Universidade do Minho

School of Engineer

Department of Information System

Quang Dang Nguyen

The role of Business Intelligence in Organizational Memory support

Dissertation Report

Master in Information System

Work performed under the guidance of

Professor Isabel Ramos

June, 2012

"Learn from yesterday, live for today, hope for tomorrow.
The important thing is not to stop questioning."

– Albert Einstein –

Acknowledgments

Firstly, I really want to give special thanks to my helpful supervisor, Professor Isabel Ramos. The supervision and support that she gave truly help the progression and smoothness of my dissertation from the beginning time. I greatly appreciate your kindness to lead me overcome my mistakes in this project despite of your busy time in researching and teaching.

My grateful thanks also go to Professor Filipe Sá Soares, Professor Ana Alice Baptista, Professor João Álvaro and Professor Luís Amaral. A big contribution and hard support from all of you during the first three semesters is very great indeed. All projects during the program would be nothing without the enthusiasm from all of you. Besides, this master program makes me realize the value of scientific research as a new experience in the academic environment, which will help me self-confident in my future study.

Great deals appreciated go to the contribution of my faculty - Faculty of Information Systems (MSI). I am also would like to thankful the International Relations Office (GRI), especially Ms. Adriana Carvalho - Head of GRI, Ms. Carina Oliveira, Ms. Maria João Teixeira and all the staff in the Department of Information Systems (DSI) office that patient in helping us complete this program.

Last but not least, I would like to thank my colleagues especially those who work together at University of Minho – Ms. Alexandra Diana and Mr. José Miguel Guedes Nunes for the wise supports throughout the project.

Finally, I would also thank the DSI, as well as all my faculty members without whom this project would have been a distant reality. I also extend my heartfelt thanks to my family and my best friends in mobilizing me for almost two years at the University of Minho.

Abstract

Nowadays, in all organizations the major challenge issue facing managers is that they must give the appropriate decisions in a fluctuating environment while the information seems very hard to recognize whether it is good or bad. However, the actions that result of the decisions made will lead the organization to be in a thriving or declining position. That is why the leaders of organization really do not want to take wrong decisions. In order to minimize the risks, the managers should use the collective knowledge and experiences sharing through the Organizational Memory effectively to reduce the rate of unsuccessful decision making. Moreover, the BI systems are also a managerial concept and tools to allow their business to improve the effectiveness of decision making and problem solving. In the light of these motivations, the aim of this dissertation is to comprehend the role of the BI systems in supporting the system of Organizational Memories more effectively in the real context of crowdsourcing initiative called CrowdUM.

Keywords: *Business Intelligence, Organizational Memory, Crowdsourcing.*

Contents

List of Figures

List of Tables

List of Abbreviations

Throughout the dissertation there are many acronyms appear with a high frequency. In addition, these acronyms are used commonly in the community of science expertise to replace the repeat of long phrases. Hence, the following list will summarize acronyms in the alphabetical order to help us make sense the abbreviations:

BI Business Intelligence

DM Data Mining

DSRM Design Science Research Methodology

DW Data Warehouse

KM Knowledge Management

OK Organizational Knowledge

OLAP On-Line Analytical Processing

OM Organizational Memory

OMIS Organizational Memory Information System

OR Organizational Representation

Chapter 1

Introduction

Nowadays, in all organizations the major challenge issue facing managers is that they must give the appropriate decisions in a fluctuating environment while the information seems very hard to recognize whether it is good or bad. However, the actions that result of the decisions made will lead the organization to be in a thriving or declining position. That is why the leaders of organization really do not want to take wrong decisions. In order to minimize the risks, the managers should use the collective knowledge and experiences sharing through the Organizational Memory effectively to reduce the rate of unsuccessful decisions and actions.

In a basic view, with the objective to enhance the decision making, one solution is proposed to manage the past knowledge shared by organizational members, as well as using information that is stored in computer systems throughout a long-time operation to increase the effect of decision making. Hereby, it helps organizations have a better and wider understanding of the objectives and past experiments throughout the aggregating individual knowledge across the years. In order to increase the effect in managing and reusing past knowledge in an organization, we need to keep track of knowledge and information in organizations via the processes such as capture, creation, storage, and retrieval information in shared spaces to help individuals might easily retrieve useful

information so that support them to solve their own problems. This concept in general is defined as the Organizational Memory (OM).

In another view, the Business Intelligence is a managerial concept and tools that allow us create and maintain a large database, retrieve of stored information, and use stored information to make effective decisions [Pirttimaki et al., 2005]. The BI then indicates a significant role to support OM in automatic collecting and updating the knowledge, proposing the useful information for decision making, and giving a better environment for organizational members to interact with the OM. Based on these advantages of BI in effectively managing information and the limited of OM in practice for managing and retrieval the organizational knowledge, this dissertation hence focuses on the role of BI systems in the light of this concept for more effectively support the Organizational Memories Information System (OMIS). As a result, this project then will propose an integrated model between the OMIS framework and the suitable BI tools to improve performance of the OMIS. Moreover, the proposed model is going to experience and validate in the real context of CrowdUM environment.

Following the research problem, the structure of this dissertation then will be divided into six chapters. The first chapter introduces the overall view of the dissertation problem. The second chapter focuses on the reason of the proposed work that includes the motivation and brief descriptions of BI and OM concepts. The third chapter is going to define the research question and investigative questions. Consequently, the next chapter is going to expand the literature research about the OM and BI concepts. Afterward, I will summarize the current researches in order to make clear the roles of BI system in supporting the OMIS. The fifth chapter focuses on introducing about the real context for experiencing the new proposed model and further clarify the detail features that will be used for the validation. Next, the sixth chapter intends to accomplish the implementation of proposed model in the CrowdUM context. Finally, the last chapter is going to explain the results that were archived from this research; indicate the recommendations; comprehend the limitations; and propose the future researches.

Chapter 2

Framework of Proposed Work

2.1 Motivation

Following the modern development progress in many organizations nowadays, the production specialization has become a new trend. Consequently, the structure of the organization is split into many divisions, in which each division will pay attention to different tasks. Thereby, the job sharing phenomenon - separating the main job into several groups - created a problem in sharing knowledge, because individuals in separate groups are neither aware other jobs nor have an understanding about new information in another department, instead they only use their own knowledge to solve their tasks. Moreover, the daily works of each individual usually relate to the decision-making that requires handling wide information and experiments overall the organization. This problem raises a demand for individuals within the organization to search and use relevant information in the past to support them solves their tasks more quickly and more accurately. By using the shared knowledge to improve performance of individuals, it enables the organization could operate with higher precision; consequently, these shared knowledge will become a primary precondition for the modern enterprise to grow up.

As we have known, the OM is a conceptual tool to support the process of knowl-

edge management by collecting knowledge and experience generated inside the organizational environment. The OM concept thus has been first presented in an integrated form by Walsh & Ungson [1991], with the OM concepts they made a first step for the OM fundamental framework to motivate scientists in business management and information system to go further. Since the 1990s, many researchers have dealt with the OM theory with the purpose of developing an effective OM application, for examples, at that time we had some representative studies as Stein et al. [1994], Schwabe [1994], Ackerman & Mandel [1995], Rao & Goldman-Segall [1995], etc. However, despite all the efforts of many researchers, the outcomes are still stopped at some limited successes of the OM concept that given the lack of methodological and technological tools to be used by practitioners. Indeed, S. Hamidi & Abdul Rahman [2006] in a case study indicated that "the implementation of OMIS is very challenging since it involved dealing with soft and hard issues such as technology and human factors". In the same aspect, Wijnhoven & Slooten [2002a] listed three problems within the OMIS development and concluded that "development activities require conceptual cohesion as far as organizationally required and technologically feasible".

Based on the limitation of OM in realization aspect, it motivated this research to find a solution for OM system to improve its performance throughout three main functions such as acquisition, retention, and retrieval that could support effective decision-making and problem-solving in a practical environment. Moreover, the advantage architecture of the Business Intelligence system is perceived as the appropriate managerial concepts and tool to enhance the performance of the decision-making. In addition, the BI comprises many useful tools to support OM in managing a large amount of information (e.g., DW, OLAP, DM). Thus, this study will focus on analyzing the limited aspects of OM that could be intervened and enhanced via the strengths of BI system.

2.2 Organizational Memory and Business Intelligence

As a concept in the knowledge management, the OM is a knowledge repository for managing a large amount of undocumented experiences, skills, and knowledge that acquired from the past. The OM pays attention to capture the organizational knowledge, which is derived from daily works of problem-solving and decision-making, then these collective knowledge will be stored in OM for later reuse. For that reason, the OM helps organizations keep track knowledge from individuals and groups, which could be transferred to newcomers later. The OM thus could not only reduce the training duration for newcomers, but also supports organization to avoid the memory loss due to depression over a long time and to preserve knowledge from the leaving of key organizational members. As a result, the knowledge retrieval could support participants to solve their own problems in an effective manner.

By the same purpose, the BI is also a useful technology to support decision making and problem solving. With more and more organizations nowadays shift to use BI in their business as a trend of the modern business to improve their success rate in decision making and lower their long-term expenditures. For this purpose, the BI concept is defined as a combines architectures, tools, databases, analytical tools, applications, and methodologies [Turban et al., 2008]. In fact, the BI mainly focuses on computer-based techniques to support the business in storage, extraction, and analysis business data; then it provides a better solution for business decision making. In summary, the BI shows an active relationship with the OM in helping organizations to manage effectively a large amount of information stored across time, as well as, support OM to search, retrieve, and deliver the past knowledge more exactly and timely.

2.3 Research question

Based on the definitive problem, this research intends to improve the performance and effectiveness of the OM in practice throughout achieving the primary goal and proposing recommendations for using the BI system to improve the OM performance. In order to address the definitive problem and then lead this research into the final goal, I suggest the main research question as: 'How can the BI systems support the OM to improve the performance of decisions and actions in organizations?'. Because this is such a complex question, hence I divided this research question into the list of investigative questions that help me investigate and make clear the important aspects of the BI in OMIS support. Moreover, the 'factors used to determine the usefulness of the OM' is key assessments, which are identified for verifying the effectiveness of the BI-OMIS integrated model in a real context. Finally, the figure 2.1 summarized the high level working map of the research question.

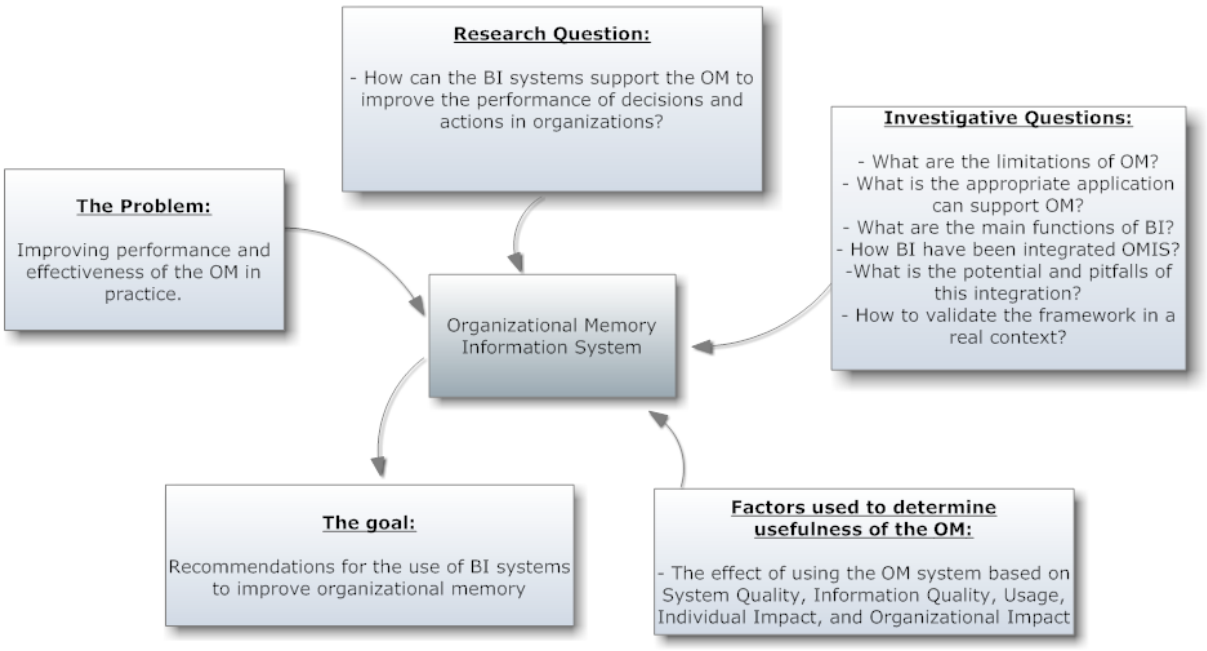


Figure 2.1: The high level working map to solve the main problem

Chapter 3

Objectives and Expected Results

To address the research question, which is depicted in the previous chapter, I divided the main research question to six investigative questions as presented in the table 3.1. These questions will support me to make sense the main problems throughout different aspects, in turn, the results will lead my research to the final goal. For this purpose, the first question intends to find out the limitations of the OM in practice that need to be enhanced. The second question concerns in understanding the BI functions and tools for further using. The third question is going to define such appropriate applications and tools of BI system for improving the OM performance. Dealing with the integrated model between the BI and the OMIS, the next question is going to identify methods to adapt BI tools into the OMIS as well as to comprehend the potential and pitfall of this integration as presented in the fifth question. The final investigative question plays a significant role in validating the proposed model in a real context of the crowdsourcing initiative called CrowdUM.

Based on six investigative questions, the main objectives that need to be solved in this dissertation will be presented in five stages as follows:

- Stage 1: Understanding the functions of OM and OMIS in the light of basic concepts and mnemonic functions. The result of this stage is the summary of recent studies

Table 3.1: The list of investigative questions

No	Investigative Questions
1.	What are the limitations of the OM?
2.	What are the main functions and tools of BI?
3.	What is the appropriate application that can support OM?
4.	How BI has been integrated to the OMIS?
5.	What are the potentials and pitfalls of this integration?
6.	How to validate the framework in a real context?

on the matrix of concepts. Furthermore, the summary also points out the challenges of OM that need to be improved in practice.

- Stage 2: Reviewing past researches on the intersection of BI and OM then organize collected information in a well structured form to support the research question. Moreover, this research focuses on finding studies and tools where the BI systems could explicitly link to the performance of OM. The main result here are the definition of appropriate BI tools that could be integrated to enhance the OMIS, and the definition of performance measurements to motivate the follow stages such as development and evaluation.

- Stage 3: This stage intends to propose an integrated model to understand how BI tools can be used to improve the OM. The result of this stage is to define the BI-OMIS integrated model, which will be implemented in the next validation step.

- Stage 4: Validating the framework in a real context; this stage is going to use the crowdsourcing initiative, called CrowdUM, as a framework for an experience of using BI systems to support the OM. In this stage, the result focuses on developing a real application of the BI-OMIS integrated model in the real context of CrowdUM.

- Stage 5: Define recommendations rely on the validating results then incorporate learned lessons into the proposed framework. The result is the recommendations

for the use of BI systems to improve the OM.

Finally, the expected result that based on the evaluating of BI-OMIS integrative model will be used to define recommendations for supporting the OMIS. Particularly, these recommendations are going to help this research analyze and improve the OMIS of CrowdUM in practical aspect.

Chapter 4

Literature Review

4.1 Cognition about Organizational Memory

4.1.1 Concepts of information, knowledge and knowledge representation

Since the appearance of OM concept, there are many researchers have dealt with the question "what resources will be dealt with the OM?" in order to make clear the OM capability such as Basaruddin & Haron [2011], Walsh & Ungson [1991], Dieng et al. [1998], and R. T. Watson [1998]. To comprehend the OM, this section will point out three main resources of OM, which are often used in the organizational environment as (i) Information, (ii) Knowledge, and (iii) Knowledge Representation. In general, the information, knowledge, and knowledge representation sometime are understood and used with a little bit dissimilar between researchers [Stenmark, 2002]. The problems here will be analyzed by exploring the definition of these concepts. In fact, the information comes from the basic factors called data or facts in the real world. Therefore, they have a close correlation such as the data may be transformed into information while it is used in specific context, and conversely. For examples, while the manuals, books, rules, and organizational procedures

are used in their context, they are called information; and conversely, when the information losses its context, it is only understood as data. At a higher level, data or information by itself is a useful material to build knowledge. Information will become knowledge when it is aggregate analyzed within the individual mind to decide what information need to use and which useful context for this information in use, e.g. the experience is a knowledge. Because the knowledge seemly exists in an informal form, it leads us to a general question about "how to represent knowledge?". As a result, the notion of knowledge representation has been offered to help us capture the knowledge and save it in a tangible form as semantic network or database system. Normally, the boundary between information and knowledge is still a gray area and they are very hard to recognize in specific contexts. Hence, I will attempt to clarify these definitions throughout the past researches as follows:

Information

Information is the higher form of the raw data and exists in a static form, which contains the meaning of a particular problem or situation; however the information, by itself, is not useful until it is used by person's awareness to deal with their problem [Amidon, 1991]. Hereby, Stenmark [2002] indicated that "what can be articulated and made tangible outside the human mind is merely information", this shows the main attribute of information is tangible and it exists outside brain of human, by this meaning he inferred that if the knowledge can be "articulated and furnished with words" then we can call it as information. In the same idea, Wiig [1994] also shows that, information is the "facts organized to describe a situation or condition"; in addition, Spek & Spijkervet [1997] intended to the meaning and defined information as the "data with meaning" or "a flow of messages or meanings which might add to, restructure, or change knowledge" [Machlup, 1983].

Conclusively, based on Stenmark [2002], this research will follow the definition of information as "the focal knowledge that can be articulated and furnished with words".

This definition shows that information is a part of knowledge, which is created by a human brain when combining words in a meaningful way. The Information usually answers the question of "What", for instance: "What is my social number?", and then the answers to this question is a kind of information (e.g., the personal information, the bank account). As a result, the information exists in a tangible form and it allows us to recall for using whenever we need.

Knowledge

Knowledge exists in the mind of human beings, for example, the knowledge appears when we are using our mind to deal with particular problems such as selecting appropriate information in the right context and meaning. The knowledge hereby is synthesized from a chain of past experiments to now. By this mean, the knowledge consists of facts, truths, and beliefs so it cannot be transferred in an exact meaning to another mind. The only way to transfer knowledge is through the information, then information by its meaning is represented in tangible forms (e.g., books, articles, papers) before re-create knowledge in another mind. The exact transfer rate of original semantic belongs to the past knowledge of the reader. So the term of knowledge usually has a wider meaning than the information; in a metaphor of the iceberg, while the information represents for the tip of the iceberg, the bulk of the iceberg - occupies a large part of the iceberg - represents for the knowledge. Hence, Stenmark [2002] implies that all knowledge is tacit and it represents for the sinking part (tacit) of the personal experiences while the information is the visible part (articulated) of it. The boundary of them is very loose, for example: if something can be processed or thought in the human brain and exist in a tacit form we call it as knowledge. On the contrary, when the knowledge is articulated or represented in an explicit form as words we can call it as information. Thus, this research will follow the definition of "knowledge is based on personal experiences and cultural inheritance and fundamentally tacit" and we could "use knowledge to perform actions such as creating information" [Stenmark, 2002].

