowing the relationships. Figure 3.3 illustrates how the concept taxonomies have been built.

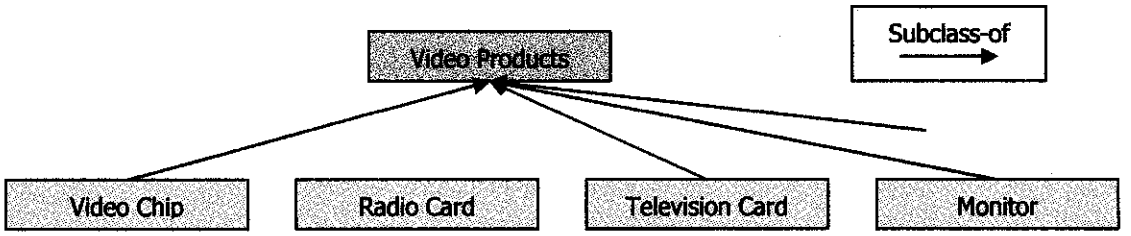


Figure 3.3: Taxonomies Concept or Concept Hierarchy for Video Product

- To build the concept dictionary

  Once the concept taxonomies have been generated, the author specifies which are properties and relations that describe each of concepts of taxonomy in a concept dictionary, and their instances. A concept dictionary contains all the domain concept, their relations, their instances, and their class and instance attributes. The relations specified for each concept are those whose domain is the concept. Relations, instance attributes and class attributes are local to concepts.

- To define instance attributes in detail.

  The main aim of this task is to describe in detail all the instance attributes which already included in the concept dictionary by means of an instance attribute table. Each row of the instance attribute table contains the detailed description of an instance attribute. Generally, instance attributes are those attributes whose value may be different for each instance of the concept. For each instance attribute, the author must specify the following fields, its name (the concept it belongs to – attributes are local to concept), its value type, its measurement unit, default value if they exists, minimum and maximum cardinality, instance attributes, class attributes, and constants used to infer values of the attribute.

- To build the class attributes in detail.

  Basically, the aim of this task is to describe in details all the class attributes already included in the concept dictionary by means of class attributes. Each row of the class attribute table contains a detailed description of the class attribute. Unlike instance attributes, which describe concept instances and take their values in instances, class attributes describe concepts and take their values in the class where they are defined. Class attributes are neither inherited by subclasses nor by the instances. For each class attribute, the author fills the following information: name, the name of concept where the attribute is defined, value type, value, measurement unit and cardinality.

- To define instance.

  When the conceptual model of the ontology has been created, we can define relevant instances that appear in the concept dictionary inside an instance table. For each instance, it contains its name, the name of the concept it belongs to, and its attributes values, if known.

3.1.3.1 e-COW Architecture

Every system must its own system architecture or architecture in order to get better understanding. Based on the analyzing all the information and resource that been on chapter 2 discussed previously, author had concluded the architecture for the system that will be developed. It consists of four areas which are user interface for information access and information retrieve, e-COW Editor for system administrator to maintain the system, Ontology Agent where the stimulation program will be processed and database repository

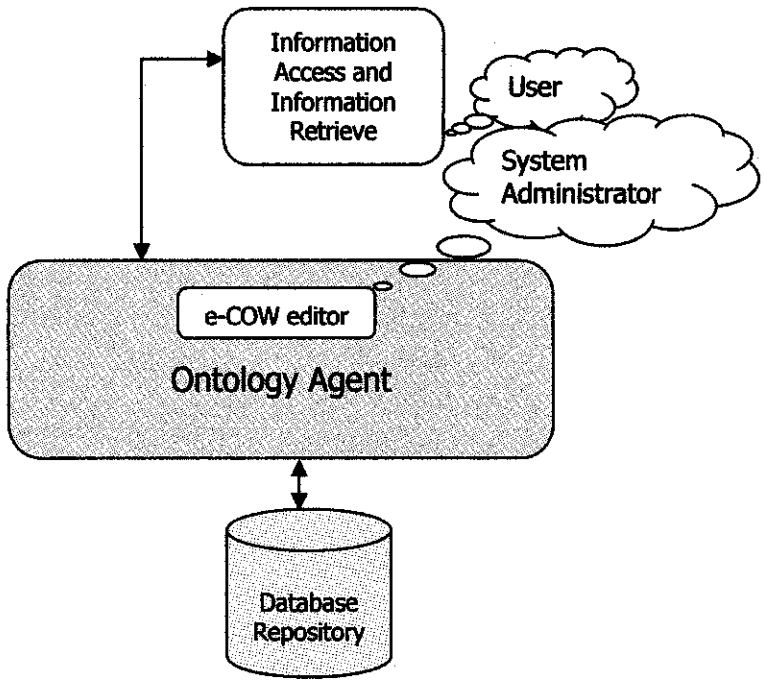


Figure 3.4: Architecture of e-COW

- Information Access and Information

In section Information Access and Information Retrieve, the user will enter the appropriate input for searching the information, for example Computer (see 3.4) .So, the query will be send or recognize by the Ontology Agent for processing. After the query has been processed the output will be displayed in showing hierarchy representation.

- e-COW Editor

e-COW Editor basically a programming language that will create ontology description for ontology agent. Author used protégé to evaluate ontology that been developed..

- Ontology Agent

The ontology agent will stimulate the query and perform the process. Based on the XML documents that been generated the query, the agent will use XML Schema to identify the taxonomy level of the query. XML Schema has a similar essential role with Document Type Definition (DTD) which is to define a grammar for XML document. Basically, a DTD defines a grammar to specify allowable combinations and nesting of tag names, attribute names and so on. Based on the taxonomy level, the ontology agent can request the information from the database

- Database Repository

The information will be stored based on the taxonomy level. There are five level taxonomy levels which are segment, family, class, commodity and business function. This database stores data based on the information that will created before through the ontology design tasks.