At the point of classifying information and knowledge, Nonaka & Takeuchi [1995] informed that information is a flow of messages while the knowledge is created and organized by the very flow of information. By this notion, they want to propose the important role of human action in knowledge when comparing with the information; furthermore, they also insisted that knowledge is derived from information. Wiig [1994] focused on the philosophy and indicated that while information is the "facts", knowledge is the "truths and beliefs, perspectives and concepts, judgments and expectations, methodologies and know how", this indicates an important role of the knowledge in handling meaning than the information. In an illustration to make clear the process of knowledge transfer between the author and readers, I assume knowledge such as the author's experiments or innovations in their research. In thus, the readers - with their own experiment - can only retrieve the information from a scientific article (as an information or explicit knowledge) and then re-create their own knowledge, nevertheless, the original knowledge only exists in the author's mind (a kind of tacit knowledge), who produced the scientific article.

Knowledge Representation

Based on the difficulty of retention and retrieval of knowledge, there are many studies to understand how knowledge can be represented and what kinds of reasoning can be done with that knowledge. Knowledge representation (KR) is a method to support KM in making knowledge explicit and reduce the ambiguity of knowledge by its explicit form. By this objective, KR defined a set of structures to contain, managing the knowledge for easy represent it in knowledge reuse.

Sowa [1999] defined KR as "the application of logic and ontology to the task of constructing computable models for most domains". With the objective to make the knowledge explicit, the definition presses a point to KR as a multidisciplinary subject, which applies theories and techniques from logic, ontology, and computation. In a review to make clear the meaning of KR, Davis et al. [1993] summarized the KR in terms of five distinct roles. The first role of KR is a surrogate; it plays a role as a substitute for natu-

ral objects in the real world that exist in both intangible and tangible forms. The second is a set of ontological commitment, which limits the KR environment to just deal with its own problem and ignore the other irrelevant rules. The third role of KR focuses on fragmentary theory of intelligent reasoning, this role defines a set of general inferences to support KR in analyzing and sorting out the issues appropriately such as (1) fundamental conception, (2) inferences that the representation sanctions, and (3) inferences that it recommends. The next role is a medium for pragmatically efficient computation to assist KR in making recommended inferences. Moreover, the final role is a medium of human expression, which focuses on the role of the transfer of messages to help us communicate with machines and other people.

4.1.2 Types of organizational knowledge

Organizational Knowledge (OK) is a branch of Knowledge Management (KM), OK has been defined by Tsoukas & Vladimirou [2001] that OK is the "capability members of an organization have developed to draw distinctions in the process of carrying out their work, in particular concrete contexts". This emphasizes the role of members' knowledge in organizations to build up a historical memory system for supporting organizational actions. In addition, this definition indicated that the OK has a specific feature that helps us create a distinct organizational cultural with other organizations. By the notion of organizational knowledge, the OK usually exists in three types:

Tacit Knowledge This knowledge implicit an abstract entity, which exists in each individual mind. The tacit knowledge is created in our mind whenever we perform our work. Because it is an intangible form, so we cannot codify it outside our mind. Moreover, tacit knowledge is hard to communicate or duplicate to other members because it is presented by action-based skills in the owner's brain. Hence, we cannot summarize them to rules or record them in normal methods. For example: the physical feeling of a driver when he drives in a crowded street that is hard to explain

and modify this knowledge at all.

Explicit Knowledge In the opposite direction with tacit knowledge, the explicit knowledge appears in tangible form, which can be codified formally using a system of writing or symbols. By this property, we can easily communicate or broadcast this knowledge to other members. Explicit Knowledge also exists in kinds of object-based or rule-based, for example: the recipe or driver guides.

Cultural Knowledge The cultural knowledge is a set of knowledge about norms, practices, and belief that was created via the communication and sharing information between members in the organization. Pemberton & Stonehouse [2000] emphasized the cultural knowledge capability in creating an atmosphere of trust that helps individuals to experiment with new opinions in business. The cultural knowledge is created spontaneously from the earliest members in the organization. Consequently, it is transferred to the newcomers, when they join in the organization. However, in some organizations the cultural knowledge is created by special purpose of organization, and its result is to build beliefs and assumptions concerning organization's goals. In effect, the cultural knowledge will help members adjust their tactics to achieve their tasks based on the organizational goals.

4.1.3 Concept of Organizational Memory (OM)

As we have known, while the OK is the capability of members in organization to help them solve works; the OM is a conceptual tool that focuses on the retention capability of organization for storing collective knowledge and experience of the organization. OM has been first presented in an integrated framework by Walsh & Ungson [1991]. According to Walsh and Ungson, the OM "refers to stored information from an organization's history that can be brought to bear on present decisions". In the high-level, Ackerman [1994] assumed "OM is an evocative metaphor, suggesting the promise of infinitely retrievable knowledge and experience" or in the lower lever Schwartz [1998] divided OM into two

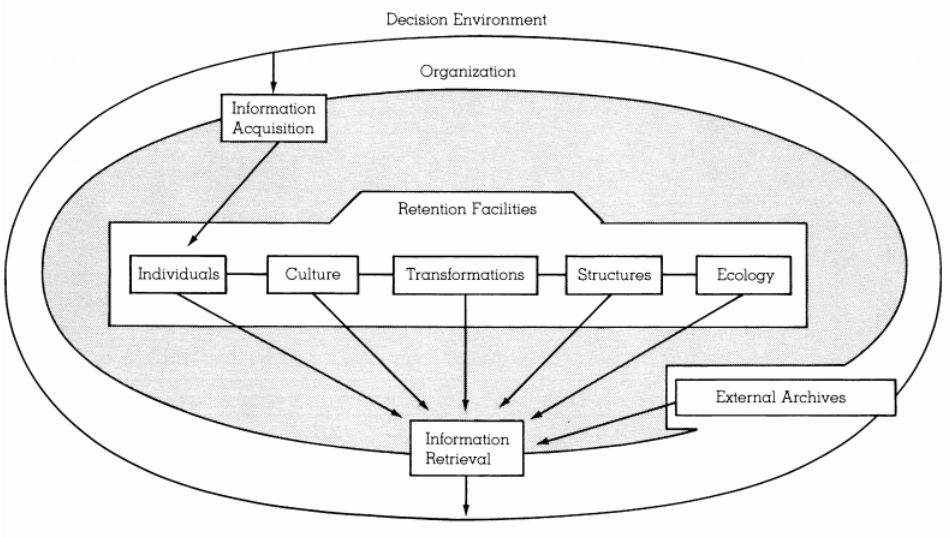


Figure 4.1: The structure of OM

key components: "(1) a knowledge base which contains the content or knowledge that is of value to the organization; and (2) a well-defined set of meta-knowledge, which is used to determine how and when the knowledge or content should be applied".

In the other words, these authors defined the OM term by comprising two main factors: the individuals acquire information from their working to solve problems, and the organizational-level construct to interpret sharing knowledge and transcends the individual level. In their model the storage components play an important role in OM that focus on answering two questions "what do we need to remember?" and "how is it remembered?". By this view, Walsh & Ungson [1991] proposed the structure of OM as represented in the figure 4.1. This structure focuses on three main functions to manage the information in an organization such as: Information Acquisition, Retention Facilities, and Information Retrieval as explained further below:

Acquisition this function is a process of capturing information expressed when organizational members solve the organizational problems. Moreover, this information is usually acquired not only from individual knowledge but also from the organizational decisions response to the stimulus [Walsh & Ungson, 1991, p. 62]. In general,

the acquisition will capture the information whenever it perceives the decision stimulus and response, which may deliver by individuals or organization via business operations. Thereby information is captured, and then it is stored in the organizational memory for future reusing.

Retention this function is used to store the information acquired from many sources of the organization. Although there are usually exist various arguments about the sources of information in OM. Walsh and Ungson in a summarized view, has concluded them to five storage components such as: individual (the memories of employees to solve the problems occurring in their job), culture (the learned way of employees to solve problems through the transmit of organization culture), transformations (the routines or procedures that guides the transformation of an input into an output), structure (the roles or rules of individuals in organization), and ecology (the physical structure and arrangement of an organization).

Retrieval the retrieval function is used to help organization access the past information (experiences and decisions) contained in OM to support the organization in solving new problems. Because the information is captured in specific context then it requires the information retrieval must be embedded in the past context to help organization restore knowledge correctly. With a large amount of information kept in five retention facilities and external archives, it requires individuals understanding the historical relevance information to retrieve information purposefully. Nevertheless, the searching and matching in a large amount of past information is such a hard problem to recognize the right solution for individual tasks in some special cases (e.g., doubt problems, controversial issues). Certainly, this is still a challenge for an information retrieval function so far.

Rely on the functions of OM structure to manage the past information; the realization of an OM is still a hard problem because it requires organizations face with issues related to human activity inside the organizations, for examples: capturing individual

knowledge or managing large unstructured information exists in the organization. To support the organization in managing the OM effectively, the functions of OM need to be improved through three main features such as (a) acquiring knowledge arose in formal and informal knowledge [Atwood, 2002]; (b) efficient managing and storage of large data with context and semantic aspect [Goesmann, 2001, Miller et al., 2008, Shepherd & Martz, 2006]; (c) providing an easy and accurate tool to help individual retrieve historical information more effectively [Ackerman, 1994, Ackerman & Halverson, 2004, Ochoa et al., 2009a, Yang & Chen, 2006].

4.1.4 Challenges in managing OM

More than two decades of developing the notion of OM, there are many researches paid attention to identify and improve not only the theory framework but also the practice uses with the objective using OM to support the working process inside the organization. However, the field of OM is still a rudimentary area that facing many challenges in practice. In this section, I will make a synthesis of these current researches in the OM field. The result has been analyzed through articles relate to OM keyword and then classified in the matrix of concepts, which covered three kinds of main researches in OM as the Managing Knowledge, the Managing People, and the Practices Use to build a system [Atwood, 2002]. The idea is to create a taxonomy of challenges in managing OM in the current environment of the organization.

To summarize the matrix of concepts by classifying each article through a set of significant concepts, the result will give us an overview about prominent researches and challenges in the OM area. These studies then will be classified in three research methods include: Field research (F) proposes the kind of collecting raw data in several particular organisms to reveal the habits and habitats in their natural surroundings and present the conclusions, the Lab research (L) represents the theory work to give the newest models, frameworks, algorithm, etc. This type of research helps improve the performance and

effectiveness of OM, and Case research (C) is the method that presents experiments when authors work directly with special case or problem in real context of the organization. After analyzing these articles in OM field, the results below point out three problems that are currently challenging in the OM management area:

Challenges in Managing Knowledge

Source types of OM Since early 80s, the term of knowledge has become popular, accordingly, various memory source types concern with KM were defined by different authors. The memory source type, by its meaning, affects directly to the fundamental functions of OM such as acquisition, retrieval, and retention. By different direction, Basaruddin & Haron [2011], Dieng et al. [1998], Walsh & Ungson [1991], R. T. Watson [1998] defined different source types, which need to be managed by OM such as text, multimedia, model, knowledge, etc. As a result, this is a challenge for OM to manage information effectively from different resources. Furthermore, another challenge for OM is to manage the tacit knowledge [Atwood, 2002]. Because the knowledge is based on skills and experiences of peoples, so it is very hard to collect and transfer to other peoples. Hence, this challenge is still a difficult problem to solve in the future.

Contexts and meaning In the view of preserving document context, this is a challenge for OM to keep the rationale and context of information, which should be acquired and stored concurrently with the information. This problem maybe makes the information lose the correct meaning in later reuse situations. Therefore, we need to maintain the context as well as relevant information for each information, which will help us restore the information lost from the OM hereafter. Concern with this problem, many authors that paid attention to enhance the context of information such as Ackerman & Halverson [2004], Halverson & Ackerman [2003], Miller et al. [2008].

Quality of information returns One problem in retrieving the past information is that we usually face with the incorrect results, which return to our queries. Hence, the challenge here is to enhance the quality of the information return from OM rely on the human beings. To deal with this problem, Bichindaritz [2006] highlighted the efficient ranking through the large case base system. By this method, the ranking could be valued by user review or from assessment of experts, then the ranking system will increase the usefulness of information in OM and supports user to make decision-making more accurate and timely.

Effective managing OM Another challenge is based on the effective managing a large amount of information in the OM. The main reason is that, because information is collected continuously in many years, it will face with the redundancy, irrelevance, and contradiction between stored information in OM as well as with the current environment in use. There are many authors paid attention to enhance this problem such as Bichindaritz [2008], Ochoa et al. [2009b], R. Watson [1995].

Challenges in Managing People

The role of older workers in OM In a research to test the relationship between OM and empowerment, Dunham & Burt [2011] has found a significant relationship between the knowledge of older workers and the requests to share knowledge. It depicts the important role of historical knowledge from the mentors is helpful for decision-making in daily activities and businesses of the organization. Thus, the challenge here is to acquire and disperse the knowledge from older workers to organizational members.

The privacy in OM In another aspect, Herskovic et al. [2006a] showed a relationship between the personal information security and the use of OM. As we know, the OM - by its functions - will capture individual knowledge and publish to other members to support daily operations. By contrast, the published opinions maybe bother the

owner benefit. Hence, the sharing knowledge could not be completely honest because of the privacy protection issues. By this way, the quality of knowledge will be decreased and then affects directly to the quality of OM in the long term. Thus, this is a challenge for the OM to secure the privacy enough to help users could feel free while sharing their knowledge honestly.

Challenges in Practices Used

Theory and practice The gap between theory and practice in OM is still a large step. Thus, Atwood [2002] indicated the challenges to develop the OMIS in practice that the implement needs intensive on the context where OM will be used, as well as, the new system of OM must be robust and complete to avoid the disruption in normal operations of organization. Moreover, the development must focus on the knowledge and competencies of employee to facilitate their access and their learning [Abel, 2008].

Intelligence factor Concern with the development costs, this challenge leads managers in building the OM system to manage a large amount of information with the lowest costs and may respond to rapidly changing environment of organizations. By these demands, it is necessary to include the intellectual factor to support the OM acquisition, retention, maintenance, and retrieval automatically. These autonomous methods will minimize the management costs of OM such as automatically collect information, enhance stored information, and deliver the right information to the right users [Abecker & Decker, 1999].

OM success factors Parallel with the development of OM system, another challenge for OM is to assess the effectiveness of the implemented OM in real environment. This aspect plays an important role in defining the OM success factors to help managers measure the effectiveness of OM in supporting organizations. Consequently, the analysis and assessment about the OM can help organization to have a better ad-

just and maintain the OM performance. Following framework 'The DeLone and McLean Model of Information Systems Success' [Delone & McLean, 2003], in a study, Hamidi & Kamaruzaman [2009] adapted these impact factors to assess the OMIS at a nuclear power plant and they concluded positively that the assessment will not only refine the current model, but also enable a comparison of the effectiveness of various OMIS models.

4.2 Organizational Memory Information System (OMIS)

4.2.1 Concept of OMIS

As defined in the last section, the OM concerns with functions such as acquisition, retention, and retrieval in order to support organizations have an effective decision-making throughout managing the past organizational experiences and information. In this view, while the OM emphasizes the point of human factor and process to manage organizational knowledge, the OMIS is a component of OM [Stein & Zwass, 1995] that deals with the technical approach to realize the OM with the support from Information System (IS). Therefore, the OMIS - which relies on the computer based - intends to realize a system to help the OM manages shared memory more effectively.

Stein & Zwass [1995] defined OMIS as "a system that functions to provide a means by which knowledge from the past is brought to bear on present activities". These authors indicated that the OMIS mainly focuses on the acquisition and distribution of information in the past, as well as using it to solve problems in the present based on computer technology support. As a result, while the OM concerns with capturing the overall knowledge from individuals and groups in organization, the OMIS deals mainly with facilitate the providing knowledge by using computer devices. The OMIS supports OM by giving a direct access to its huge repository of documents, experiences, and reports; in addition, the OMIS also supplies indirect access links to external sources. As an illustration, the

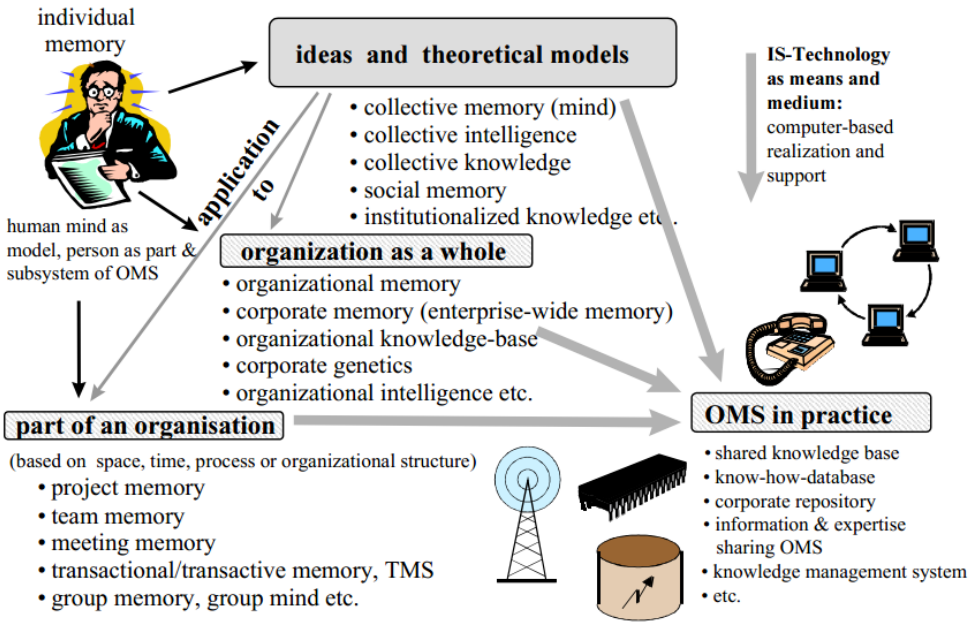


Figure 4.2: Relationship of OM-related terms.

relationship between OM and OMIS has been indicated by Lehner [2000] in the figure 4.2 that proposes the critical role of each term. The authors indicated OM is a part of the organization and it derived from the term of individual memory. In another view, the term OMIS is adapted from OM concept and IS-Technology (computer-based realization and support). Hence, the OMIS has a close connect with the term OM in the light of their functions, by contrast, the OMIS is intensive in technical aspect more than the OM concept.

In order to make clear the OMIS concept in practices, Wijnhoven & Slooten [2002b] indicated three problems, which are big challenges for OMIS, such as the integration, diversity, and closeness. The first challenge is that the integration of OM content needs to be easy enough to link and transfer information among people and other systems. The second challenge is to handle a diversity of memory contents in a clear structure and united system. The third challenge is intensive on keeping the OMIS as close as possible to person's memory that supports intelligent learning.

4.2.2 Functional structure of OMIS

As we know, the term Organizational Memory Information System is combined from two areas of the Organizational Memory and the Information System. Hence, OMIS is a part of OM and it still deals with main functions of OM as acquisition, retention, and retrieval but intensives on information system aspects. Following the framework of Stein & Zwass [1995], the figure 4.3 shows us the two layers of OMIS. The first layer includes four subsystems for instance: integrative, adaptive, goal attainment, and pattern maintenance. Each of them is the effective function proposed in the OMIS framework as described below:

Integration is a main subsystem of OMIS, which has responsibility in providing the coordination and management information across the organization. Thus, this function requires integrating the system over space and time.

Adaptive is a function of the organization to recognize, capture, organize, and distribute memory of the organization's environment and support the OMIS an ability to adapt to changes in its environment.

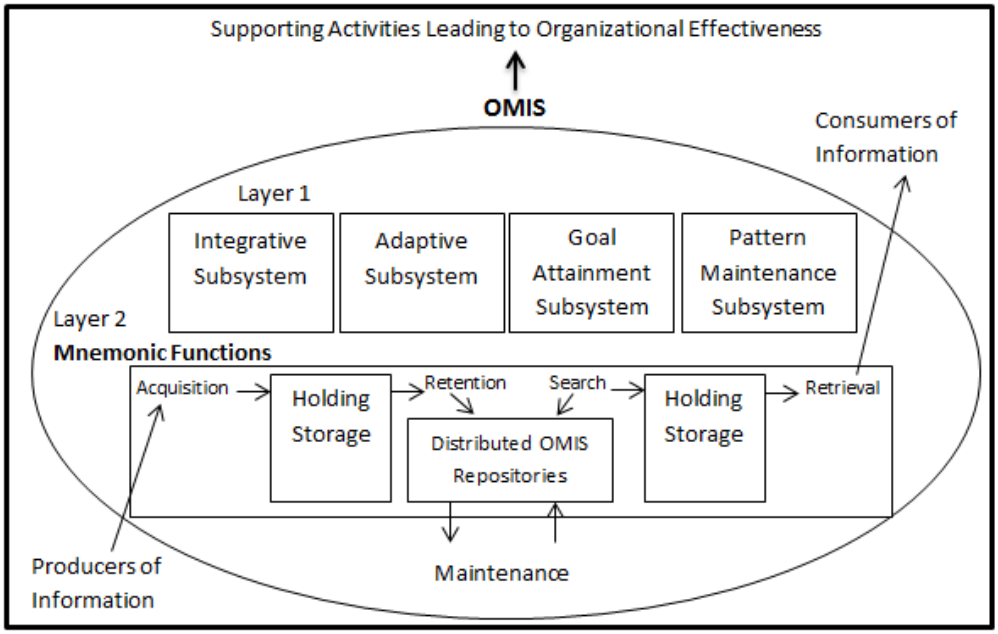


Figure 4.3: Framework for an OMIS

Goal attainment focuses on supporting organizations in planning and controlling evaluation models, plans, and performance by templates. This function shows the ability of the organization in setting goals and evaluating the degree of their fulfillment.

Pattern maintenance is a specific function that relies on human resources to enable organization remember the work history of organizational member (e.g., attitude, value, and norm). Moreover, the pattern maintenance also supports the OMIS an ability in maintaining the cohesion and the morale of the workforce.

The second layer of OMIS focuses on the Mnemonic functions, which encloses knowledge acquisition, retention, maintenance, search and retrieval. These functions have a higher definition in technical aspect than the first layer. In this view, the Mnemonic functions intend on the workflow of information in OMIS and respond to the question 'how information can be managed in OMIS?'. Consequently, the acquisition function of OMIS plays a role of the input process for both internal and external data sources. For example, the acquisition concerns with input data such as emails, documents, databases, bar-code data, reports, etc. The retention function is to remember relevant information of the organization and it is influenced by the communicability and consensus among organizational members. The maintenance function has an important role in refining (e.g., update, discard, delete, and integrate) the data and information kept in OMIS. The two final functions are used to search and retrieve information in OMIS, which have a well defined structure to support searching and retrieving information faster and more extensively than manual systems.

4.2.3 The roles of the OMIS in supporting the OM

Based on the objective of OM to store information collected from organization's history for future reusing, the OMIS is an IT-based solution to realize the OM by the using of information system. The OMIS in thus deals with technologies that come from advantages of IT-based to support the OM as represented below:

- The first ability of OMIS is to memorize a large amount of information by using the permanent device such as disk, tape, or optical disc storage. Nowadays, with the evolution of a high-technology, the memory of OMIS is large enough to support OM with a huge terabyte space that can hold all information and data of organization over the decades.

- The fast search and retrieve information is the second advantage of OMIS in supporting OM to retrieve information from the mass storage devices quickly by special queries, filters, and random searches from the structural records such as meta-data, with the advantage techniques in content management (e.g., ontology, CBR, pattern matching, natural language processing). Thus, this method is used more effectively to support the OM than the paperwork or the manual systems such as file management system.

- Following, the maintenance function of OMIS is also effective in managing the current knowledge in OM. By the support of information system, we can define automatic processes to refine information in the OM by specific algorithms that could maintain the stored information in order to avoid the redundancy, irrelevance, and inconsistency problems.

- Finally, with adaptive function, the OMIS is a solution for organizations to adapt to rapid changes of business environment such as the change of techniques, markets, or customer trends. This function allows managers to set goals and evaluate the strength and fulfillment of the OM.

4.2.4 Previous studies of OMIS

Since 1995, after four years from the first presentation of the OM concept by Walsh and Ungson (as an integrated form); Stein and Zwass enhanced OM concept to a higher level that intends to realize the OM concept with Information System aspect and these au-

thors called this new concept was the organizational memory information system (OMIS). This was an important milestone in the development duration of organizational management theory; and it motivated many researchers rely on the bright future of OM's implementation by embedding OMIS in a real organizational workflow. Based on the OMIS framework and the requirement of this dissertation, the research strategy will find out the concepts that deal with the OM functions and then summary these studies on the matrix of concepts to help solve the main research question. Through the bibliographic research strategy, I selected 27 articles that published in the science technology category (e.g., Google Scholar, DBLP, Web of Knowledge), which are refined throughout two main keywords 'Organizational Memory' and 'Business Intelligence' for finding the relevant researches in the past. Consequently, these articles will be analyzed and separated by the matrix of concepts as indicated in the table 4.1.

These 27 articles are collected from published articles between the years from 1995 to 2009. In addition, they are classified in six concepts as acquisition, retention, maintenance, search & retrieval, case research, and evaluation. The first four categories adapted from four main functions in 'mnemonic layer' of the OMIS basic framework [Stein & Zwass, 1995]. The case research indicates research methods that apply practices on organizational environment to test the effects of OMIS in real organizational environments. The last concept is the evaluation, which will help us maintain and improve the quality of the OMIS by defining success factors of OMIS, this research also helps organizations to assess and build up the keys of success when implementing their own OM.

With the cross analysis main studies in OMIS, we can highlight authors who deal with refining the OMIS framework as Linger et al. [1999] proposed 3-layered framework for a dynamic OMIS, Mokhtar et al. [2008] with an adapted framework of the Faculty Memory Information System, and the integrated OMIS framework based on transaction memory of Nevo & Wand [2005]. In the real case studies of OMIS in organization, we have three studies from Croasdell & Paradice [2002], Lang & Schmidt [2002], and Jennex & Olfman [1997b] deal with the practices to realize and refine the OMIS framework. In

another view, there are some authors focused on defining the success factors of OMIS, this is the keys to lead the organization into successful while implementing the OM system. With the research leading of DeLone & McLean [2003] in identifying seven dimensions to evaluate the success of an information system, Jennex et al. [1998] gave an extension from the DeLone and McLean's I/S success model to introduce the OMIS success model. In the same orientation, there are Hamidi & Kamaruzaman [2009], Wijnhoven & Slooten [2000], and Wargitsch et al. [1998] also intensive in this aspect. Finally, concerning with traditional way, all last authors mainly focused on refining four basic functions of OMIS such as acquisition, retention, maintenance, and search & retrieval functions to enhance the performance, processes, and models of the OMIS.

Table 4.1: Main researches on the OMIS area.

Study	Acquisition	Retention	Maintenance	Search & Retrieval	Case Evaluation
Ochoa et al. [2009c]	x	x	x	x	
Hamidi & Kamaruzaman [2009]					x
Mokhtar et al. [2008]	x	x			
Herskovic et al. [2006b]				x	
Chou & Cheng [2006]			x	x	
Hwang & Salvendy [2005]		x		x	
Nevo & Wand [2005]		x			
Ji & Salvendy [2004]				x	
Vasconcelos et al. [2003]		x			
Fortier & Kassel [2003]		x		x	
Ji & Salvendy [2002a]				x	
Ji & Salvendy [2002b]				x	

Continued on next page

Table 4.1 – *continued from previous page*

Study	Acquisition	Retention	Maintenance	Search & Retrieval	Case	Evaluation
Croasdell & Paradice [2002]					x	
Lang & Schmidt [2002]					x	
Goesmann [2001]				x		
Kuhlman & Deiters [2000]	x	x		x		
Wijnhoven & Slooten [2000]						x
Linger et al. [1999]	x	x	x	x		
Hackbarth & Grover [1999]	x					
Wargitsch et al. [1998]						x
Burstein et al. [1998]			x	x		
Jennex et al. [1998]						x
Morrison [1997]	x	x				
Burstein et al. [1997]		x				
Jennex & Olfman [1997b]					x	
Jennex & Olfman [1997a]				x		
Stein & Zwass [1995]	x	x	x	x		

4.3 Business Intelligence for OMIS

4.3.1 Business Intelligence Concept

The term Business Intelligence comes from a branch of Artificial Intelligence science that initiated since the decades of 1980. However, in that time, the BI application was not popular because of the lack of technical support. In the 21st century, the explosive growth of BI, with many published articles and applications of BI, is gradually realized through-

out the organizational environment all over the world. The BI functions mainly focus on using intelligent tools from information technology, especially in AI area, to support the organizational members in making better and faster decisions. In an integrated view, Cody et al. [2002] defined BI as a system combining data with analytical tools, in order to provide relevant information for decision making. The purpose of these systems is to improve the availability and quality of complex information, BI has been adopted in Aerospace, Banking, Chemicals, Health case, etc. In addition, it exhibits the advantages of helping organizations realize their strategy.

Based on the definition of BI system, Chaudhuri et al. [2011] summarized and exhibited the traditional architecture of a BI system as presented in the figure 4.4. This model classified the basic BI architecture into five tiers by their specific objectives to support the BI system as follows:

Data source is an input of the BI system, it includes organization's internal data (e.g., ERP, CRM, POS, CMS) and external data sources cover information that comes from the outside environment for instances: the stock data from the security company, the market data from specific service providers. These data usually exist in raw forms, which do not contain any context, purpose, or meaning, and stay in diverse formats from multiple sources (e.g., structured, unstructured).

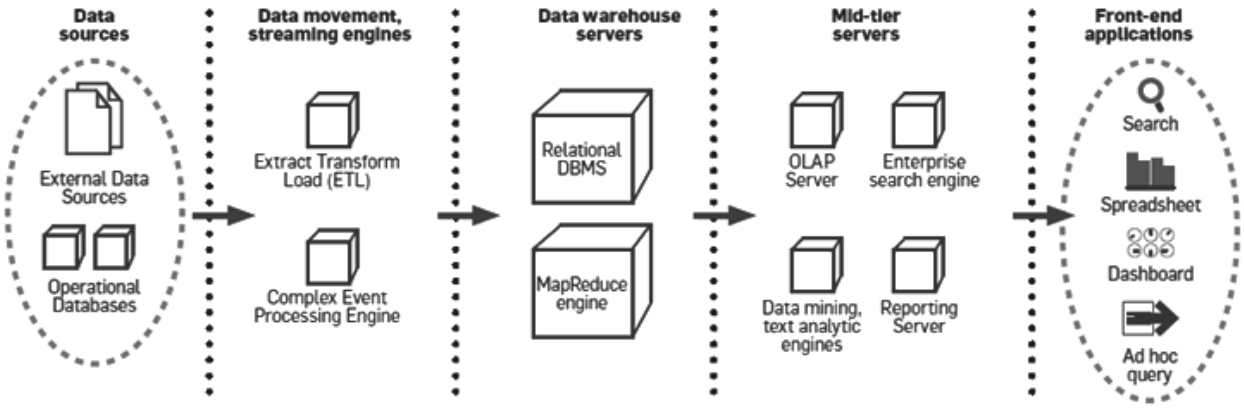


Figure 4.4: The basic architecture of BI system

Data movement, streaming engines play a challenging role in classifying, integrating, cleaning, and standardizing all raw data from Data Sources. This is a complex component because it is required to collect data reliably, timely, and integrally. Any loosing information may affect widely and directly on future analysis results. Thus, we normally use the Extract-Transform-Load (ETL) tools to adapt the input data and use the Complex Event Processing engine (CEP) to interact with the real time system as the stock-exchange session.

Data warehouse servers are the back-end storage of BI system. This tier support BI system to handle the structured data (from previous ETL, CEP tools) in a huge repository. Therefore, the DW repository must support enough space to store working information/data continuously and requires an appropriate structure to enhance the response time for performing complex queries and analysis in historical data repository. Normally, we use the Relational Database Management System (DBMS) to store and manage data for the DW (e.g., MySQL, Oracle, and MSSQL), which are compatible with the complex SQL queries over huge records of the DW. Moreover, with the globalization trend nowadays the DW is required to support the Distributed Database Management System (DDBMS) to connect the data in different places or countries all over the world.

Mid-tier server contains the main components of BI that enhance the business analytic environment. The goal is to support the BI system in creating reports, query, and analyzing information inside the data warehouse. In Mid-tier server, the OLAP and Reporting Server play an important role to help user manage, expose, and create the static reports as well as dynamic reports from complex queries; especially in multidimensional view such as drill-down, filtering, and aggregation. In addition, the data mining and text mining have a responsibility in deeper analyzing and classifying irrelevant or unstructured data in the DW. The main functions of DM, which are integrated from the AI techniques such as prediction, association, and cluster, will identify the patterns based on attributes of information and data in the DW.

Front-end applications constitute the User Interface (UI) tier. This part keeps user transparent with all previous technical tiers, but still support complete management functions of BI. In general, this tier includes the search, spreadsheet, dashboard, and ad-hoc query to connect to back-end tiers and support user in a friendly way. One of the prominent functions in this tier is the dashboard technique, this provides a complete visual view of corporate performance measures (KPI), trends, and expectations. Dashboard presents KPIs by graphic view and allows users to interact with respond immediately. By this functionality, the dashboard provides an overview and multidimensional view from the updated information that will support users in effective decision-making. Moreover, the UI tier supports automatic update the newest information based on the continuous collect data from the previous tiers.

4.3.2 How BI has been integrated in OMIS

In order to make sense the feasibility of BI systems in supporting the OM, this section refers to tools and studies of BI to deal with functions of OMIS such as acquisition, retention, and retrieval. The BI systems, which based on architectures and tools, then will be clarified why and how they are integrated in OMIS as the indication follows.

- Acquisition:
  - According data acquired, there is an important demand to collect data from the dynamic environment, which consists of unpredictable changes from directly affected agents to the organization, for instances: dynamic changes in technology, politic, trend, and economic. To interact with dynamic environment we can use the ETL engine, this supports OM to adapt to dynamically change as well as to extract, transform and load data from many data sources as transaction systems and the Internet into the warehouse repository [Blue et al., 2011].
  - Based on Semantic Web tools to support networked knowledge acquisition. The Semantic Web Tools are compatible tools for supporting semantic knowl-

edge in collecting and sharing more exactly and effectively. With the ontology architecture, Chapman et al. [2009] in an application demonstrated that semantics alone cannot solve a problem of knowledge sharing and reusing, and then he proposed a hybrid search modality of K-tools as a possible system to support the KM. By this way, he proposed the K-Forms and K-Search that support knowledge acquisition and representation, as well as, enable flexible searching and sharing the semantic knowledge.

  - To help managers have adequate assessments about their organizational products and services, it requires the OMIS could gather comments and assessments from direct users to their products and services via their website and forum. This information will help the organization in discovering the most important topics, trends, or mistakes in business operations; and then managers can adjust or create new knowledge for OMIS. Thus, Zhan et al. [2009] depicted a new approach to enhance the organizational perception by using BI to gather customer concerns from on-line product reviews.
  - Another requirement for OMIS is to manage a number of unstructured data such as documents in organizations and unrelated data/records from difference transaction systems. Based on the advantages of the mining engine to analyze raw data, the OMIS exposes the urgent need for integrating the data mining and text mining in order to help the OMIS analyze and classify the organizational knowledge automatically. In a study, Wei & Qing-pu [2007] highlighted a significant role of the mining engine of BI and established solutions for knowledge mining model to improve the accuracy and efficiency of knowledge acquisition.

- Retention

  - To manage knowledge, the ontology shows a significant role as a formal description of concepts and relationships to keep track of knowledge. Cheng et

al. [2009] in the implementation view presented an ontology-based approach for BI applications that focus on statistical analysis and data mining in retention and retrieval knowledge.

- Mining method is another approach of BI to improve the accuracy and efficiency of knowledge stored in OM. In a study, Nasukawa & Nagano [2001] defined a basic prototype system named TAKMI to find the valuable patterns and rules that hide in very large amounts of textual data contained in OMIS. By this way, the text mining and data mining in the same purpose are useful tools to help OM in mining the rule extraction, which rely on a large set of methods and algorithms of DM, for instance Farquad et al. [2009] used Support Vector Machine and Naive Bayes Tree methods to perform the rules extract in bank credit card successfully.
- Case-based reasoning (CBR) is a powerful method of machine learning to help solve new problems by using past solutions. By its own purposes in decision-making, CBR plays a coherent role in the OMIS as a method to support the explicating of knowledge. The BI that relies on its advantage will support the OMIS in managing the CBR such as reusing experiences stored in knowledge-based systems to solve the current problems quickly [Mahapatra, 1997], as well as, predicting possible failure cases by using the past experiments [Borrajo et al., 2011].
- Another advantage of BI is the OLAP cube; this is also an easy way to present the historical data and information contained in OMIS through multidimensional schema. In this view, Sciarrone et al. [2009] presented a multidimensional OLAP schema to support information retrieval in an ontology-based environment. Their research indicated a method to integrate the hierarchic structure of ontology into the OLAP to manage effectively the non-structured documents stored in DW.
- As a requirement of OM to manage knowledge acquired, BI needs to support

the knowledge container for easy extraction, cleaning, and storage. Nemati et al. [2002] used the knowledge warehouse (KW) architecture that might be integrated with the OMIS, and they concluded that the KW has capability of capturing and coding knowledge, enhancing the retrieval, and sharing knowledge across the organization.

- Retrieval
  - In retrieval components, the OMIS need to be supported the ability to search and query in large amount of data and information by using the search engine of BI system. The BI supports an intellectual factor to analysis and classify the patents from irrelevant and unstructured information into separate categories and groups. Hence, it gives OMIS an ability to speed up the search and retrieval relevant patents effectively.
  - With orientation to help OMIS becomes more friendly and effectively in problem solving, the BI with powerful User Interface (UI) components such as dashboard, browser, and portal can support OMIS to illustrate the information more effectively. In an experimental research, Chung et al. [2005] with the visual framework for knowledge discovery, finally, highlighted the usefulness of graphical displays that was found to be more effective, efficient, and usable than the result list. Hence the BI, based on the UI components, plays an important role to interact with the user as well as representing information stored in OMIS via compatible tools.

4.3.3 Potentials and pitfalls of BI in OMIS support

Despite the fact that BI system is such a useful technique to support the OMIS in managing a large amount of data and information collected across many years, so that improve the problem solving and decision making. However, there still have some pitfalls and potentials of the BI while integrating with the OMIS as follows.

Potential of BI to support OMIS

- The first potential of BI to support OMIS is to utilize data warehouse with the OLAP cube, which might support the OMIS to collect data and information from many data sources and then disperse information to particular organizational members to perform specified tasks. Furthermore, it can quickly create multidimensional reports to help enterprises realize their strategic objectives easier and more efficient.