3.1.3.4 System Workflows

There are the process workflows for the e-COW. Figure 3.5 (see appendix) shows the general workflow for the system. The details for the workflows are as follows:

i. The user will view the product details by follow the navigation within the system. At first, user will see the landing page for the e-Commerce Ontology Web. The users will be welcome to use the system for assistance in helping them to search certain products. There is no security part for this system where the users need to enter the password because this system is for general use.

ii. After that user need to select either to choose advance search or normal search. Advance search is an intelligent search where by its help user to search a certain product more specific and accurate and its use ontology agent to find that particular products. Whereas for normal search, it guides user to find the available product in the supermarket.

iii. Advance search particular mainly for guiding user in more specific items or products. Basically is an intelligent search that used ontology agent to find more specific products for users. Example if the user wants to find a present for his/her parents. Then the ontology agent will ask certain criteria as inputs in form of set of cases as guidelines to search those specific products. It will give available results for user to select and view those particular products.

iv. Normal search particular is normal base searching where the users need to search based on the items they want. There are two types of search mode which are description search and hierarchy search. For description search will display a text box asking for user to enter a specific keyword for searching the products. The keywords can be classes, subclasses and

instances. Whereas for hierarchy search, it will display the information in hierarchy model where the information in tree view.

v. When the users select the result it will display the description about those particular products.

3.1.3.4 Use Case Modeling

The Use Case in Figure 3.6 describes the actions that occur between the users and the system. Basically, there are two users for this systems but only one is applicable in production environment which is customer and the other one system administrator or ontologist that will monitor the ontology agent. The customers for the supermarket will become primary user for this system where they will use the system in helping them for purchasing the products. But another entities that will communicate with the customers is ontology agent that been monitored by system administrator. Ontology agent will become the entities to provide relevant searching information to customer.

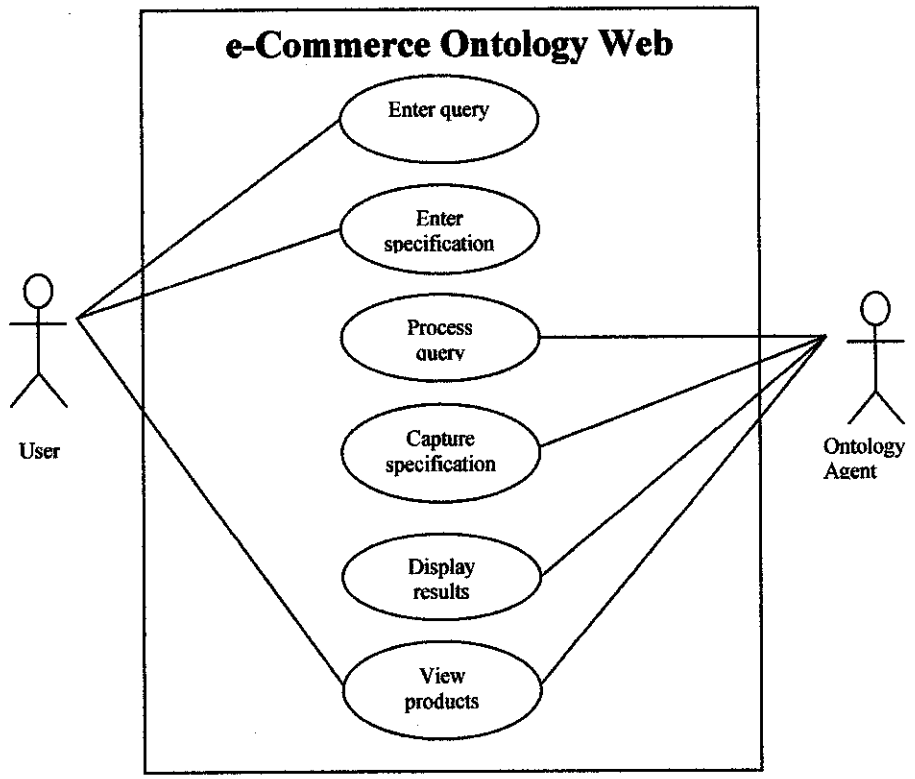


Figure 3.6: UML Use Case Diagram of the System

Customer will enter the query for searching products or use the navigation for hierarchy display of the products. Besides that, customer can also use advance searching where ontology agent will help the user to find specific or meaningful information about certain by entering the specification in advance searching.

For ontology agent, it will process the query entered by user. Then for advance search, it will capture the specification given by user for reasoning and processing the query for specific information. After that ontology agent will display all possible result to users and guiding them for better decision in purchasing the products. Ontology agent will display the products details based on the customer action to view the product details.

3.1.4 Coding, Testing

With the designed specification, the next phase is developing phase and testing out the prototype which are Coding and Testing Phase. The purpose of Coding and Testing phase is to build a system that fulfills system requirements and design specification and to implement the interfaces between the modules. Among the task to be implemented are designing the database, web application programming and user interfaces.

Testing is very essential to keep track the requirement is available to fulfill or not. So every aspect has been tested during development to see whether to see the requirement is suitable or not with the system developments. The prototypes for this system are developed using ASP.NET and XML Dom and ontologies design is tested using Protégé version 3.2. Testing has been conducted by checking the ontology design and also the efficiency and the effectiveness of the prototype in performing the query functions.

3.1.1 Implementation

After the testing is done, the prototype now advances to this phase where the implementation for the prototype will be done. For this phase the prototype will be operate in operating environment, that mean the system can be used by end users (customers).

3.1.1 Maintenance

When the system in operating environment, it need to be monitored frequently to make sure the system is up to date in terms of products especially the ontology of products and also the ontology agents. The maintenance process can be done for every quarter of year to make system is reliable and usable to users.

3.2 System Tools

Basically there are two main categories of tools that are necessary in this project development work, there are software and hardware. These equipments are also necessary to stimulate a real working environment of a running production system. However, some items are not available due to constraints, thus they are replaced by either emulated devices and assumed to work the same in real world.

3.2.1 Software

Among the software that been used by in this project are:

- Microsoft Windows XP Professional Edition
- Microsoft Visual Studio.NET 2003
- Microsoft SQL Server 2000 Developer Edition
- Microsoft Internet Information Services 5.1
- Microsoft.NET Framework 1.1
- Microsoft Internet Explorer 6.0
- Protégé 3.2

For the purpose of development work, Microsoft Windows XP Professional Edition is chosen because the platform offers more stability and realibility over other Windows platforms and others operating system. Furthermore Professional Edition is bundled with Internet Explorer 6.0 and Internet Information Services (IIS) 5.1 which is vital for the purpose of web hosting ASP.NET pages. For the purpose of database storage, Microsoft SQL Server 2000 Developer Edition is used because this edition allows developer to build any thype of application on top of SQL Server. It

includes all the functionality of Enterprise Edition but limited for the purpose of development and not production.

As for the development environment, the tool that is necessary is Microsoft Visual Studio.NET 2003. It comes bundled with Microsoft.NET Framework 1.1. Visual Studio.NET is a complete set of development tools for building ASP Web Application, XML Web Services and other programming tools, and all use the same integrated development environment (IDE), which allows them to share tools and facilitates in creation of mixed language solutions. In addition, these languages leverage the functionality of the .NET Framework, which provides access to key technologies that simplify the development of ASP Web applications and XML Web Services.

With the .NET Framework and ASP.NET, Microsoft has not just shown itself to be a contender in Web development technologies, but many commentators also believe Microsoft has taken the lead.ASP.NET is well equipped for any task you want to put to it, from building intranets to e-business or e-commerce megasites. With ASP.NET we now have a true choice of languages. All the .NET languages have access to the same foundation class libraries, the same type of systems, equal object orientation and inheritance abilities, and full interoperability with existing COM components. We can use the same knowledge and code investment for everything from Web development to component development or enterprise systems, and developers do not have to be concerned about differences in APIs or variable type conversions, or even deployment. ASP.NET incorporates all the important standards of our time, such as XML, it is arguably easier to implement than in any other technology, including Java. An ASP.NET programmer still only needs a computer with Notepad to write ASP code.

XML Document Object Model (DOM) that is the heart of XML definition which can be found in Visual Studio.NET. The DOM is a recommendation made by the W3C to enable a standard way of providing access to the data contained in an XML document. It is essentially a list of interfaces and class recommendations to allow anyone to internally implement an XML parser and expose a common interface to enable a developer access to his or her XML data. This interface is based on the idea

that the XML is in a hierarchical or tree-like structure, where certain nodes might contain child nodes, which might also contain child nodes, and so on. The DOM provides the capabilities to navigate this tree of nodes in a relatively simple programmatic manner. Some of the DOM API capabilities include:

- Find the root node in a XML document.
- Find a list of elements with a given tag name.
- Get a list of children of a given node.
- Get the parent of a given node.
- Get the tag name of an element.
- Get the data associated with an element.
- Get a list of attributes of an element.
- Get the tag name of an attribute.
- Get the value of an attribute.
- Add, modify, or delete an element in the document.
- Add, modify, or delete an attribute in the document.
- Copy a node in a document (including subnodes).

For Protégé, is an integrated software tool used by system developers and domain experts to develop *knowledge-based systems*. It helps user to create an applications for problem –solving and decision making for particular domain. The author uses this software is to create and to test for every of ontology have been created. It really helps in determines the concept (known as class in Protégé) and instance attributes (known as slot in Protégé).

As general definition, ontology is a specification of a conceptualization that is designed for reuse across multiple applications and implementations. Ontology also is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for knowledge base.

3.2.2 Hardware

A development machine is vital for the purpose of web hosting, storage, system development, testing and simulation. The hardware that will be used is:

i. Computer; act as a server

The server will be used to store the database of the system. The system will basically have a centralized database. All the information will be stored inside the database. The server will be center of all the information related to the system. The computer that is dedicated as the server for the system, will need to fulfilled several requirements. The requirements are:

- Pentium 4 – 1.4 Ghz
- 348 Mb of Ram
- At least of 20GB of hard disk space for operating system, relevant software installation

ii. Computer, act as platform to end user

The computer will need the same specification and also it can be the same machine for dedicated server. This computer will become the platform of the system for end user to use the system.

CHAPTER 4
RESULT AND DISCUSSION

4.1 Findings

Basically the author has come out with the ontology for apparel as basic ontology through exercising the task of designing the ontology. In this section author shows the result from implementing the task of designing the ontology for apparel and show its in tables. Besides that, the author explains the e-Commerce Ontology Web itself.

4.1.1 Tasks of Ontology Development

- **Build glossary of terms**

As shown in the table 4.1, these are the basic terms use in e-COW Apparel Ontology. For example we take Men's Apparel as an example. Through the table 4.1, we know the description of Men's Apparel which is all types of apparels for men. Moreover, we also know the type of 'men's apparel' is a concept or class where it becomes the one of main classes. 'Men's Apparel' becomes the super class for 'Men's Short Sleeves' because in the table above it shows that biscuit is a subclass.

Name	Description	Types
Men's Apparel	All types of apparels for men	Concept
Women's Apparel	All types of apparels for women	Concept
Apparel Price	A price for a particular apparel	Instance Attribute
Apparel Size	A size for a particular apparel	Instance Attributes
PADINI	A brand name of apparel	Sub - concept
Men's Short Sleeve	Types of apparel for men's apparel	Sub - Concept

Table 4.1: A selection of the Glossary of Terms of e-COW Apparel Ontology

- Build concept of taxonomies

Based on the glossary of terms, the author builds the concept of taxonomies to define the concept the hierarchy. Figure 4.1 illustrates on the concept taxonomies are build based on the e-COW Apparel Ontology, Basically the main concept or domain for this hierarchy is Apparel. Then it divides in several classes such as 'Men's Apparel'

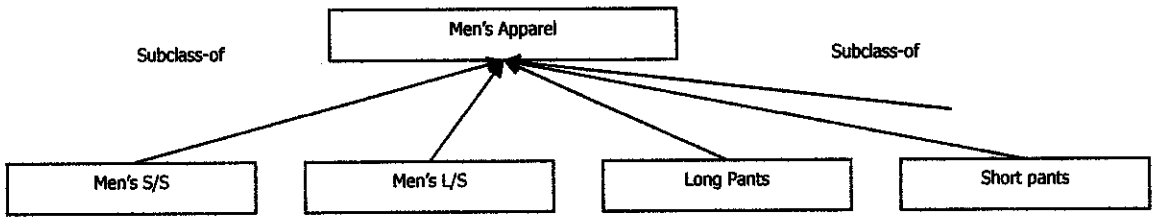


Figure 4.1: An excerpt the Concept of Taxonomy of Men's Apparel in Apparel Ontology

- Build concept of dictionaries

As defined in the chapter 3, the author has defined the properties and relations that describe each concept of the taxonomy in a concept dictionary. It contains the entire domain concept, their class attributes and their instance attributes. See table 4.2 for example of concept dictionaries. For 'PADINI', it has several instances attributes such as apparel price, apparel size, apparel color, apparel material, and apparel fabric.