- Second, in the user view the dashboards and other information broadcasting tools are such an effective way to present the synthesis of past information in graphical view. The BI then has significant supports the OMIS to compatible with user working environment (e.g., word, spreadsheet) by offering visualization tools and allowing interaction with user in problem solving and decision making.

- The other advantage of BI in supporting OMIS is the ETL engine, it is a strong engine to support the acquisition process in collecting and standardizing data from many sources, which can run timely and exactly based on the defined rules.

- Consequently, the BI uses AI engine such as Data Mining and Text Mining to support OMIS maintenance and refine information. The AI plays a main role in automatic analysis and classify the useful information with the support from autolearning methods. Hence, these methods will enhance the cost efficient and lower the organization's expenditures.

- Finally, the BI not only supports OMIS to actively collect and deliver relevant information for re-creating knowledge, but also supplies to forecast the future information from past information. By this way, the BI is an appropriate method to help organization avoid the repeat of previous actions that have been done in the past.

Pitfalls of BI to support OMIS

- As we know, the organization, via their works in a long time, generated so much data and information. Instead of the advantage support from BI tools such as DW to handle large data stream, however the pitfall here is that we usually keep all data in the DW beyond what we need and believe that they will be useful in the future, even not for now. This problem will limit the effectiveness and efficiency of the BI in identifying appropriate results and slowing down the system. Therefore, we need to understand clearly what data sources should we use and how can these data support OMIS in problem-solving and decision-making. Moreover, the OMIS should limit and filter the data sources before saving into DW, it is recommended for covering information just simple enough to support appropriate solutions according to organizational requirements.

- In technical aspect, because each organization has different environment to implement the BI system to support OMIS. Hence, the BI system for this organization is not used correctly for another organization and it should be reimplemented in the specific environment of that organization. In addition, the BI system provides a wide range of tools; hence, the development of BI-OMIS integrated model is required to identify which appropriate tools are required for the particular case of an organization, as well as, to avoid the abuse of BI tools like a metaphor of using a sledge-hammer to crack a nut.

- The BI is a good tool with the support of ETL to help cover all kinds of information in an organization but it could not guarantee to encompass all the knowledge in a real organizational environment. Especially, BI was limited in managing the tacit knowledge, and the informal information such as experiences, thinking, aptitudes, perceptions, and insights.

4.3.4 Directions for further work

At first, based on the four main functions of OMIS framework, one objective of BI in supporting OMIS is to improve and maintain the process of information management. Hence, this research needs to define the retention, acquisition, and retrieval tools of BI to support the OMIS manages formal knowledge as well as refine the model of BI to support OMIS for better decision-making and problem-solving.

The other aspect to make clear the BI process is to enhance the functions of Data Mining, Text Mining, and Knowledge Mining. This is required to define a detailed process with the support of rules and methods to integrate the raw data and transform them to semantic information.

With the on-line analytical processing (OLAP), it is also a point to analysis what is the dimensions need to be setup with OM environment. Based on the advantage of OLAP cube, the future work for this research is to define facts as well as dimensions concern with OMIS, especially, in the crowdsourcing initiative.

With the proposed BI-OMIS model, the last improvement requires to be validated in a real context. The result will be analyzed and assessed by using the crowdsourcing initiative called CrowdUM. The final result from assessment methods then will propose an experience of using the BI system to support OM and show the effectiveness of this proposed framework.

4.3.5 A proposed framework for evaluating

Through all previous analyzes on both of the advantages and limits of BI system in supporting the OMIS, this section is going to propose an adaptive model as indicated in the figure 4.5. The model, which is called the BI-OMIS integrated model, is combined between relevant tools of the BI system with the OMIS framework in order to support the OMIS by improving knowledge sharing and effective decision making.

By this way, the main BI tools that are used in this proposal then will be classified into four layers such as: Acquisition, Retention, Maintenance, and Search & Retrieval layers. These layers correspond with four main functions of the mnemonic layer in OMIS framework. Consequently, for each layer of the BI-OMIS integrated model, I will present descriptions about BI tools that make sense why and how these tools might be adapted to the OMIS framework as follows.

Firstly, the Acquisition layer plays an importance role in capturing and adapting data as well as information from many resources concerns with OM. Thus, I identify two tools, which have a significant role for OMIS, to enhance the acquisition functions as:

1. **Web Crawler**: this is a tool to collect the data from the web sources as forum, web, on-line services, etc. The web sources are one of the external knowledge that contains wide information with a rapidly updated. Hence, the web crawler is an appropriate tool to support OMIS in integrating web data and searching relevant information about competitors, customers' trend, and market segment.

2. **ETL**: this tool encompasses the functions such as extract, transform, and load data in many data sources and then classify, normalize, and load data into an organizational repository (e.g., data warehouse, data mart) based on the defined structures (e.g., meta-data, ontology structures). This tool plays a critical role for the OMIS to acquire almost data in organizations.

The second component of the BI-OMIS integrated model is Retention layer, this layer has a responsibility in managing capability of organizational repository. Thus it mainly focuses on the database tool called the Data Base Management System (e.g., MSSQL, Oracle, MySQL). Moreover, one of the main problems of OMIS retention is to store a large historical data across many years, hence, the Data Warehouse and Data Mart are compatible repositories to manage a number of historical data throughout the organization.

The third layer is the maintenance component, in this layer, BI tools will be adapted to enhance, analysis, and classify the stored data, which has a definite structure in the

organizational repositories such as data warehouse and data mart. Like a metaphor of intelligent engine for the BI-OMIS integrated model, this layer includes five components of BI system to help maintaining the information and knowledge in an effective way as indicating follow:

1. **OLAP server**: this is a multidimensional data-managing server to help the user interact with the data warehouse visually. Consequently, the DW might be filtered, aggregated, and drill-down throughout the multidimensional views. The OLAP also allows user to reorganize and calculate a massive historical data quickly based on functions of the OLAP cubes. By this way, user could use the OLAP cubes to create their own star schema in various dimensions to speed up the retrieval and query processing from large amounts of data stored in the OMIS.

2. **Ranking engine**: is one of the database tools proposed by Fan et al. [2005] to enhance the search engine retrieval performance. Through the ranking function research, the authors concluded that "the retrieval performance of these newly discovered ranking functions has been found to be superior to the performance obtained by well-known ranking strategies in the information retrieval literature". In general, the ranking in BI-OMIS system can be assessed by all members in organization to refine the usefulness of past information stored in the OMIS.

3. **Search engine**: based on the main functions of BI search engine, this function supports employees to perform the searching based on the structured data that is stored in knowledge repositories such as email searching, document searching, task searching, and best practices finding. Moreover, this function also allows users to retrieve the history of data modified via the time line of the document. Thus, this is an effective method to help user retrieve the past context of knowledge throughout highlight the new updated contents, as well as identifying these authors who performed on the document.

4. **Reporting server**: the reporting tools of BI provide suitable mechanisms that en-

able definition, efficient execution and rendering of reports [Chaudhuri et al., 2011]. This tool allows managers to export the information from an OLAP cube in dynamic forms/templates, which will facilitate the operations of managers in creating, modifying, and presenting a number of reports such as the balance sheet, product reviews, or annual financial report without requiring so much knowledge in database technique.

5. **Data Mining and Text Mining engine**: they are the main components of BI that enclosed intelligent methods to help us analyze the complex and unstructured problems such as classifying, knowledge learning, problem solving, heuristics, etc. The mining methods use AI algorithms to deduce and propose new knowledge beyond what were saved in DW and OLAP. In addition, the mining engines provide an ability for OMIS to forecast the new knowledge relies on intensive analysis and synthesis historical data. In another hand, the text analytic engine helps us understand more about the large amounts of text data contained in organization's documents like business contracts, customer information, annual reports, customer reports, etc. The most popular functions of text mining are to solve problems such as information extraction (identifying key phrases and relationships of text), topic tracking (predicting relevant documents based on user's profile), summarization (proposing the main content), categorization, clustering (grouping similar documents), concept linking, and question answering [Turban et al., 2010, p. 289].

The last part is the front-end layer that is intensive on user interface (UI) methods to create an interactive channel with users. The user then might operate with all functions in the third layer of BI-OMIS model through these UI components. Thus, this layer includes tools such as Knowledge searching, Spreadsheet, and Dashboard. Firstly, the knowledge searching improves the basic search of BI by enclosing with knowledge searching support (e.g., rules, case based reasoning), it gives user a friendly interface with search fields and search conditions to help users find out the most appropriate information for solving problems. Next, the spreadsheet is still a useful environment for users to manipulate

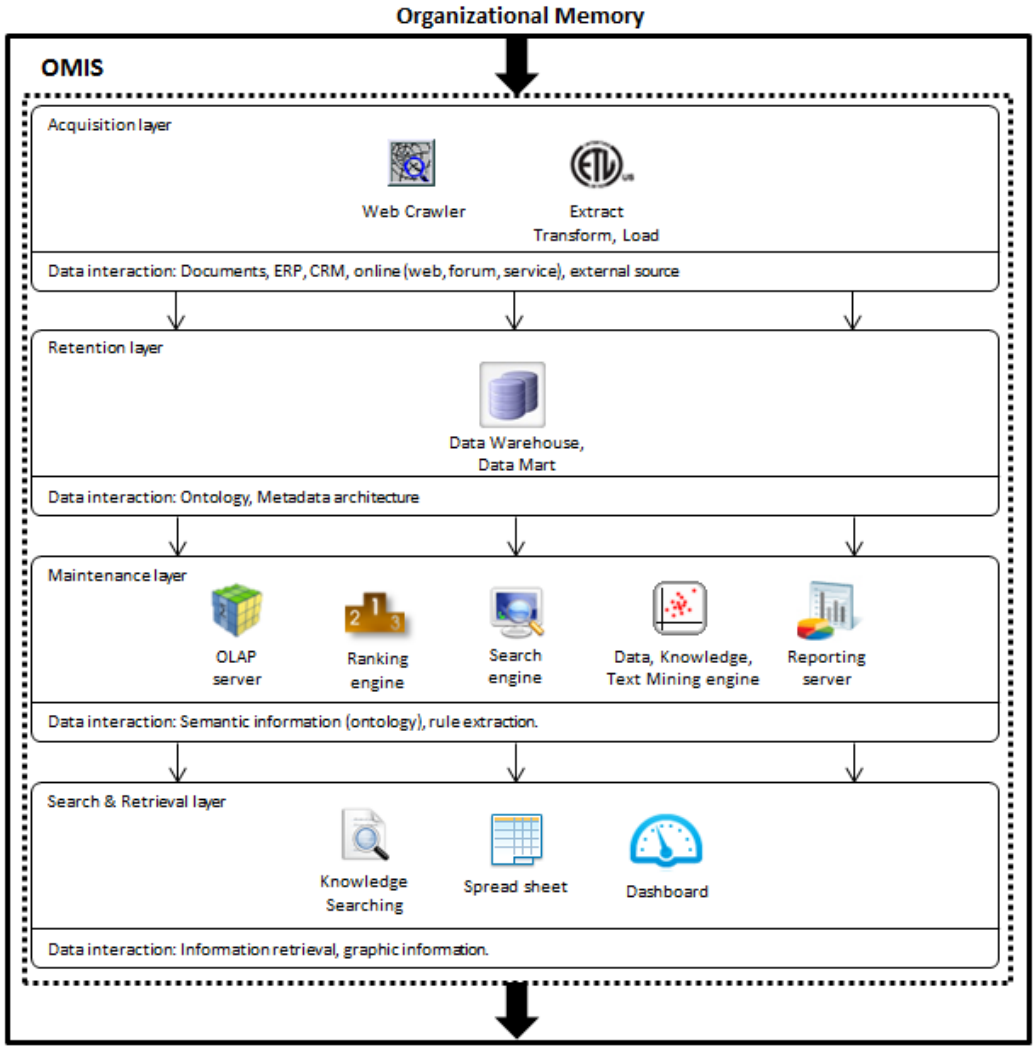


Figure 4.5: The BI-OMIS integrated model.

with the OLAP, which can quickly retrieve the data from OLAP in various dimensions and present in drill-down view. Finally, the dashboard is a graphical interface for indicating the key performance indicators (KPIs) of OMIS. In addition, the dashboard enables decision makers to keep track of the current status of organizational environment with a simple visual view.

Based on the objective of this research, the future works will intend to use this BI-OMIS integrated model to analyze and improve the OM of a crowdsourcing initiative in order to refine the model and improve the memory processes of the CrowdUM.

Chapter 5

BI-OMIS proposal for CrowdUM

5.1 Crowdsourcing concept

In the recent decades, the internet is gradually becoming the backbone of organizations. It not only connects people in the organization closely throughout time and locations, but also keeps our works up to date by the immediately sharing their own information. With the main objective to make use of these advantages of the internet to establish a crowd network between the requester and the large groups of people in the purpose to enhance the working performance, as well as gaining more profits for participants in the crowd. By this aspect, the term 'crowdsourcing' is derived from the business area to indicate the process of connecting a large crowd together via the internet, which allows seekers to address almost large projects by using crowd's creativity and talent. For that reason, these projects usually require a large amount of people who have a moderate ability such as students, normal employees, or local residents all over the world.

In a definition, Howe [2006] concluded that "crowdsourcing is the act of taking a job traditionally performed by a designated agent and outsourcing it to an undefined, generally large group of people in the form of an open call". This definition focuses on discriminating the traditional job from the crowdsourcing. In which, the authors emphasized the

role of large groups in selecting the jobs through the open call instead of dealing with specific customer directly. As a result, the crowdsourcing could create solutions more quickly and cheaply than the traditional method such as hiding new workers/agents, and likewise the customers could also collect more suitable solutions for their problem.

Throughout the evolving time of the crowdsourcing as well as based on the purpose of each crowdsourcing group, the crowd is often classified into six categories [Ramos, 2011] as follows:

- **Crowd wisdom:** The crowd wisdom focuses on collecting the innovations via asking the crowd, inside and also outside company, to solve problem, forecast future, or propose a strategy for a specific environment. Then, the best wisdom is going to help their company solve specific problem. In this way, we have the simExchange[1], Iowa Elections Market[2], and Hollywood Stock Exchange[3] as examples.

- **Crowd creation:** The crowd creation has a little bit difference from the crowd wisdom. In this purpose, the crowd and requester have to work together to co-create the new information/knowledge for examples: NASA's Clickworkers, Innocentive[4].

- **Crowd Voting:** The voting is a method of inquiry into the opinions of the crowd to collect their ideologies about the designs, campaigns, and trends of a product or technology. The result then supports company to modify the product before they turn in bulk production such as American Idol, or Threadless T-shirt company.

- **Crowd Funding:** The main objective of crowd funding is to aggregate a group of shareholders, who will build up the funding of projects or creation of assets as Kiva and Sellaband companies. The outcome and responsibility of the crowd in this type is higher and more important than other types of the crowdsourcing.

[1]http://www.simexchange.com
[2]http://tippie.uiowa.edu/iem/index.cfm
[3]http://www.hsx.com
[4]http://www.innocentive.com

- **Crowd Democracy:** This kind of crowd mainly focuses on the participants who give decisions of local, regional and national governments. This field is so close and the result will affect directly to the thriving or declination of a country.

- **Crowd Review:** The crowd review pays attention on the available objects such as products, services, or events from requesters (e.g., companies or manufacturers). In order that, the crowd is going to share review information about their activities while using it, for instances: visited monuments, tried restaurants, watched films, verified the projects or services. These reviews then support companies have a better understanding about the quality of their products.

Throughout these six categories, the crowdsourcing shows advantages in various areas based on the potential of crowd that could reduce the cost and time in solving problems. In another hand, the crowdsourcing still exists several pitfalls that prevent the expansion of crowdsourcing. First, there is a difficulty in managing a large scale of crowd such as jobs, culture, human being, knowledge, etc. The second obstacle is the management of an immense results/solutions returned from the crowd, hence it requires the crowdsourcing to add more costs in order to filter and select such a suitable solution. The third problem is that the crowd usually creates a lower quality of work, thus this requires an appropriate process and technique to manage jobs tightly and clearly. Fourth, the cooperation of the crowd to solve projects is still a prominent problem. In order that, it needs an appropriate mechanism to support the working process, and then increase the quality of crowdsourcing solutions. As an important impact factor to determine the thriving of crowdsourcing in the long-term, this is one of the objectives of my dissertation to improve the working environment of the real crowdsourcing context.

5.2 Functional structure of CrowdUM

To make clear the features of crowdsourcing in a real context, the figure 5.1 below indicated the use-case diagram of the CrowdUM that relies on the fundamental structure of crowdsourcing. In general, the CrowdUM plays a main role in the crowdsourcing model as a brokering service or intermediary. The CrowdUM is designed for organizations such as companies, personals, or agencies who have ideas, problems, and requirements that need to be accomplished through simple activities but creatively; this actor called the 'Seeker'. On the contrary, the crowd may be called the 'Solver', who are able to gain experience, create knowledge and become more entrepreneurial, they are almost the current students and former students of the University of Minho. Moreover, this is a dynamic group that could propose more creativeness in exploring the solutions regardless of their nominal cost. The third component in CrowdUM appears as a service brokering, called the intermediary, who pays attention in finding requirements from seekers, and then spread them to right solvers. For more detail, the figure 5.2, which was introduced by Professor Isabel Ramos [2011], represented main functions of crowdsourcing initiative in the form of action diagram to help us comprehend the CrowdUM's process.

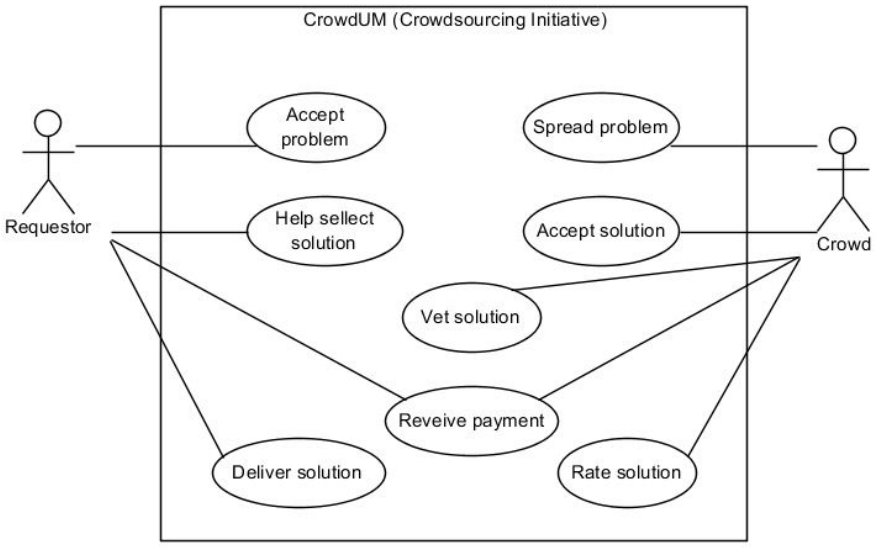


Figure 5.1: The crowdsourcing use-case diagram

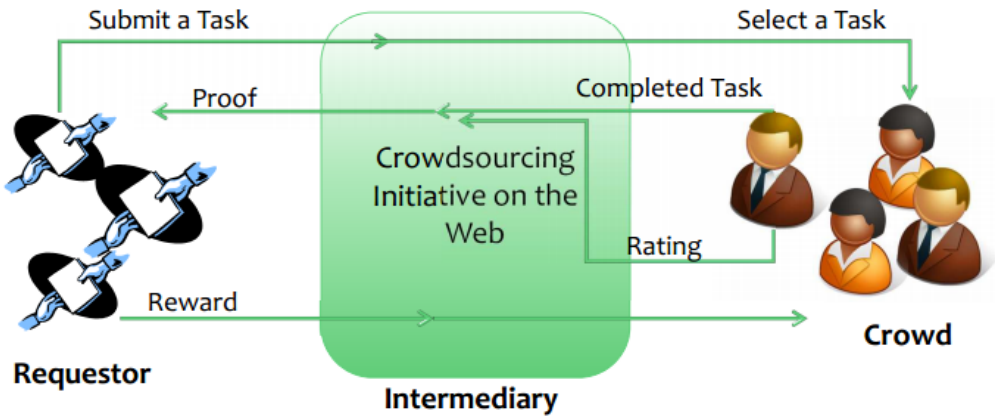


Figure 5.2: The CrowdUM active diagram

With the aim of creating intellectual assets and providing creative solutions to existing problems, the CrowdUM allows seekers and solvers to access to a wide range of skills and collective knowledge via the web base platform as an intermediary. Moreover, the CrowdUM is a tool for innovation, which used properly to generate new ideas, reduce the time to research and development projects, lower costs, having the power to create a direct relationship between the seekers and solvers. This phenomenon arose from the perception that many small contributions linked together are more relevant and effective than a large contribution of a single person.