Name	Class Attributes	Instance Attributes
Men's Apparel	--	---
PADINI		Apparel Price Apparel Size Apparel Color Apparel Material Apparel Fabric

Table 4.2: A selection of the Concept dictionaries of e-COW Apparel Ontology

- Describe instance attributes

Instance attributes are those attributes whose values maybe for each instances of concept. For each instance attributes, the author specifies the following fields: its name, its value type, its measurement, default value it exits, minimum and maximum, and cardinality. Table 4.3 shows a fragment of instance attribute table of Apparel ontology, some of previous field are not shown. This table contains the attributes price, fabric and size of apparel. This is how the instance attributes has been described.

Instance Attributes Name	Concept Name	Value Type	Measurement Unit	Default Value	Cardinality
Apparel Price	PADINI	Float	Currency Quantity	---	(0,1)
Apparel Fabric	PADINI	String	---	---	(0,1)
Apparel Size	Biscuit	String	-----	XL, L, M, S or 29 - 36	(0,1)

Table 4.3: A selection of Instance Attributes Table of e-COW Apparel Ontology

- Describe the class attributes in detail

Within the apparel ontology, the author does not define any class attributes. The reason because the class attributes is unlike instance attributes, which describe concept instances and take their values in instances, class attributes describe concepts and take their values in the class where they are defined. Class attributes are neither inherited by subclasses nor by the instances. For each class attribute, the

author fills the following information: name, the name of concept where the attribute is defined, value type, value, measurement unit and cardinality

- Describe instance

The author defines the instance based on the conceptual model of the apparel ontology. For each instance, the author defines: its name, the name of the concept it belong to, and its attributes values. Table 4.4 presents some off the instance table of e-COW apparel ontology. When the instance is '101101 apparel', so we know that the class for '101101 apparel' is PADINI and the price is RM 99.00, size is XL, L, M, and S and the last values is color for that instance is green.

Instance Name	Concept Name	Attribute	Values
101101	PADINI	Apparel Price	RM 99.00
		Apparel Size	XL, L, M, S
		Apparel Material	Cotton
		Apparel Fabric	Knit
		Apparel Color	Green

Table 4.4: A sample of instance in e-COW Food Ontology

4.1.3 Interface e-Commerce Ontology Web

For e-Commerce Ontology Web, the interface has been put at appendix section.

- Welcome Page

This page contains welcoming text to user and benefit of the system. There are several navigation buttons on this page which are Home, Search, Advertisement, About Us, and Contact Us. (See appendix Figure 4.2). Besides, it displays the Popular search for ease navigation to user

- Search Page

This page will display type of search available within the system (See Figure 4.3 in appendix)

- Category Page

Consists of category for apparel ontology only (see Figure 4.4 in appendix)

- Apparel Page

Consists of apparel ontology views in tree diagram (see Figure 4.5 in appendix)

- Result Page

Display the instance of class or result of the search (see Figure 4.6 in appendix)

4.1 Discussion

The response from various people was very interesting and encouraging about the product. Many wanted to know what ontology whole thing is all about. Explaining to the users took some time because they did not understand the concept of ontology. After demonstrating the system, they have a brief idea on how the system is all about. After they used the system, they gave their opinion and feedback on it. Most of them realized the benefits and negative aspects of the relying on the ontology and e-COW application.

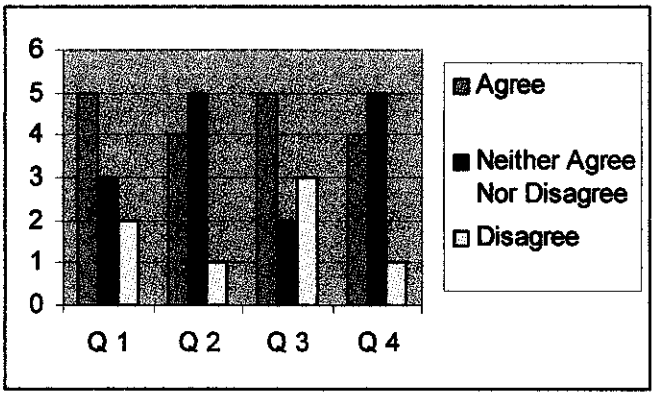


Figure 4.7: Respondent on the usage of e-Commerce Ontology Web

	Question	Rating		
1	The system is easy to use.	1	2	3
2	The system flow is easy to follow.	1	2	3
3	The content of the result / is understandable	1	2	3
4	The system is valuable and important to people.	1	2	3

Rating	**Q 1**	**Q 2**	**Q 3**	**Q 4**
Agree	5	4	5	4
Neither Agree Nor Disagree	3	5	2	5
Disagree	2	1	3	1

In this bar chart, we can see that many respondents agreed to the usage of the application but what probably caused the disagreement of other seven respondents. They would be thinking that this technology is not applicable in Malaysia.

CHAPTER 5
CONCLUSION AND RECOMMENDATION

5.1 Conclusion

Global economy is rapidly becoming more and more knowledge intensive. Knowledge is now widely recognized as the fourth production factor, on an equal footing with the traditional production factors of labor, capital and material. Semantic Web technology can make an important contribution to knowledge management.

The Semantic Web is a loosely-coupled set of technologies designed to create a semantically based network of documents across the World-Wide Web. Semantic Web languages, such as RDF, DAML, and OWL, are built on top of XML, but in addition have the ability to define object-oriented relationships. Document schemas can extend to another document schema, and typed relationships between documents can be created. Semantic Web technologies support the concept of the expression of meaning and of knowledge representation in web documents by using available and previously defined ontologies. Ontologies are central to the Semantic Web, where they describe an explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in specific domain of interest and the relationships that hold among them. With ontology development process, the process of designing the ontology becomes easier to complete. In addition, the tasks of designing ontology have become guidelines for ontologist to create ontology.

For e-commerce communities, the Semantic Web can be used as a model for specifying a highly expressive and open language that can be used not only in advertisements, but also covering all relevant business activities (manufacturing, transportation, purchasing, etc.), of supply chain model - all in a uniform way. The Semantic Web is intended to facilitate a full automation of Web service tasks including automated Web service discovery, execution, interoperation, composition and execution monitoring. But, this kind of evolutional automation is has yet to be fully outlined in any form of usable detail.

For this project, the successful of delivering the application can make the system become more useful in future time. Information system and knowledge management is very essential in our daily life especially for business organization. So, author hopes that the system will meet the requirement of today technologies and also future technologies.

5.2 Recommendation of Future Works

Up to the current stage of project development, the suggested next task is to implement into real supermarket environment. Additional function also can be developing such as online purchasing where the users can purchase the product through the system. The system also can be upgraded by adding the features for system administrator to maintain the ontology agent using the system itself. The capacity of ontology can increase by upgrading the concept and class of the ontology to become more details and specific.

5.2.1 Apply in real organization.

Currently, the prototype is mainly for Final Year Project (FYP) only but not applicable to real environment. So, the author suggests the system can be implemented in real supermarket environment where the user can go to one stop corner for searching the products they want.

5.2.2 Purchasing feature

The prototype is only applicable for searching out the product but not for purchasing the products. In the future time, author suggests for adding feature of online purchasing for the products available in the particular supermarket.

5.2.3 Features for System Administration

The prototype does not contain features for system administration to upgrading the ontology agent within the prototype itself. System administrator needs to enter the e-COW editor by bringing back the system into Coding and Testing phase for customization. The adding of system administrator to updating ontology agent will make the system more dynamics in future.

5.2.4 Capacity of Ontology

The time constraint is very crucial for this project. Most of the time has been spent out only for designing the ontology for this system. it has effect the developing ontology where the author has cut some of the products to be viewed in the ontology. If the ontology can be upgraded and covering almost all product, therefore, it become dynamic and efficient in the real implementation

REFERENCES:

Liang T.P., Aronson J.E., Turban E., (2005). Decision Support System and Intelligent
Sytem. 7th Edition. United States: Pearson Prentice Hall.

Gomez-Perez A., Fernandez-Lopez M., Corcho O. (2004). Ontological Engineering: *with examples from the areas of knowledge management, e-commerce, and the*
semantic web. London: Springer.

Fensel D., Hendler J., Lieberman H., Washlster W., (2003). Spinning the Semantic Web Bringing the World Wide Web to Is Full Pontential. London: The MIT Press.

Davies J., Fensel D., Harmelen F.V. (2003) Ontology-Driven Knowledge Management: *Towards The Semantic Web.* West Sussex: John Wiley & Sons Inc.

Geroimenko V., Chen C., (2003). Visualizing the Semantic Web: *XML-based internet*
and information visualization. London: Springer.

Berners-Lee, T., et al.(2001). "The Semantic Web." *Scientific American*, Vol. 284.