In normally, the process of CrowdUM is executed throughout four steps. At the beginning, the seeker describes the project's requirements with the limited time and budget for the challenge and awaits the solutions from the students (solvers). After receiving and resolving challenges, solvers can submit their solutions to the CrowdUM website from anywhere and anytime. In the third step, the seekers can choose the best solutions, as well as they can respond to some solutions to provide precise indications in order to help solvers could improve the quality of their final product. Finally, the best solutions are selected from all solvers in crowd and then the selected solution owner will receive incentives. In which, the major incentive will be a monetary award that given to winners, as well as the social recognition will be served as a lever to their good performances.

5.3 The role of BI-OMIS integrated model in CrowdUM

Following the advantages of crowdsourcing, the CrowdUM was built with the purpose to connect the seekers and solvers together for quickly solving problems. However, with the maturation of project scale and the size of crowd, which increase continuously through years, the CrowdUM system showed the lack of an overall solution in management such as: (1) the autonomous ability to help the crowd could be notified the new project more intelligently based on their hobbies; (2) the demand to adapt the BI-OMIS into the CrowdUM in order to enhance the quality of solutions; (3) and the ability to manage more effectively projects, solutions, documents, and also the business of CrowdUM.

Depend on the crowdsourcing architecture and the previous researches in the BI-OMIS integrated model, which was proposed in the figure 4.5; this section is going to clarify five useful aspects of the BI system, meanwhile, adapting the BI-OMIS model into the crowdsourcing initiative (CrowdUM). As a result, the working efficiency of the BI-OMIS model will enhance the performance of CrowdUM in various areas such as managing solutions, projects, and human being as follows:

- **Using the BI system to support the CrowdUM's OMIS:** this feature enables the CrowdUM to manage the guides, documents, and relevant information inside the CrowdUM; as well as support users to share their knowledge with other members. This helps the crowd members use the past knowledge for solving their new problems. To develop this feature at first, the ETL tool of BI system will extract and load data and document's content from the CrowdUM database into DW at the scheduled time. Following, the DW and OLAP cubes play a significant role in building a multidimensional cube for quickly classifying documents in various dimensions. Finally, the crowd members could find relevant documents via the dynamic interface of the QbE system to select desirable dimensions, which are based on their demands and priorities.

- **Using the Data Mining to classify the crowd:** the main objective of this feature is

to provide the intermediary a better awareness about user's ability. Afterward, this feature gives the CrowdUM an ability to find and map the right members in crowd (solvers) with specific challenges from seeker in an intelligent manner. In technical aspect, we use the Data Mining to classify the user's groups, who have maximum similarity based on their personal information and their usage behaviors. The DM classification is going to use the cluster method (e.g., K-means, ANN/SOM) to isolate groups, after that the suitable jobs/tasks will be sent to the desirable groups automatically. Hence, this method is more flexible and active than the traditional method such as email subscription, or using the method that sends the same challenge to participants in all categories, despite of their expectation, as a spam message.

- **Supporting customers to manage the solutions**: in order to enable the seeker easier to find the most feasible solutions, and also ignores the unfitting solutions. The BI-OMIS supports the decision support system to refine the large amount of solutions from the crowd, which includes two methods user's ranking and comment classification. In which, the user's ranking focuses on the number of votes from crowd members at each solution. In contrast, the comment classification is developed from the DM classify functions (e.g., J45, Naïve Bayes), that classifies comments in each solution into two classes as good assessment and bad assessment. Consequently, seekers could quickly recognize the value of the solution based on its good and bad rating. By this way, the DSS helps customers reduce the cost and time, as well as improve the quality in finding a suitable solution to their problem.

- **Putting in place effective incentive mechanism:** this feature is going to use the Dashboard component of BI-OMIS model to update the list of best crowd members in each category, who have high contribution. The incentive mechanism will promote the excellent members to attend further projects, as well as, create a motivation for new members to improve their ability to achieve this level.

- **Supporting web platform:** in the purpose to manage a large scale of crowd members from various places and time, the BI-OMIS supports the web-based platform for all the BI components. Therefore, the final BI-OMIS system could be easily adapted to the CrowdUM and facilitate the interaction between the participants more effectively.

5.4 The BI-OMIS key features proposal for CrowdUM

According to the BI-OMIS integrated model (see the figure 4.5) and the significant role of the BI-OMIS in the CrowdUM, in this section, I am going to clarify the detail components of the CrowdUM and the SpagoBI open source suite, which will be implemented for validating the proposed model in a real context of CrowdUM.

In technical view, because the CrowdUM is developed from the Joomla framework as a modular and extensible Content Management System (CMS), it plays a main role in managing the contents, news, and users. Moreover, the CrowdUM was built with extensions to manage the challenges and solutions for the crowdsourcing system such as creating a connection between the CrowdUM internal team (intermediary), participants (crowd), and commercial enterprises (requester). Depend on the existed CrowdUM website, my dissertation is going to reuse this framework to develop new components, which are embedded in the CrowdUM website to supply necessary data for the OMIS, and to display the results generated from the BI-OMIS system.

Besides the implementation extensions for CrowdUM, the BI system is another important component that supports this dissertation to complete the proposed model. In recent times, there are many leading BI suits that allow us to build up an overall BI system for various purposes. However, although the most common BI suites are provided with the expensive license such as SAP, IBM, SAS, Oracle, and Microsoft; there are still some leading open source BI suites for small and medium companies such as the JasperSoft, Pentaho, and SpagoBI, etc.

Figure 5.3: The main components of SpagoBI

After reviewing the advantages and disadvantages of these leading open source BI suites, these BI tools also offer more or less a full range of BI components like ETL, DW, DM, ad-hoc analysis, report tools, etc. In general, the Pentaho has a wider community of users, because it is more advantageous and simpler for integrating and modifying than the others. However, the Jaspersoft and the Pentaho are commercial open sources, which have the limited time of using. In turn, the SpagoBI is a non-commercial open source BI platform, which is sponsored by the Engineering Group (one of the Italy's leading IT consultancies). This tool shows the best capability to experiment with the CrowdUM, and it includes an overall set of separate BI tools as presented in the figure 5.3. Finally, this dissertation is going to use the SpagoBI suite in order to experience the role of BI in OM support.

As a result, the table 5.1 clarified eleven features of the BI system in two separate views of the CrowdUM and the BI-OMIS proposed model. To this end, these features will be used to complete implement and evaluate the project in the following chapter.

Table 5.1: The BI-OMIS workable proposal for CrowdUM.

Functions	CrowdUM	BI-OMIS
1. Web crawler (Acquisition layer)	Developing a new extension for CrowdUM (in Joomla version 1.5) to support users to write comments about the tasks and projects of the CrowdUM, as well as proposes feedback to improve the performance, processes, and rules of the organization.	Inside: Capturing these comments and feedback from CrowdUM website to store in the BI-OMIS. Then the DM will classify the main topic and category of them by using the text classify function of Weka's library. In addition, the tool is going to send the current issues to the groups or persons who are in charge (have responsibility in these issues). The main technical here is to develop a small Java project using the Weka text classify function such as Naïve Bayes Multinomial (NBM) to deliver the most appropriate information to the right users via their emails. Outside: Creating a connection with social network such as Facebook, LinkedIn; or from jobseekers website as Adecco, Bolsa Emprego, or Egor. Then we can extract the profiles of available companies and employees to enlarge the size of the crowd.

Continued on next page

Table 5.1 – *continued from previous page*

Functions	CrowdUM	BI-OMIS
2. ETL (Acquisition layer)		- Using the ETL tool to collect customers and crowd members' data, information about tasks, problems, and solutions that stored in CrowdUM DB (MySQL). And then reorganize them in the fact tables of DW for further reusing by the OLAP, Reports engine, etc. This feature is supported by the Talend component in SpagoBI suite.
3. Data Warehouse (Retention layer)		- Designing the data warehouse in MySQL that is connected with the ETL tool (Talend). The structure of DW will be defined to compatible with the OLAP schema (fact tables) that will be indicated in the following chapter.
4. OLAP (Maintenance layer)		Managing the tasks and documents in the OLAP cube with various dimensions such as time, geography, products and services, user roles, and clients. In technical view, I will use the Mondrian tool to deploy the OLAP server, which runs on the Tomcat server (Java framework). It enables users to interact and analyze a large datasets stored in DW through the special query language called Multidimensional Expressions (MDX). In which, the MDX is supported by the Palo component of SpagoBI suite

Continued on next page

Table 5.1 – *continued from previous page*

Functions	CrowdUM	BI-OMIS
5. Ranking (Maintenance layer)	Developing a new component in CrowdUM that allows crowd could vote, as well as writing comments about their ideal for each solution in different projects. This helps our customers find the perfect solution for their requirements.	Using this information to support the DSS to re-order the most useful solutions by its ranking. Moreover, the BI-OMIS will use the DM to classify these comments into good or bad class, and then return these two indicators for DSS purpose such as the like indicator and the comment indicator.
6. Search engine (Maintenance layer)		As the main service of the BI system, the search engine provides a full text search for the CrowdUM's OMIS. In which, users can quickly find out the suitable documents via specific dimensions of the OLAP cube. In SpagoBI, the search engine called 'Query by Example' (QbE) that can support user to create, modify the particular search and then save it for later reusing or sharing.
7. Data Mining (Maintenance layer)		Integrating the Weka's library in the BI as a Data Mining tool to support the CrowdUM in removing the irrelevant solutions (spam), classifying appropriate information for each group/member in an organization based on the DM classify functions such as ANN and SVM. This feature is supported by the Weka component in SpagoBI suite.

Continued on next page

Table 5.1 – *continued from previous page*

Functions	CrowdUM	BI-OMIS
8. Reports (Maintenance layer)		Using the report engine to support the managers to generate various report types such as making reports to summarize the current status of tasks, job schedule, information of crowd members, or the financial outcomes of the organization. This feature is supported by the JasperReport and BIRT components in SpagoBI suite.
9. Search form (Retrieval layer)		This is the dynamic form of QbE, that supports users query the OMIS by simply creating the search conditions following their purposes. The queries could be saved in order to reuse in the future or share with other users. This feature is supported by the QbE component in SpagoBI suite.
10. Spreadsheet (Retrieval layer)		Supporting users to interact with OLAP cube follows the defined dimensions in the web-based, which is supported by the BI tool. This also includes the drill-down operations for obtaining a data examination at various detail levels and from various perspectives. The main technique here is to use web-based (Html and JavaScript) that can connect to the OLAP server for presenting the dynamic spreadsheet as included in the Palo OLAP component in SpagoBI suite.

Continued on next page

Table 5.1 – *continued from previous page*

Functions	CrowdUM	BI-OMIS
11. Dashboard (Retrieval layer)		Presenting the KPIs about the current projects/tasks status in CrowdUM such as: - Indicating the power of crowd through the number of members and projects by category; or the chart about the number of solutions in each project by time; - Presenting the most popular projects/tasks that are running recently (rely on the number of responses or solutions from participants in each project and the total fund of it); - Indicating the ranking of users to promote them over the community based on their best responses. - This feature is supported by the Dashboard component in SpagoBI suite.

Chapter 6

Implementing the BI-OMIS in CrowdUM

6.1 The BI-OMIS workflow model in CrowdUM

In practical view, the workable proposal of the BI-OMIS integrated model is modeled and represented in the figure 6.1. It illustrated processes of the CrowdUM throughout seven steps that enable participants to interact with the BI-OMIS system.

The first step intends to initiate crowd members, this step focuses on finding the new members for crowd by using the Web Crawler Tools (e.g., Facebook SDK, LinkedIn SDK). The tool will connect to user's account and collect their friends list and private information of user in the leading social networks such as Facebook or LinkedIn via their permission in social networks. In order to expand the crowd through friends of CrowdUM's members, the tool gives crowd an ability to invite their friends to join the CrowdUM site just only simple steps. By this way, the size of crowd will be enlarged time by time.

In the second step, the requester will register a project to database of CrowdUM (Joomla framework), as well as submit relevant documents to make sense the project's

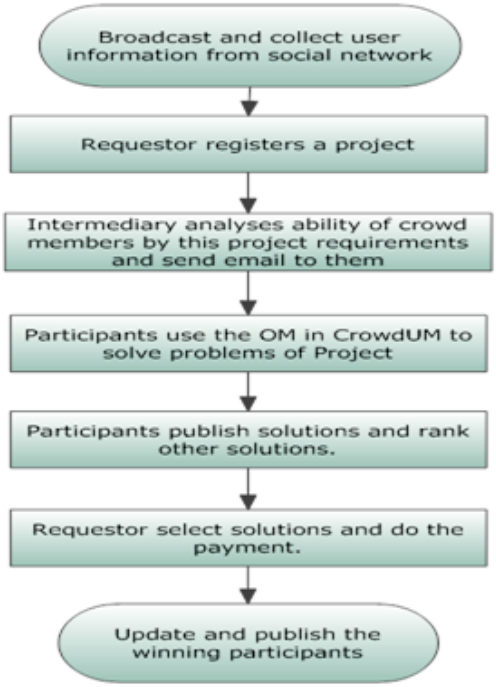


Figure 6.1: The BI-OMIS workflow model in CrowdUM

guidance and requirement. Then the ETL tool will capture the changing of database in CrowdUM and do the extract, transfer, and load project data to DW that relies on the star schema of BI-OMIS model as presented in the figure 6.5.

The third step is to find appropriate members in CrowdUM to solve the new projects by using the DM methods of BI-OMIS system. Hence, in this step I will use the DM classify functions to support the intermediary (CrowdUM managers) in analyzing the ability of crowd members based on the project requirement, and then the intermediary will send an email to particular groups in CrowdUM to notice them about the new projects. In order to assign special projects automatically to relevant groups of CrowdUM, the BI-OMIS system uses the text mining technique for automatic text categorization, which is depended on the WEKA text classify functions such as SMO, NaiveBayesMultinomial, Simple KMean.

In following step, another component of the BI-OMIS system is the OLAP, which supports CrowdUM to manage the OMIS more effectively throughout the multidimen-

sional views. With the support of the OMIS, crowd members could quickly query the guidance, comments, answers for their own problems from the past projects' information that are stored in the OLAP Document Fact Table. Moreover, the BI-OMIS also support the QbE tool to allow participants quickly create their specific query statement (MDX) to find out special information (e.g., documents, projects, problems, solutions) in the DW. In turn, the participants can send a new problem, give a ranking, or propose a solution for special problem in the project via the CrowdUM's components (Joomla site). Finally, this information will be transferred to the DW of the BI-OMIS system through the ETL tool at definitive time of the schedule plan.

After completing the project, participants are able to submit their solutions throughout the CrowdUM's component (Joomla site). Consequently, the crowd members can propose their ranking and comments to the other solutions. This data is going to support the seekers to have a better decision while selecting the best solutions via the DSS feature of BI-OMIS system.

Finally, based on the selections of seekers, the winners will be noticed and received the payments hereafter. Besides, the new information and level about the winners will be updated and published into the Dashboard component of BI-OMIS system, which will promote their best effort in further projects. Consequently, the information about current project such as problems, solutions, comments, and documents - generated throughout the working duration - will be stored in DW to enhance the CrowdUM's OMIS in future.

6.2 The social network interacts with CrowdUM's users

Communication is an important demand of every people, it helps us share information, build a better relationship, and also connect with more new friends. Nowadays, with the popularity of internet all over the world, the social network is becoming a daily demand of many peoples because this is an easy method to connect people with their friends via the visual social channel for communication. By this way, we can share information to our

friends quickly and dynamically than the traditional methods such as letter, phone, or direct meeting. For instance, with the maturity of common social networks like Facebook, Twinter, and LinkedIn; we can not only send the simple text messages, but also include the multimedia messages like musics, images, and videos. Moreover, another useful of social network is to provide the high interaction support between users notwithstanding the limit of their time and space. Hence, the social network gradually becomes an important service, and also a huge storage which keeping many useful information about users' private such as personal information, hobbies, and their recent activities.

In another hand, because the CrowdUM activities are performed mostly depend on the significant role of inside crowd members, who can help the crowdsourcing initiative working model to increase the quantity and quality of solutions. Thus, it exhibits a high requisition to strengthen the number of crowd users (e.g., solvers, seekers) throughout the finding and connecting more and more participants all over the world. Based on this demand of CrowdUM and the huge user resource from the social network, this section is going to create a social network application to help CrowdUM collect users' information, introduce the CrowdUM for newbies, and share this application to their friends on social network via various methods (e.g., posting on user's wall, sending message, sending invitation), and finally encourage them to register with CrowdUM organization.

As an experiment, this application just focuses on the Facebook - the most famous social network over the world in recent years. In the traditional way, we can use some web crawler tools such as Nutch[1] or Crawler4j[2] to help us crawling over the Facebook website and retrieving user's information. However, this method is abandoned by the Facebook regulation[3], so that the only way to interact with user's information is to develop an application based on the Facebook SDK, which will be adapted into the CrowdUM's Facebook page as a component called 'Following CrowdUM' app. As a result, the Following CrowdUM application supports CrowdUM's users in four main functions as

[1]http://nutch.apache.org
[2]http://code.google.com/p/crawler4j
[3]https://www.facebook.com/legal/terms

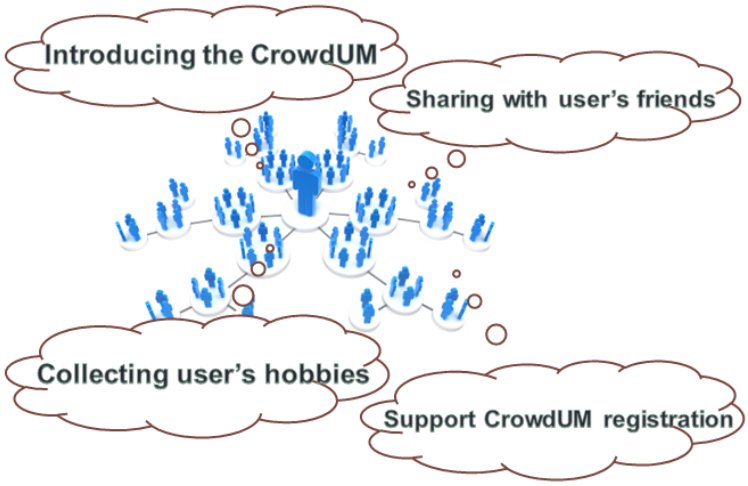


Figure 6.2: The main Following CrowdUM functions

indicated in the figure 6.2.