http://cse.unl.edu/~smodali/research.htm

APPENDIX

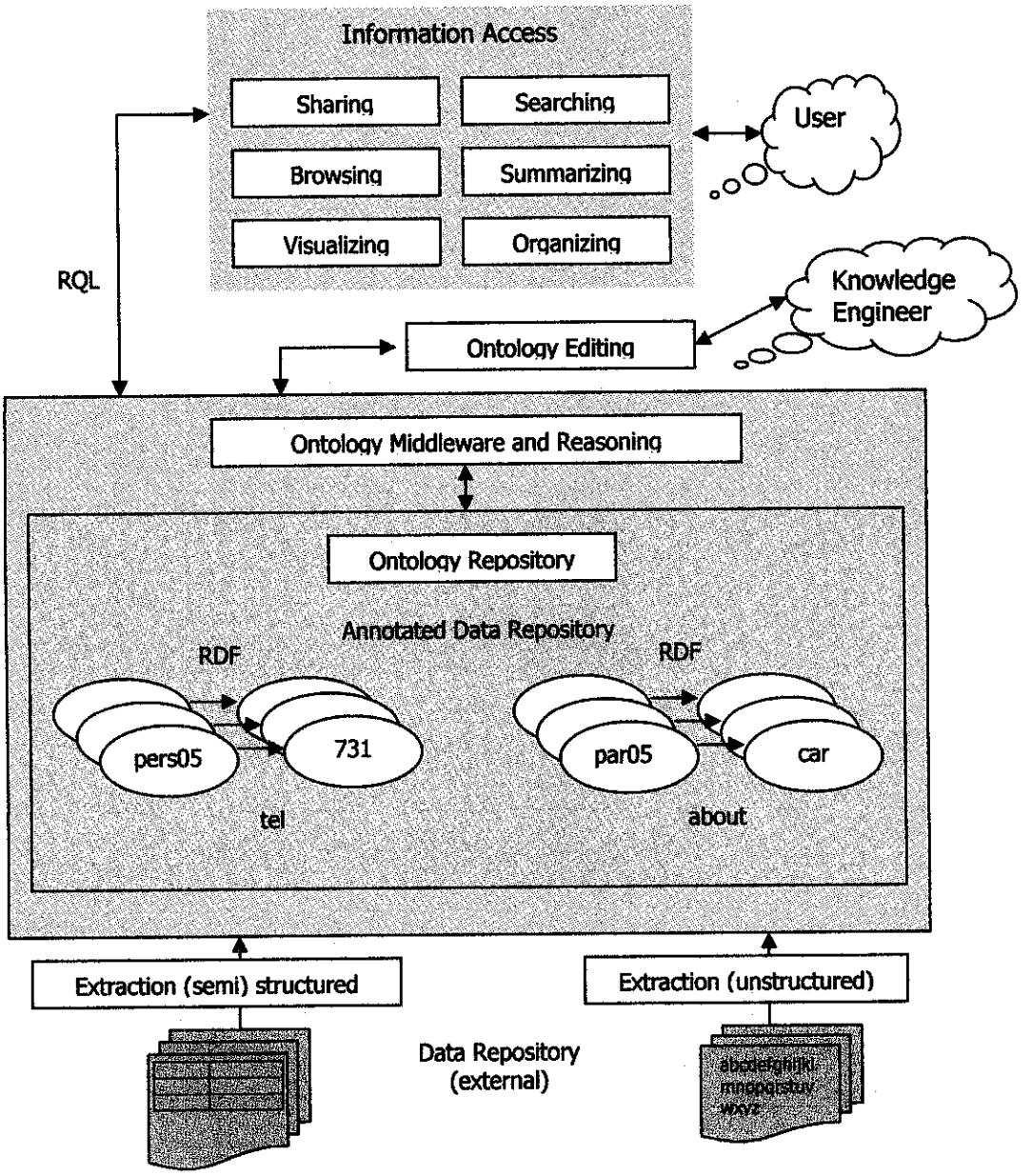


Figure 2.2: Semantic Web-based knowledge management Architecture

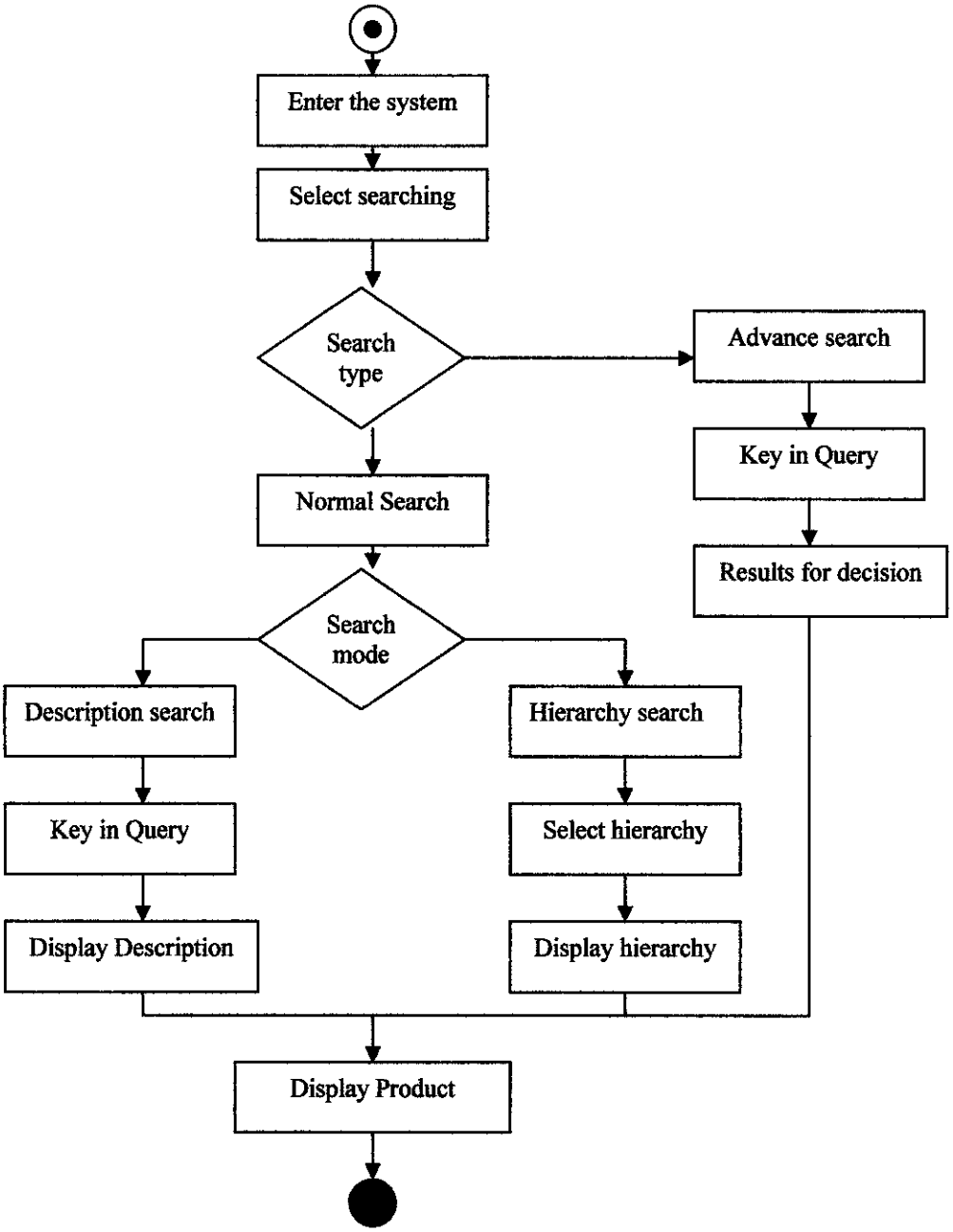


Figure 3.5: e-Commerce Ontology Web process workflow

e-Commerce Ontology Web (e-COW)

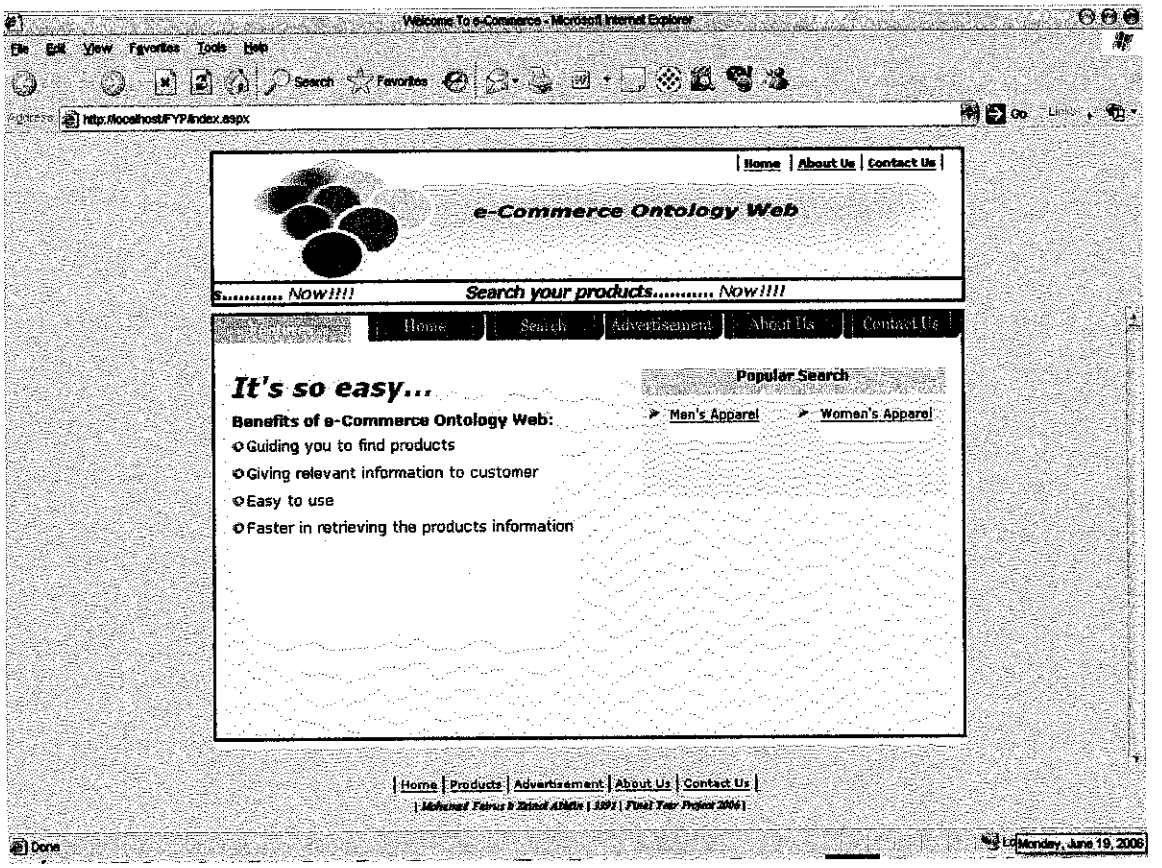


Figure 4.2: Welcome Page

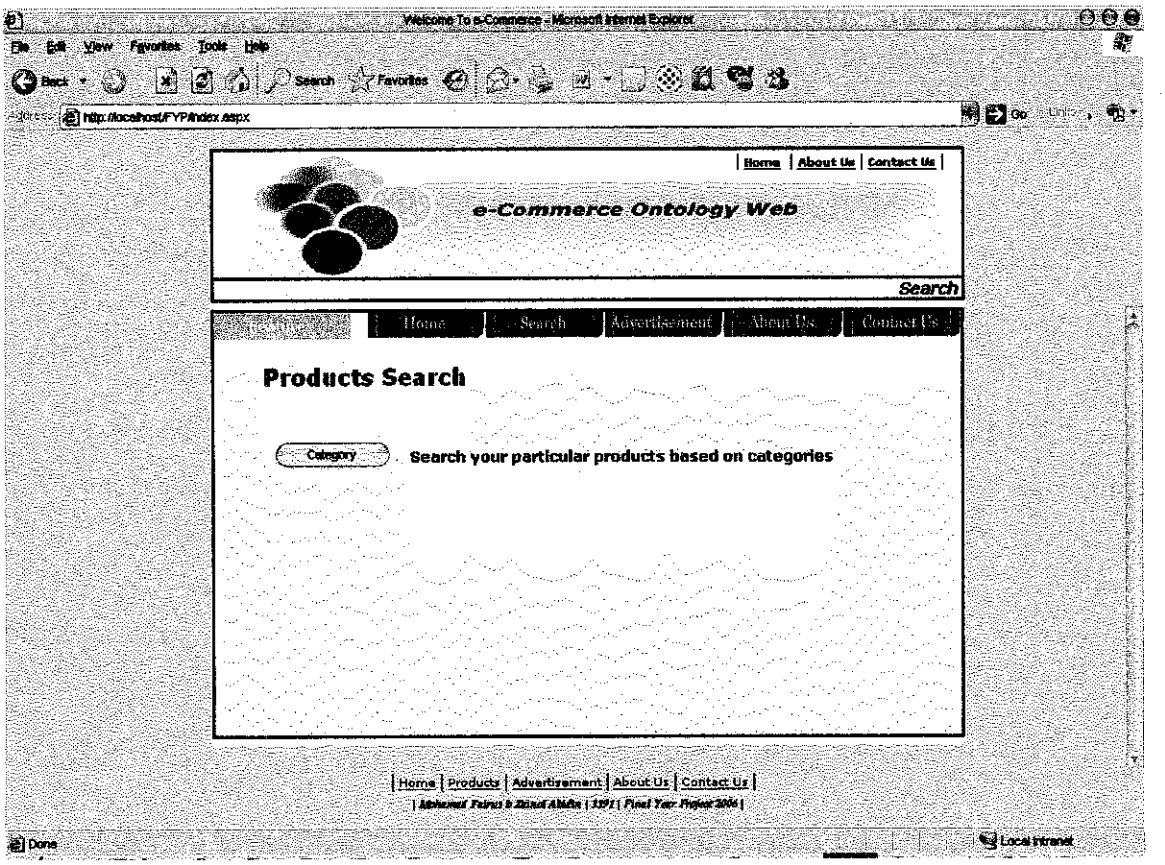


Figure 4.3: Product Search Page