- **Sharing with user's friends:** this part is included in the top of this application. The main objective intends to the spreading features that are supported by the Facebook SDK. The first method 'Post to your wall' is going to publish a new message to current user's wall that will display the welcome message with the link to this application, then all their friends will be noticed a new welcome message for inviting them to use this application. Similarly, the second method 'Send Message to your friends' is going to post a new message to specific user's friends without publishing to all other friends in order to avoid disturbing the others. The final method 'Send Requests to your friends' allows user to choose several or all friends in current friends list and give them an invitation to use this application.

- **Collecting user's hobbies:** This is an important feature that helps the BI-OMIS system crawling user's information (e.g., user's friends, user's private hobbies, friends who used this application). These hobbies then will be mapped through the defined table, which can reflect the user's like to the specific hobbies as indicated in the figure 6.3. Then this information is going to store in the user database of CrowdUM following the register process supported by the CrowdUM website. It is important

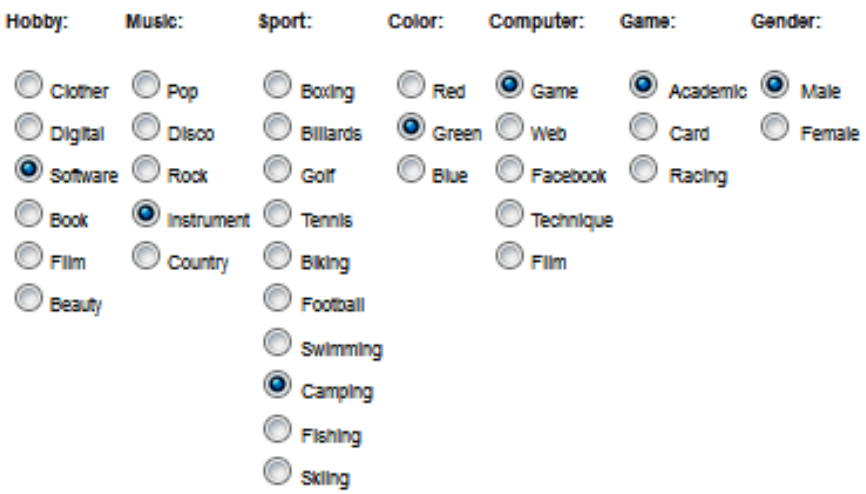


Figure 6.3: Mapping hobbies with user's information

because this information will be used to classify the group/category of users automatically based on their hobbies and the DM algorithms in the following sections.

- **Support register CrowdUM:** Based on the user's information, which is collected from the previous step, this function allows user to quickly redirect to the CrowdUM registration process with the collected hobbies from a social network (Facebook). In this process, the user can refine the hobbies before submitting the form and become a member of CrowdUM.

- **Introducing the CrowdUM information:** this section is displayed at the bottom of the application with four significant contents to help newbies learn more about CrowdUM services such as CrowdUM definition, Rule of the CrowdUM site; the link of CrowdUM website; the quickly update link will show user the newest projects that are running in CrowdUM; and the link of CrowdUM Facebook page where user could communicate with CrowdUM's members.

In summary, the interface of the 'Following CrowdUM' Facebook application is presented in the figure 6.4. The figure gives us a general view of this application that is enclosed with four main functions as introduced in this section.

Figure 6.4: The Following CrowdUM Facebook application.

6.3 The ETL and OLAP functions in BI-OMIS system

To help the BI-OMIS system saving more time and to enhance the effect in managing the data warehouse, the ETL tools are developed in order to simplify the frequently jobs of database administrators such as connecting to various kinds of databases (e.g., MSSQL, MySQL, Oracle), synchronization, transformation, and loading data from several existed databases into new data warehouse for special purposes without requiring the adminis-

trators any insight knowledge about the programing languages. Indeed, there are many available ETL tools that can support these fundamental functions of ETL, for instances: Information Server Infosphere platform from IBM corporation, SQL Server Integration Services of Microsoft, SAP Data Integrator, Talend Integration Suite, SAS, Sybase, etc. Although all most ETL tools are commercial, the Talend Integration Suite is one of the open source tools that built with many powerful features for the ETL process. Moreover, the Talend tool is supported by some well-known open source BI suites (e.g., JasperSoft[4], Pentaho[5], and SpagoBI[6]). In order to develop the ETL component to work with the SpagoBI suite, in this dissertation, I decide to use the Talend tool to experience the ETL component in the BI-OMIS system.

6.3.1 ETL structure

To build the data warehouse for the BI-OMIS system, I am going to use the Talend Integration Suite to extract, transform, and load data from the CrowdUM database into the business fact table and document fact table as indicated in the figure 6.5.

Following this diagram, the BI-OMIS data warehouse is built up from eight dimensions, which holding different roles such as: (1) The Geography dimension contains the location of projects and documents in the CrowdUM, in which the Country and City fields play a role of level-based dimensions; (2) The Project dimension contains project information, in which the Group and Type fields play a role of level-based dimensions and the other fields are the value-based dimensions; (3) The Date dimension shows time of project's orders and documents in the system; (4) The Customer dimension contains user's information on the type of value-based dimension; (5) The Problem dimension deals with questions and requirements from users; (6) The Solution dimension indicates rank and content of solutions in each project; (7) The Document dimension contains the

[4]http://www.jaspersoft.com

[5]http://www.pentaho.com

[6]http://www.spagoworld.org/xwiki/bin/view/SpagoBI

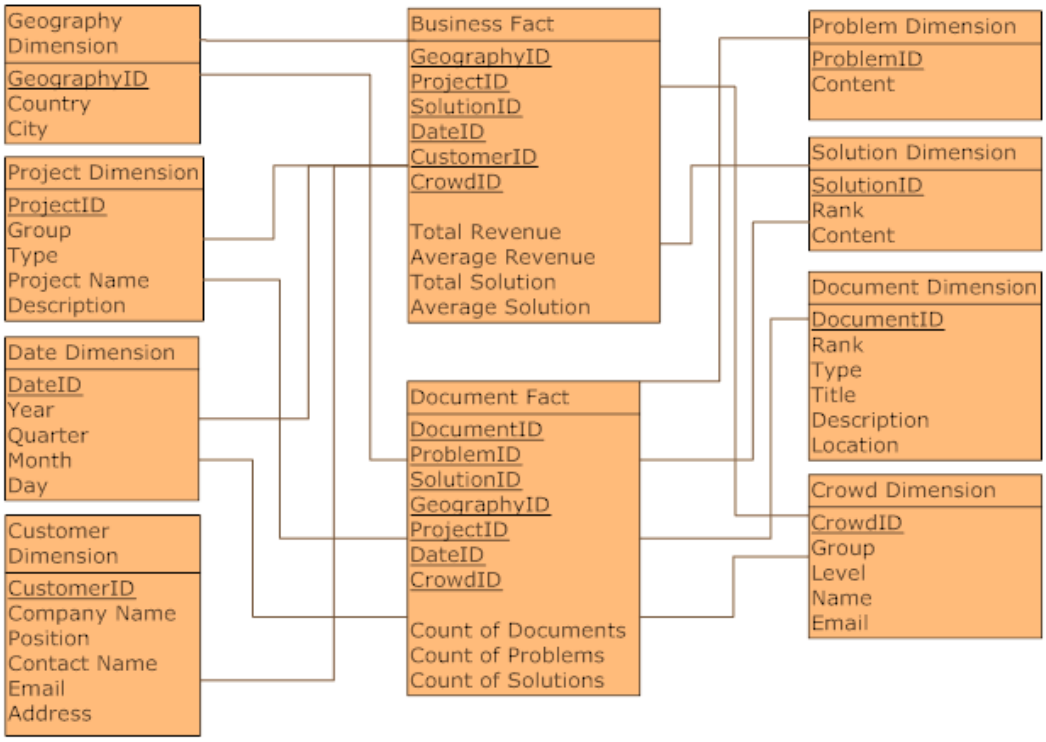


Figure 6.5: The fact tables of BI-OMIS system.

detail information of documents that support solvers to correctly execute projects; (8) The last dimension contains the user information such as their level, group, and contact to help the BI-OMIS system classify the ability of each CrowdUM member.

With two fact tables and eight dimensions, the ETL system plays the leading role to collect special data, and build the BI-OMIS data warehouse based on the CrowdUM database as illustrated in the figure 6.6. Moreover, the Talend tool supports various advanced ETL functions to interact with database system such as connecting to various kinds of databases; and support many features for joining tables, filtering redundant data, string manipulations, etc. Throughout the ETL process of BI-OMIS, the tMap object is used frequently to extract data from each CrowdUM table then allows us to transform and load data into the BI-OMIS's dimensions and fact tables.

Finally, to help the current ETL model integrate with the SpagoBI system, the Talend suite also developed a special feature that allows us to deploy the ETL processes directly

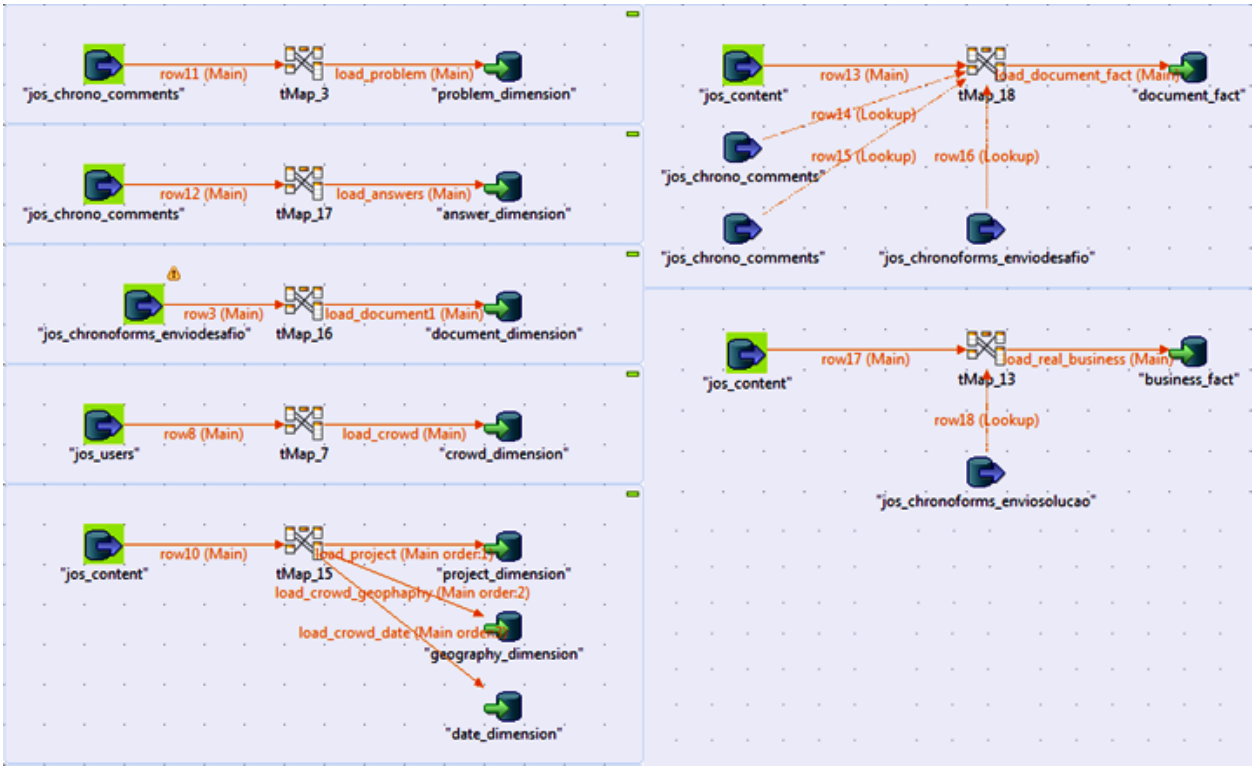


Figure 6.6: The ETL diagram for the BI-OMIS system

into the SpagoBI suite. This deployment feature will automatically create and deploy a final job script (meta-data file) relies on the SpagoBI Talend Engine specifications. So that, the SpagoBI managers can schedule and execute the ETL process model that exported from the Talend open source suite.

6.3.2 OLAP cube diagram

To manage the BI-OMIS data warehouse, the OLAP cubes is a common technique that used to explore data cube created from the multidimensional views of data warehouse. This allows analysts to navigate, define, and analyze organizational data organized in various dimensions. Hence, the OLAP is a useful tool to design, publish, and query multidimensional cube quickly and visually.

As a component in the SpagoBI suite, the Mondrian OLAP server is included in the

back-end of SpagoBI that allows business managers to real-time analyze the data warehouse via the OLAP cube. In the front-end, the JPalo component is a presentation layer in web based platform to communicate with the Mondrian OLAP server and support users to perform with OLAP cube throughout drill down/up and cross-tabulating information with many complicated analytical queries. In summarize, the figure 6.7 presented main components for building the BI-OMIS OLAP cube, this action diagram also shows the working model between the JPalo and Mondrian in SpagoBI environment.

In this diagram, the BI-OMIS data warehouse plays the leading role as storage of fact tables and other dimensions built by the ETL (Talend) system. Following, the SpagoBI server is a web server deployed on the Apache Tomcat server. It supports the Mondrian ROLAP engine and allows analysts to perform with the business cube and document cube via the MDX queries. In which, the MDX is a multidimensional query language as an exchange standard XML/SOAP for supporting the Web Service to communicate with the Mondrian ROLAP Engine through the HTTP protocol.

On the contrary, the BI-OMIS cube schema is modeled by the Mondrian Schema

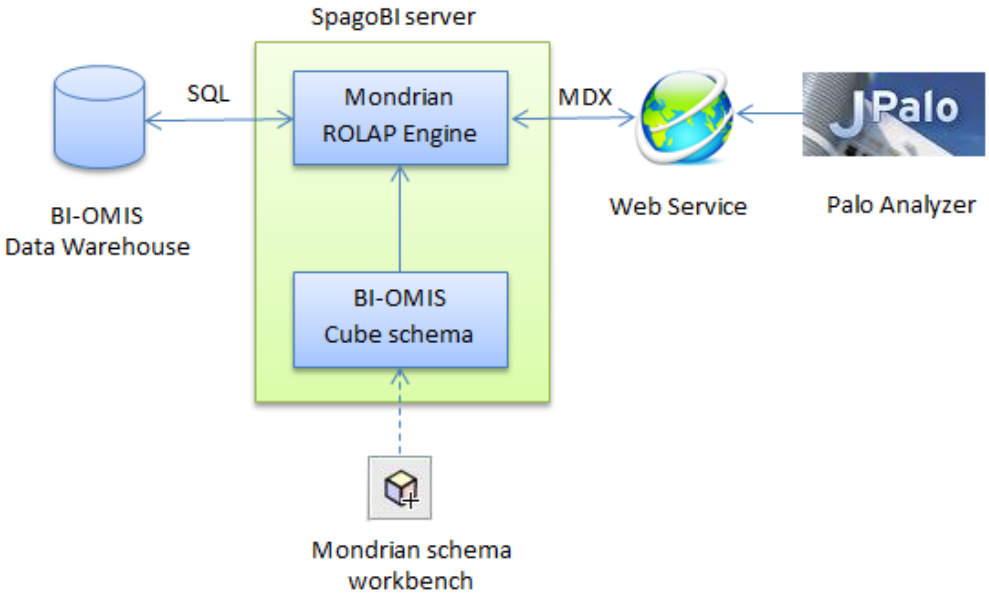


Figure 6.7: The working model between JPalo and Mondrian server

Workbench and then this defined schema will be used to build the OLAP cube in the Mondrian ROLAP Engine. The Mondrian Schema Workbench is an open source tool for modeling and publishing special star schema OLAP cubes according to the Mondrian defined patterns. Besides, in business view, the CrowdUM users can access the BI-OMIS OLAP cubes by interacting with the SpagoBI web service that could run on common web browsers for instances: IE, Firefox, Chrome, etc. Consequently, the results returned from the Mondiran engine will be managed by the JPalo system, which included many advanced analysis features such as: navigation tools (e.g., drill down/up, slice, dice), filter layers, computation methods (e.g., sum, count, avg), and present multidimensional views on spread-sheets, graphs, and reports.

Finally, the figure 6.8 illustrated the user interface of JPalo OLAP cube, in which the user performed the drill down function on the OLAP document cube. Following this concrete case, the OLAP cube was built up with Geographic dimension and Document dimension in the left axis. In contrast, the measurements were calculated and showed in the other axis with three values as Document Count, Problem Count, and Answer Count. Moreover, the user could also choose other dimensions and drop into the left and top axes

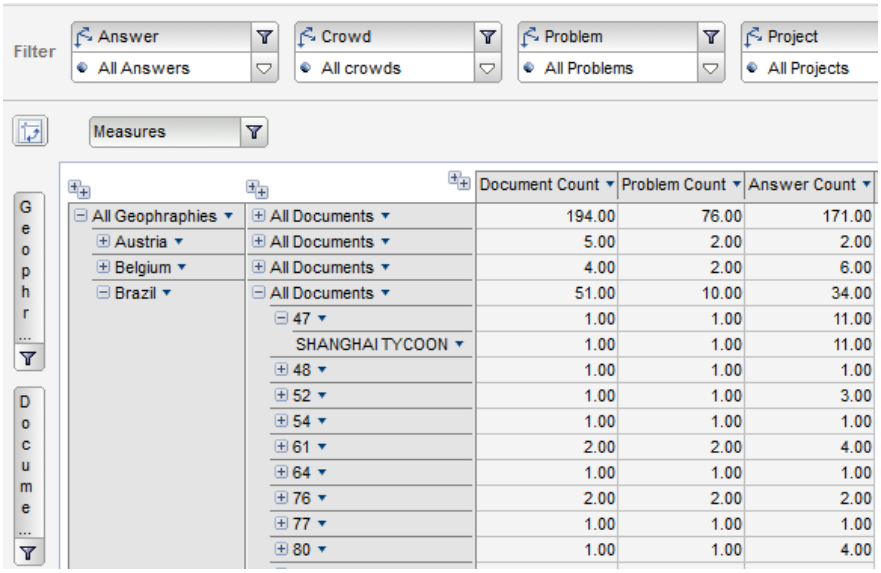


Figure 6.8: The OLAP cube in action

to modify the result by their own demands.

6.4 The BI-OMIS in supporting CrowdUM's OMIS

As mentioned in chapter 4, the OMIS is the main component to help CrowdUM in managing the information, and formal knowledge. The information in the CrowdUM usually exists in various forms of document such as pdf, word, text, or comment which are submitted by the seekers, and also from the crowd members. In which, these problems and projects are divided into six categories for instances: website, design, logo, translation, marketing, and other. Hence, this section will extend the role of the BI-OMIS system in supporting the CrowdUM's OMIS throughout four main features as follows.

6.4.1 OLAP and QbE in managing documents of the CrowdUM's OMIS

In the BI-OMIS system, while the data warehouse and OLAP cube are keeping significant positions in creating and establishing the multidimensional view to manage the business and documents fact tables, the Query-by-Example focuses on facilitating the user performance in most inquiry tasks, which is also known as the QbE in abbreviation. By its meaning, the QbE is another kind of the traditional query language such as SQL, however it supports user to perform the query in the graphical interface that contains dynamic elements and fields. In addition, the QbE allows user to interact visually with OLAP multidimensional data by defining relevant condition fields, then users can select possible values from the dynamic form of their query. By this way, the QbE proposes crowd members an easy method to retrieve suitable documents from the multidimensional views without limiting users in a few fixed forms, as well as it does not require user have an insight knowledge about the SQL language at all.