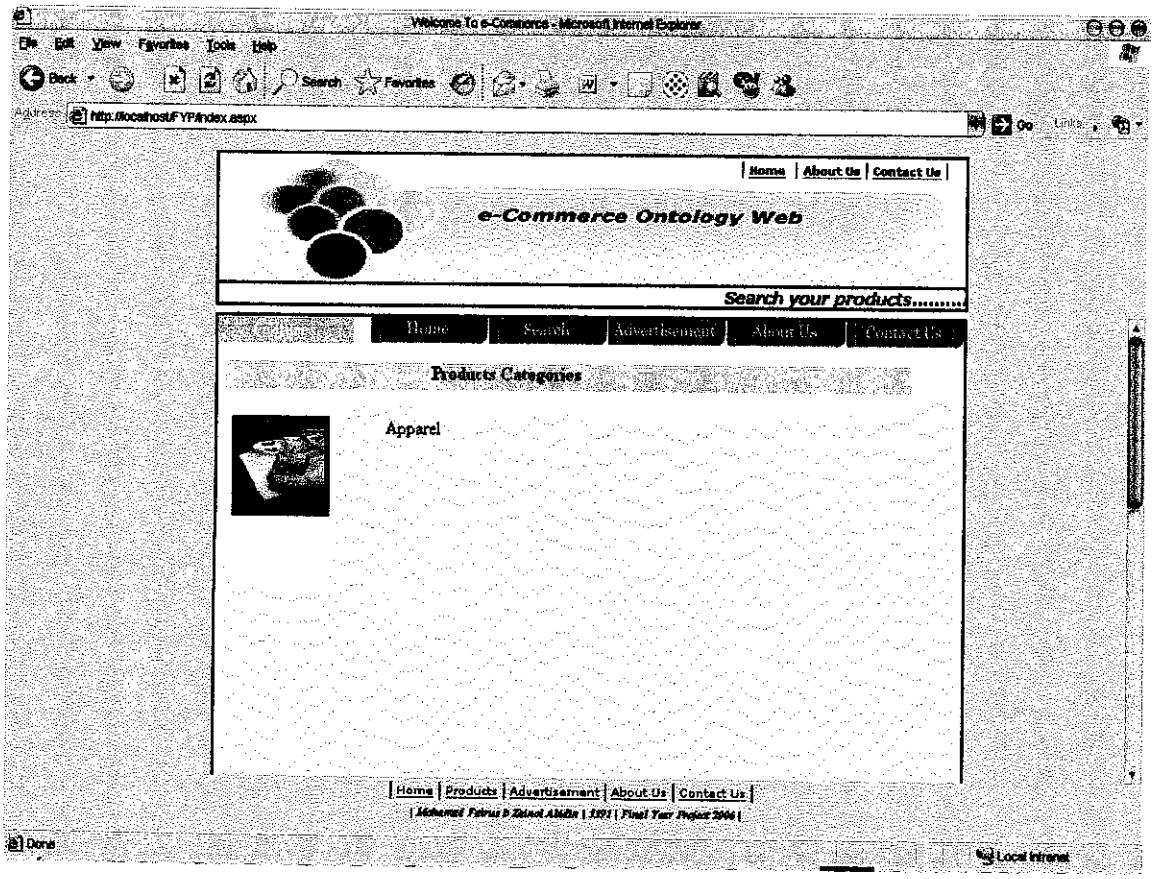


Figure 4.4: Category Search Page

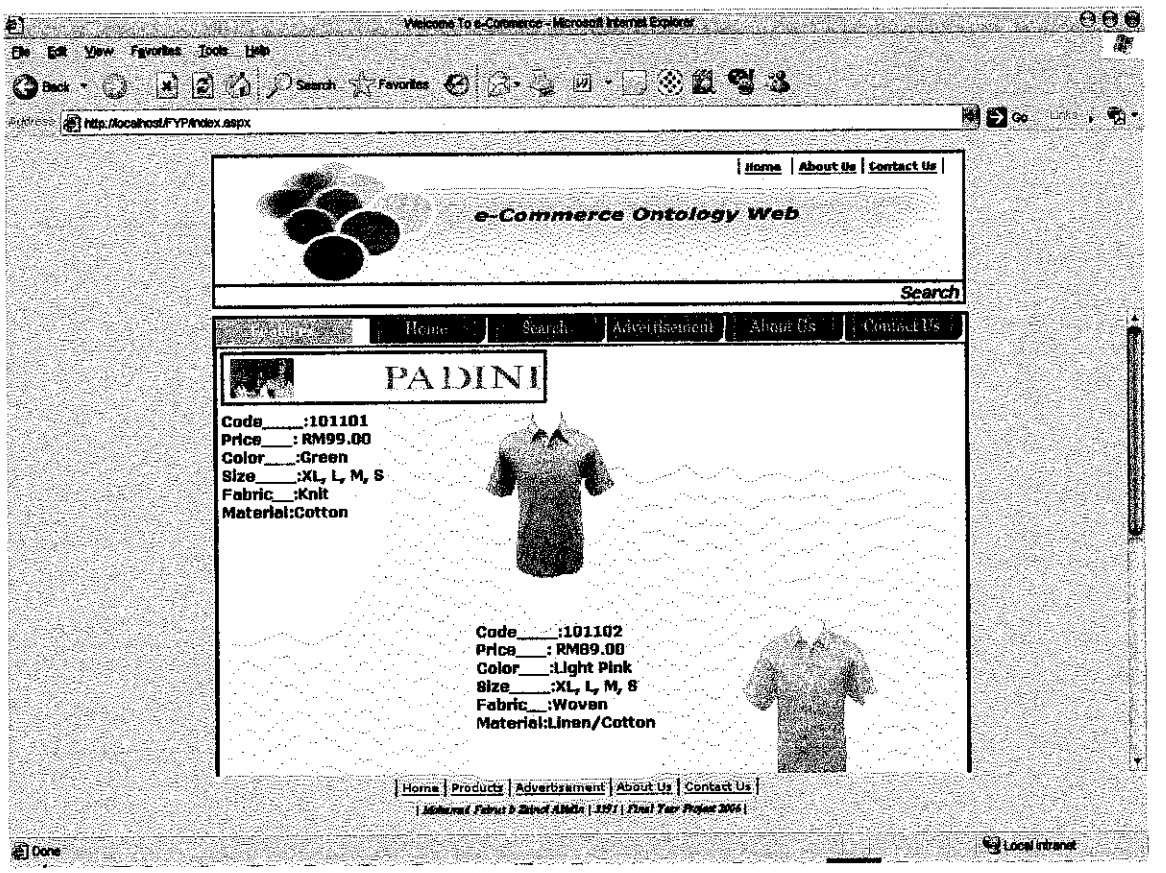


Figure 4.6: Result Page