As a component in the SpagoBI open suite, the QbE integrates closely with the document OLAP Cube. Hence, it allows user to quickly build the dynamic document search

form based on the QbE's features. Consequently, crowd members can query relevant documents in an intuitive way through a set of filters as indicated in the table 6.1. In which, the 'Field' column presents field names that are used to retrieve relevant documents; the 'Dimension' column indicates the location of correlative fields in OLAP Cube; and also, the 'Type' and 'Filters' present the method and specific filter elements that concern with the 'Field' column. By this way, each field is classified into three types: the 'Data' indicates that this field is used just to display information; the 'Dynamic' shows that the elements in this field will be loaded dynamically from the data warehouse at the running time; and the 'Static' indicates static elements that are clarified in 'Filters' column.

Table 6.1: The QbE filters to retrieve CrowdUM documents

Field	Dimension	Type	Filters
Document Title	Document	Data	Display Document Title
Document Rank	Document	Dynamic	Allow user to specify the limit docs rank (ex: filter docs that have the ranking higher than '7')
Document Location	Document	Data	Display the direct link to open the document
Country	Geography	Dynamic	Load from Country field
City	Geography	Dynamic	Load from City field
Year	Date	Dynamic	Load from Year field
Month	Date	Data	Display the Month of year
Crowd Name	Crowd	Dynamic	Load from Crowd Name field
Project Type	Project	Static	'Finish', 'Fail', 'Running'
Project Name	Project	Data	Display the Project Name
Project Status	Project	Static	'Urgent', 'Normal'
Category Name	Project	Static	'Logo', 'Web site', 'Translation', 'Design', 'Marketing', 'Other'

In addition to support crowd members to comprehend the past context of documents, user could add more fields that existed in the document fact table such as specific problems, questions, and answers by the order of time dimension and project dimension. Throughout retrieving this information, users can understand more about the scenario of these documents at the created time, so that they can better recover the past knowledge and experiments while using the CrowdUM's OMIS. Moreover, the QbE component gives users an ability to save the query result in order to share with other crowd members via the SpagoBI suite. Thus, these supports from the QbE component have proposed the most significant role of BI-OMIS that allows participants to use the OMIS more effectively via the supports from the BI system.

As a result, the figure 6.9 represents an example by query documents with the Document Type is 'Urgent', the Category is 'Web sites', and have 'Finish' status as indicated in top of the figure 6.9. Consequently, the results are presented at the bottom of the figure 6.9, which have been grouped by Geography dimension as Country and City; and in each group the list of relevant documents showed on the right side of figure 6.9.

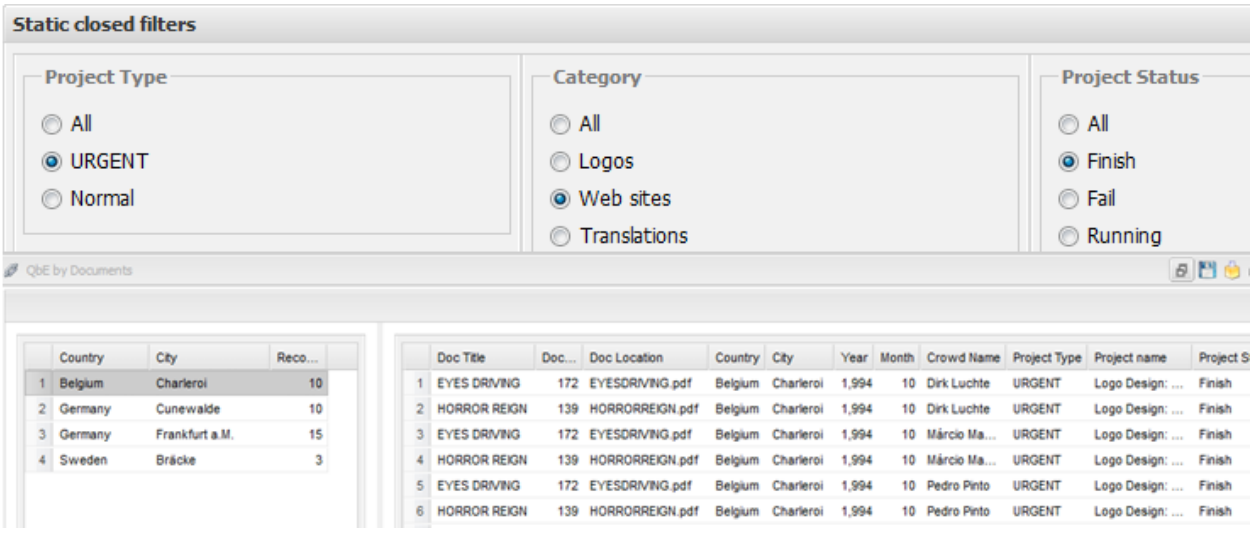


Figure 6.9: The QbE in retrieving the documents

6.4.2 DM in clustering and searching the OMIS

In another aspect of the BI-OMIS system, the Data Mining has an important role in supporting the CrowdUM's OMIS to classify and divide documents into separate groups from a large scale of data. Hence, this section will illustrate the OMIS cluster model in the light of reorganizing the content of all documents in the OMIS into various groups that have the similar content. For that purpose, each cluster will contain a set of relevant key-terms, which represent for the universal meaning of that group. By this method, whenever crowd members want to find out an information in OMIS, they can propose specific key-terms concerning with their problems, consequently, the DM will indicate the most appropriate cluster with documents that have the similar meaning with user's demand.

To implement this feature, I used the clustering method supported by the Weka DM to connect relevant documents together not only by their categories, but also by the probability of each word in their content. As a result, the figure 6.10 below presents the process flow of DM to cluster all documents in the CrowdUM's OMIS.

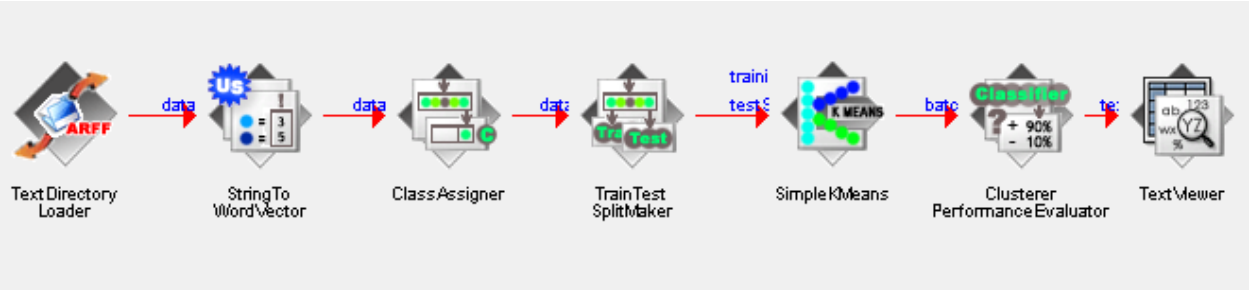


Figure 6.10: The Data Mining process flow for document classifier

Following this process flow, I used the 'TextDirectoryLoader', a function of Weka DM, to load all documents in OMIS and then I converted them to the vector format by using the 'StringToWordVector' function before we can use them to perform the clustering method. However, because the documents usually contain irrelevant or meaningless words (e.g., $, #, number), hence in the converting process I defined the stop-words list

to eliminate these useless words in order to enhance the quality of clustering process. Finally, the 'SimpleKMeans' function was used to classify documents into 8 clusters, which gave a better result after several tests with the current documents in CrowdUM's OMIS.

By focusing on the sequence of exchange messages, the figure 6.11 showed a sequence diagram of the OMIS clustering process in order to describe interactions between the user and these BI-OMIS components. In which, the CrowdUM is the current website

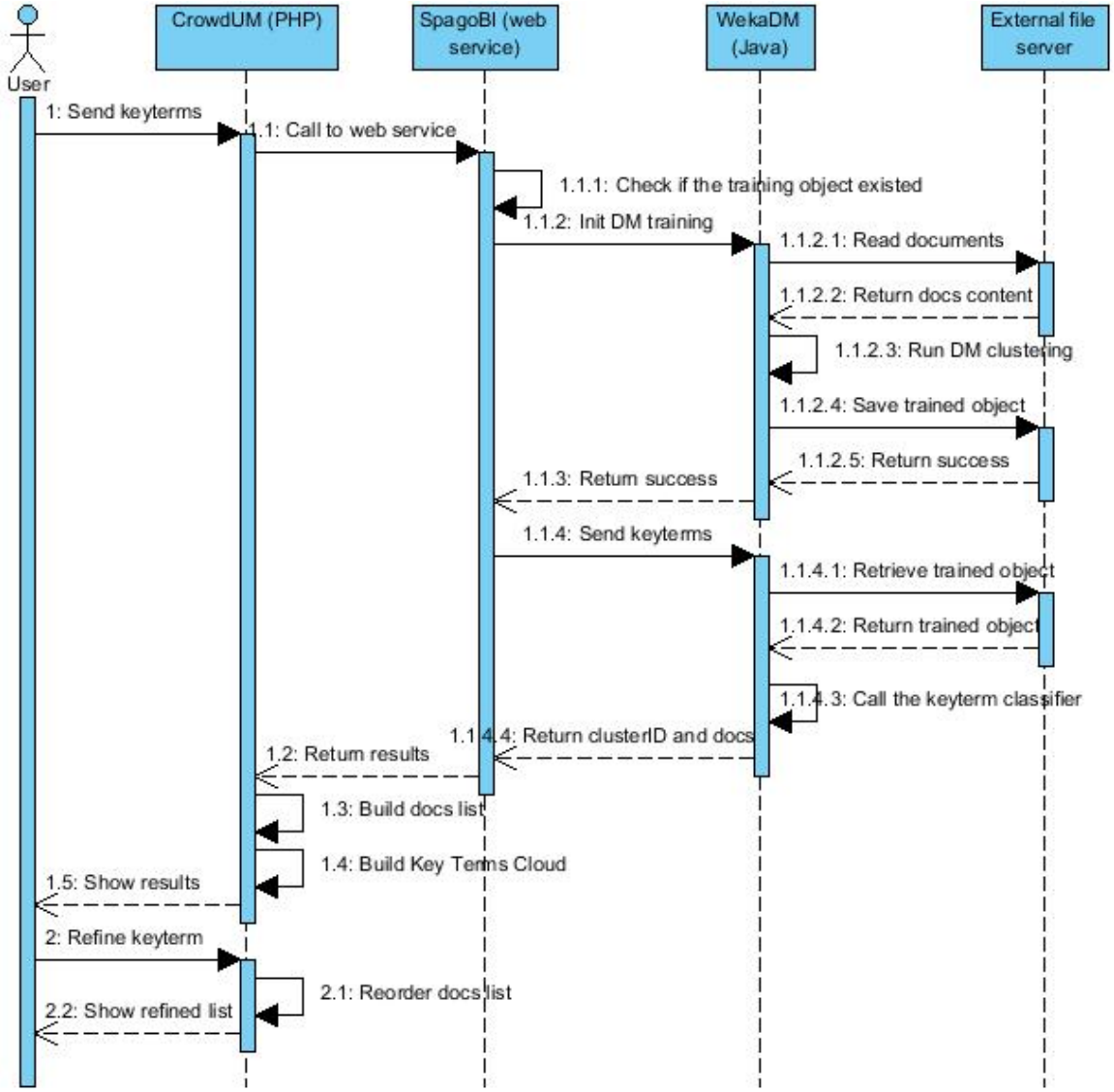


Figure 6.11: The sequence diagram for the OMIS clustering

of CrowdUM, and the SpagoBI web service is an extended component that is supported by the SpagoBI suite. In addition, the WekaDM is a Java application based on the WEKA library for reusing the defined data mining functions.

In this sequence diagram, user can interact with the OMIS through two methods, the first one is to input key-terms to find relevant documents via the OMIS clustering method; the second method is to refine the order of return documents via the Key-terms Cloud as illustrated in the figure 6.12. In this feature, the most popular key-terms will be shown in a kind of dynamic cloud with different color and size, which are depended on its probability of appearance in the current cluster. By this way, whenever user click on each key-term, the results list will be reordered based on its probability in the current documents list. As a result, the OMIS clustering method and the Key-term Cloud will

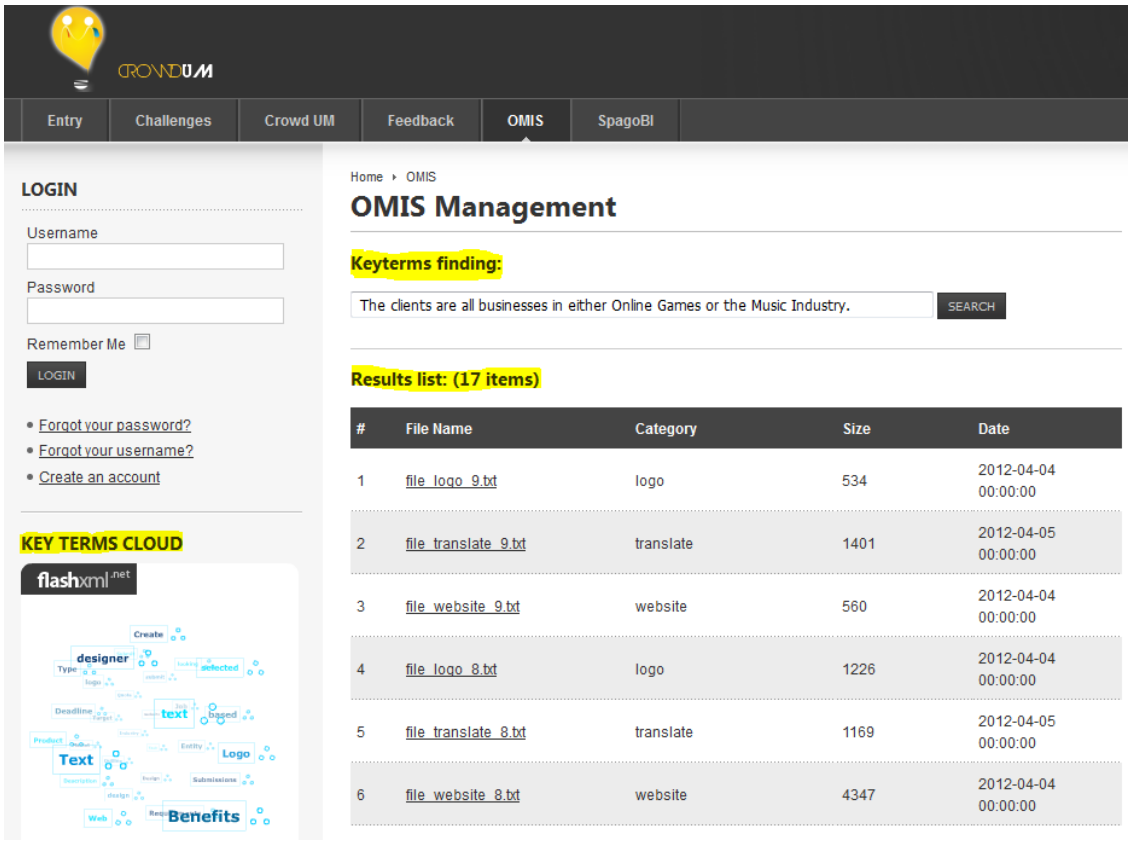


Figure 6.12: The front-end of the BI-OMIS system

enhance the OMIS process of CrowdUM in the friendly manner, as well as support crowd members for finding suitable documents quickly and effectively just through some simple and visual steps.

6.4.3 DM in mapping the project with user's hobbies

As mentioned in the chapter 6.2, the personal information plays an important role in the BI-OMIS system that enables the CrowdUM system to comprehend user's demand more exactly. Based on this information, it helps the BI-OMIS system could spread new projects from customers to the suitable crowd members. Moreover this feature will reduce the irrelevant rate of invitation letter which automatically send to subscribers whenever receive a new project.

In normally, the CrowdUM will send email about the new project to all members of CrowdUM who subscribed into CrowdUM site. However, the CrowdUM's members usually have different objectives and abilities as classified in six categories of CrowdUM, so that it is not proper to send the same projects/jobs to all participants. In some cases, these emails usually disturb other members in different fields/areas, consequently for a long time these emails will be ignored and moved to spam folder. For that reason, the third feature of BI-OMIS system is going to support the CrowdUM's OMIS to create a better dynamic connection between the seeker and solvers based on the project's category and the user's hobbies classifier. The final result focuses on increasing accurate rate of the project notification when sending to CrowdUM's members, which based on the users' hobbies collected from the Facebook's profile.

In practical view, the figure 6.13 presented a general process about the user's hobbies classifier and the requirement mapping tool, which were implemented from the DM classify functions of Weka's library. In which, the Requirement classifier tool is built upon two sub-processes. The first process focuses on collecting the user's hobbies (see chapter 6.2) and then they will be classified into different categories of CrowdUM (e.g., logo,

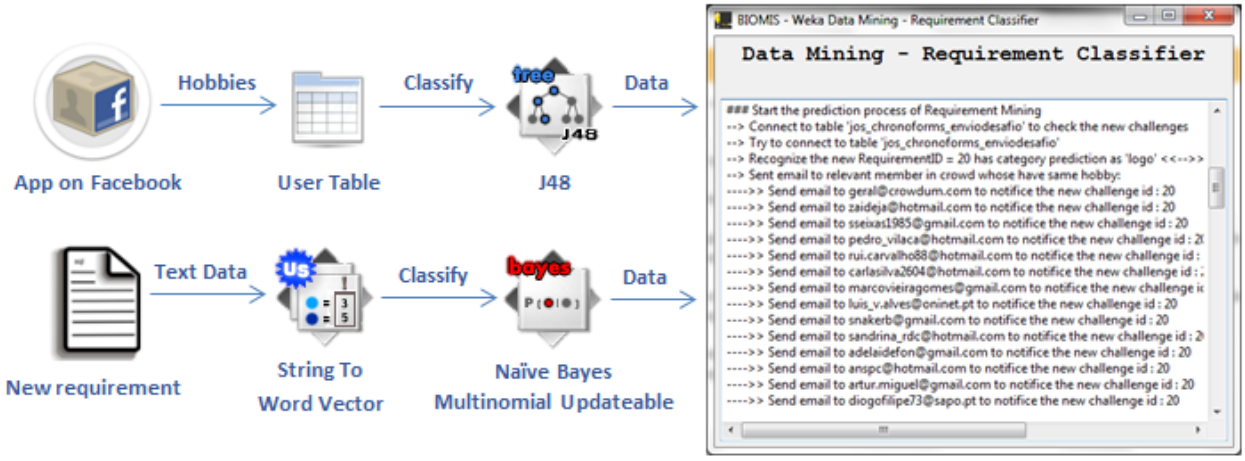


Figure 6.13: The user's hobbies and requirement mapping model

translate, website) by using the J48 algorithm from DM tree classifier. In order to build the classifier, I used the hobbies training set to help the J48 algorithm build the DM classify model as proposed in the table 6.2. In the same way, the second process focuses on the Naive Bayes Multinomial Updateable function to categorize the new requirement into specific category of CrowdUM. Finally, the Requirement Classifier tool, which is represented on the right side of the figure 6.13, will execute these two processes every 60 seconds (this value can be configured by CrowdUM's managers). Afterward, whenever

Table 6.2: The training data to classify user's hobbies from the Facebook's profile

Hobby	Music	Sport	Color	Computer	Game	Gender	Category
film	country	biliards	green	technique	racing	female	translate
clother	rock	fishing	green	web	sport	female	logo
film	rock	boxing	green	facebook	racing	male	logo
clother	pop	golf	blue	facebook	racing	female	translate
film	rock	tennis	red	web	action	male	translate
book	instrument	swimming	red	technique	action	male	logo
digital	country	biking	blue	technique	action	male	logo
beauty	instrument	skiing	blue	film	sport	female	website

the DM tool discovers a new requirement submitted by seekers, this tool will map the new requirement/project with similar solvers by its category and then the Requirement Classifier tool will send email to concerned members to notice them about the new project.

6.4.4 DM in classifying the user's feedback

Despite there are some available methods to deal with the CrowdUM's OMIS such as using the QbE component to query knowledge in the CrowdUM OLAP cubes, as well as using the OMIS clustering method to retrieve similar documents via the DM clustering model. However, with some newbie members, they still need a direct inquiry for special information or difficult issues that are not existed in the OMIS. So that, in this section I am going to propose a direct feedback component that supports the CrowdUM's OMIS to quickly classify the category of feedback, determine the concerning peoples (e.g, supervisors, masters, older workers) to answer the feedback, and then notice these problems directly to them (see the figure 6.14).

Through the feedback classifier component, it allows the OMIS to respond to user's feedback more effectively. In addition, this feature is not only used for the newbie members, but also for crowd members as a backup solution to support them resolve their problems whenever they can't find out any suitable information in the OMIS. Hence, in this method the role of the BI-OMIS system is to classify the messages from the user and then create a direct connection between members with the older workers in CrowdUM, who have an important historical knowledge for decision-making support.

In summary, throughout the OLAP cube and DM, the BI-OMIS system plays a significant role in improving the effectiveness of CrowdUM's OMIS. That is to say, there is such a benefit in adapting the business intelligence in the organizational memory, that is going to support crowd members to solve their problems better and more effective based on the stored information in the past and the support from DM functions.

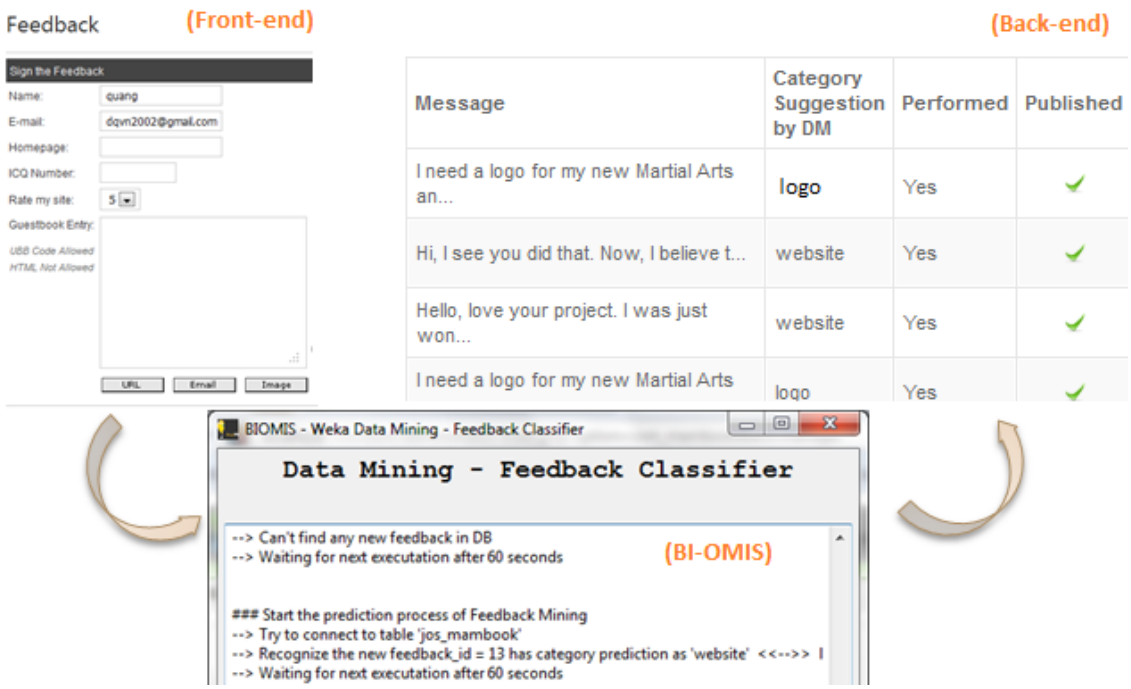


Figure 6.14: The User's Feedback classifier process flow

6.5 The DSS for supporting customer to select solution

On the contrary with the section 6.4 that mainly supports crowd members to solve their own problems by using the CrowdUM's OMIS, this section will intensify on the customer support to help them select the best solution from a large amount of submissions. Because one of the problems inside the crowdsourcing is that the crowd usually produces many unwanted or junk solutions, hence it creates such a difficult problem for seekers to find out and select the best solution from massive solutions. To deal with this problem, the BI-OMIS provides the Ranking and DM features to estimate the solutions as called the design support system (DSS). In thus, the DSS allows seekers easily find an appropriate solution via the voting and comments from crowd as follows.

6.5.1 Data Mining system for solution classifier

Based on the DM system, the first feature of the BI-OMIS will allow the seeker quickly classify and count the rate between the number of 'good comments' and 'bad comments'. In which the good ranking supposes that the current solution is satisfactory to choose, and the bad ranking shows the dissatisfaction with this solution. Then, the seeker can monitor the rate between good/bad comments for each solution, hence seeker can elicit the suitable solutions effectively than get tired reading every comment among a massive solutions in their project.

To implement this feature, I used the DM classifier function to separate the kind of comments into two types: 'good' class and 'bad' class as modeled in the figure 6.15. In the first step, I created a training data set as indicated on the top left of the figure 6.15, in which, there are about one hundred samples representative for the 'good' class and about fifty samples typical for 'bad' class. Consequently, the data are converted to the vector format by the 'String To Word Vector' function of Weka's library. Next, the 'Class Assigner' function will indicate the main class of each message and the 'Train Test Split Maker' will create a training set and test set for classify function. For classify function, I

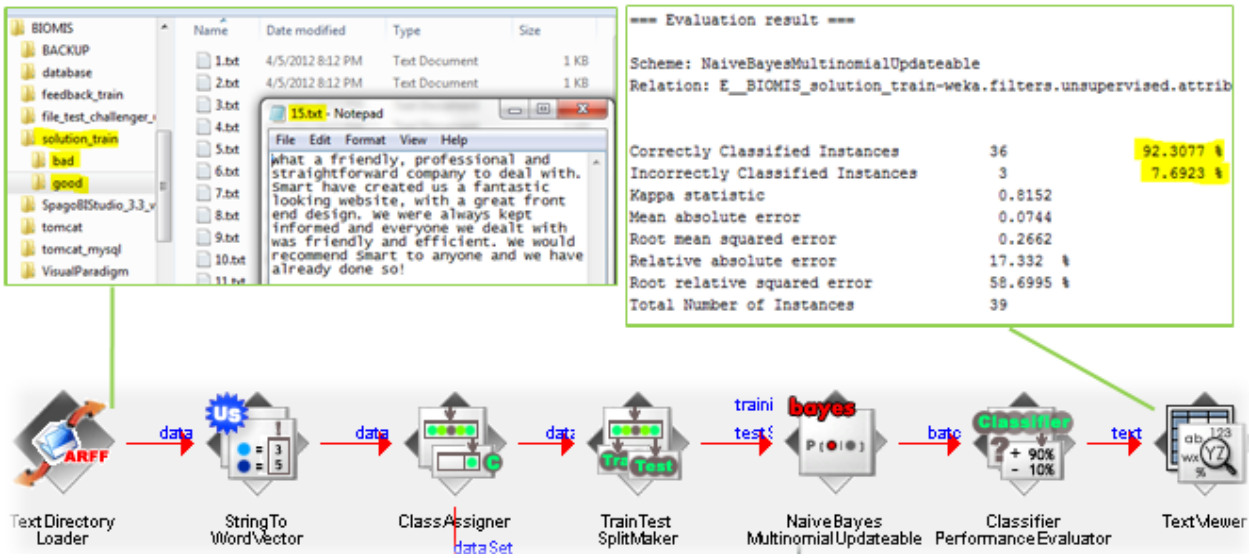


Figure 6.15: The solution's comments classifier.

chose the Naive Bayes Multinomial Updateable function to identify the 'good' and 'bad' messages. As a result, the evaluation result of this classification method is about 92.31 percent for the Correctly Classified Instances and 7.79 percent for the Incorrectly Classified Instances as highlighted on the top right of figure 6.15. Finally, this result indicates a good training performance for the comment classifier function.

6.5.2 Ranking system for solution classifier

While the DM is highly based on the comment classifier as an indirect method to set off the better solutions. In another aspect, the ranking system plays a significant role in the DSS as a direct method for the crowd. By this aspect, the ranking system mainly focuses on the direct votes (likes) from crowd members who reviewed the solution. This method is more direct than the previous method of comments classifier, because the user can decide clearly their opinion about the solution. The ranking allows members to decide which solutions are good enough to get the award. Moreover, to reduce the spam from the crowd such as the multiple votes for their own solution in the order to dominate the rankings of solution, the rating feature limits user by just voting one time per user for each solution.

To illustrate the DSS features in supporting seekers to select solutions more effectively, the figure 6.16 presented the main interface of the Solution management with two DSS indicators as follows:

- **The like indicator:** Rely on the number of likes from crowd members in each project, the Solution management interface will list all solutions with their rating values in the 'Likes' column by decrease order. With this order, the rating indicator allows the seekers quickly review the most prominent solutions at first. Following, the interface also gives users a direct link to review detail information of the selected solution as indicated in the top-left of the figure 6.16.
- **The comment indicator:** The comment classifier is a feature of the BI-OMIS system

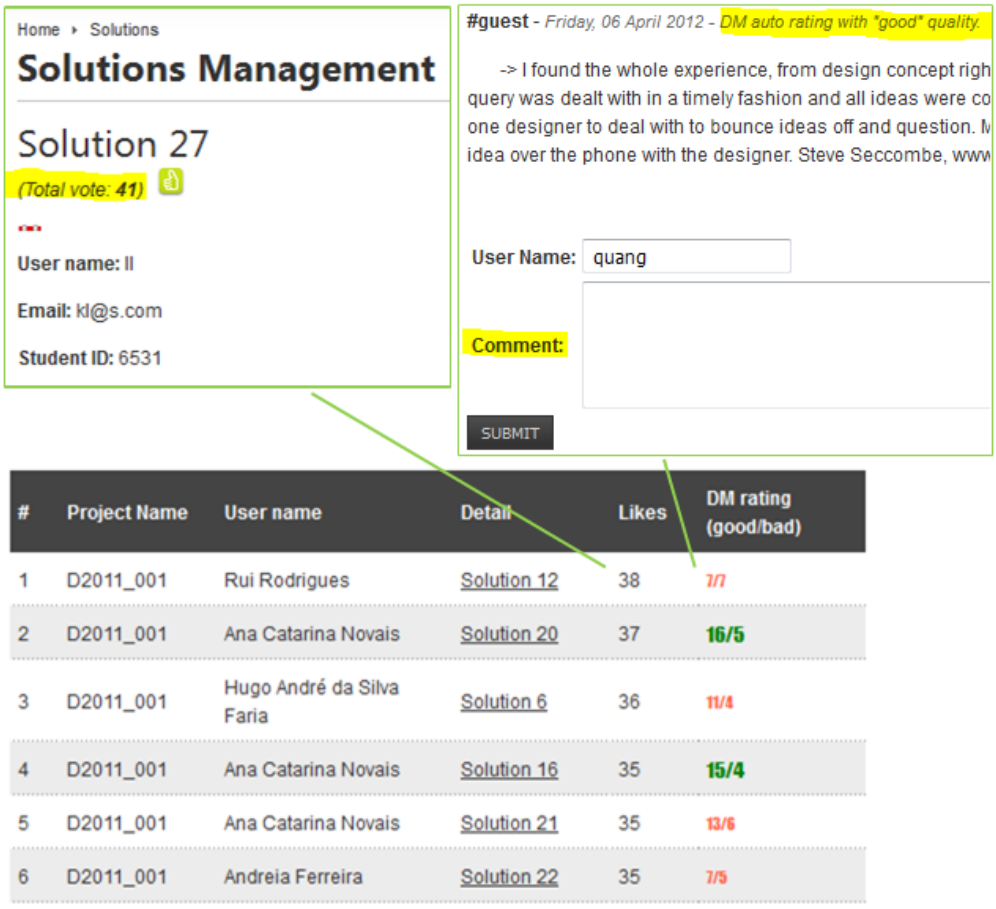


Figure 6.16: Solution management (DSS)

as introduced before. In the Solution management interface, this indicator is presented in the 'DM rating' column which displays the total of good comments and the total of bad comments (separated by the slash mark). Moreover, these indicators not only show the values of good/bad comments, but also indicate the worth of solutions rely on their different rate between the number of good and bad comments. To explain, if the number of good comments is higher than the number of bad comments three times (equal or greater), these values will be shown with the green color and large size; that is to fascinate the seeker's eyes on these prominence solutions. On the contrary, the other solutions will be held in red color and small size for meaning that these solutions are not worth enough to choose.

With both of these indicators, the seekers can inspect a large amount of solutions proposed by crowd quickly and suitably. Indeed, seekers may choose the solutions which have higher voting numbers, higher good/bad rating numbers, or both of them according to their experiments. Anyway, throughout the using of these two indicators in the CrowdUM environment, the strong of DSS in particular and the BI-OMIS system in general has been examined to facilitate the solution management process, as well as enhance the effect of crowdsourcing projects based on the crowd's knowledge proposed in the past (OM).

6.6 The BI system in managing the CrowdUM business

As another powerful resource of the BI-OMIS system, the business performance management, called BPM, keeps an important position in the BI-OMIS system that focuses on the business processes, metrics, reports, etc. By this aspect, the BPM allows the CrowdUM's managers comprehend about the current business environment of CrowdUM throughout measuring, monitoring, and managing business performance. In order that, this section is going to examine the usefulness of BI components in supporting the managers of CrowdUM to enhance decision making, as well as measure key performance indicators (KPIs) such as: creating dynamic financial report; performing evaluation and notification the current status of CrowdUM (e.g., total of revenues, problems, solutions, projects); and enabling managers to manually query the OLAP multidimensional business cube via the QbE component.

6.6.1 Report system

Normally, the report system is a common feature in the BI system by its facilitation in presenting the business information and statistics to CrowdUM's managers throughout various chart types and tables such as bar charts, line charts, and pie charts. In an exam-

ple, the figure 6.17 indicates the number of projects versus with the number of solutions from different countries. At which, the total projects raised by seekers are illustrated by violet color on the front line chart; and the total solutions submitted by crowd members are presented in the blue color of the back line chart. By this chart, CrowdUM managers can observe the current projects and solutions by each country, then they can understand which countries have a low interaction or response rate such as Poland (7 projects; 16 solutions), or Norway (6 projects; 15 solutions). Thus, they can find out new strategies to enhance the low attention in these countries; as well as to give awards/promotions to encourage members in countries which have higher responses such as Germany (122 projects; 328 solutions), and USA (122 projects; 352 solutions).

In the same purpose, the figure 6.18 represents six important reports to help managers to maintain the business performance, and also quickly export these charts into the financial report, which are updated dynamically rely on the execution schedule of ETL component on real data. These reports include: (1) The 'Revenue by Month' report shows

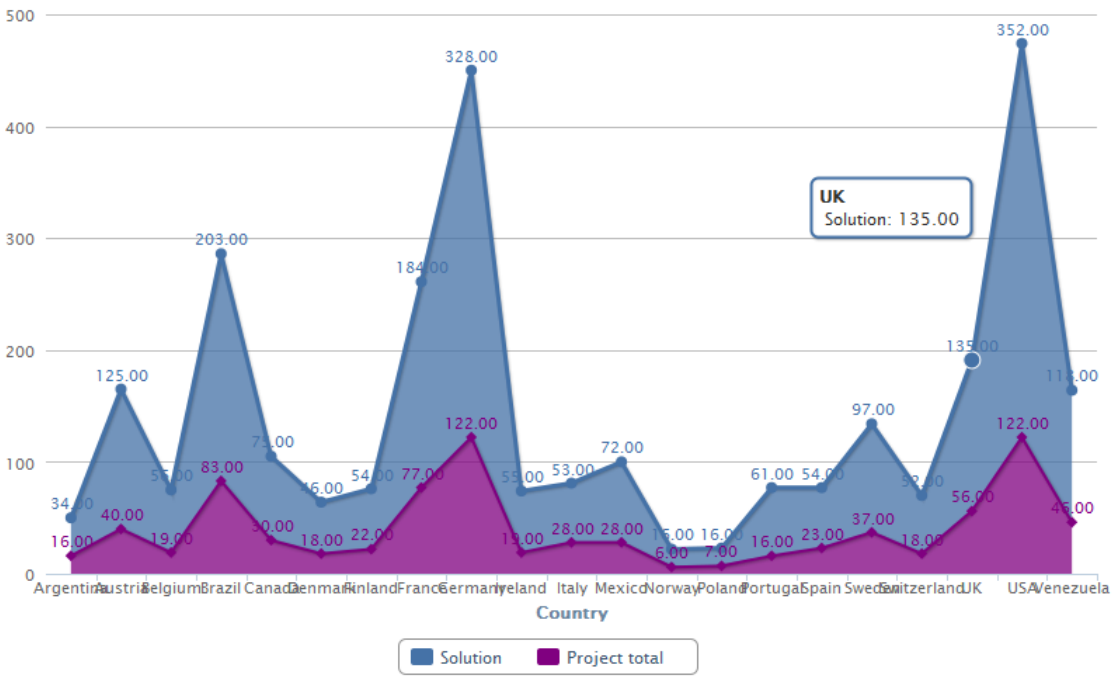


Figure 6.17: The projects and solutions report

the monthly revenue in the current year, in which the blue bar represents for revenue and the orange bar shows the total finished projects in the correlative month. (2) The 'Revenue by Category' shows revenue by the category dimension. (3) The 'Summary documents' table shows the number of documents submitted into the CrowdUM's OMIS system by category dimension based on two types as 'Normal' and 'Urgent'. (4) The 'Projects status' presents the comparison among total documents (the dark blue bars), total unresolved problems (the light blue bars), and total answers/replies (the orange bars) divided by the category dimension. (5) The 'Chart of Total Revenue' indicates the total of revenue by category dimension and percent of revenue held by each category. (6) Finally, the 'Solution by country' report focuses on the total projects and solutions by geography dimension as

Figure 6.18: The example charts to create business reports.

introduced in earlier.

6.6.2 Dashboard system

Following, the Dashboard system is a dynamic component that is updated continuously with the OLAP multidimensional cubes to quickly notify the managers about the status of CrowdUM. To demonstrate, I implemented a compound dashboard that includes 8 indicators as showed in the figure 6.19:

- **The revenue indicator** shows the total revenue in year 2011 (knew as the current year), this indicator also measures the risk of current revenue through three segments for instances: the red part (from 0 to 200 thousand USD) indicates the low level (high risk), the yellow part (up to 700 thousand USD) indicates the medium

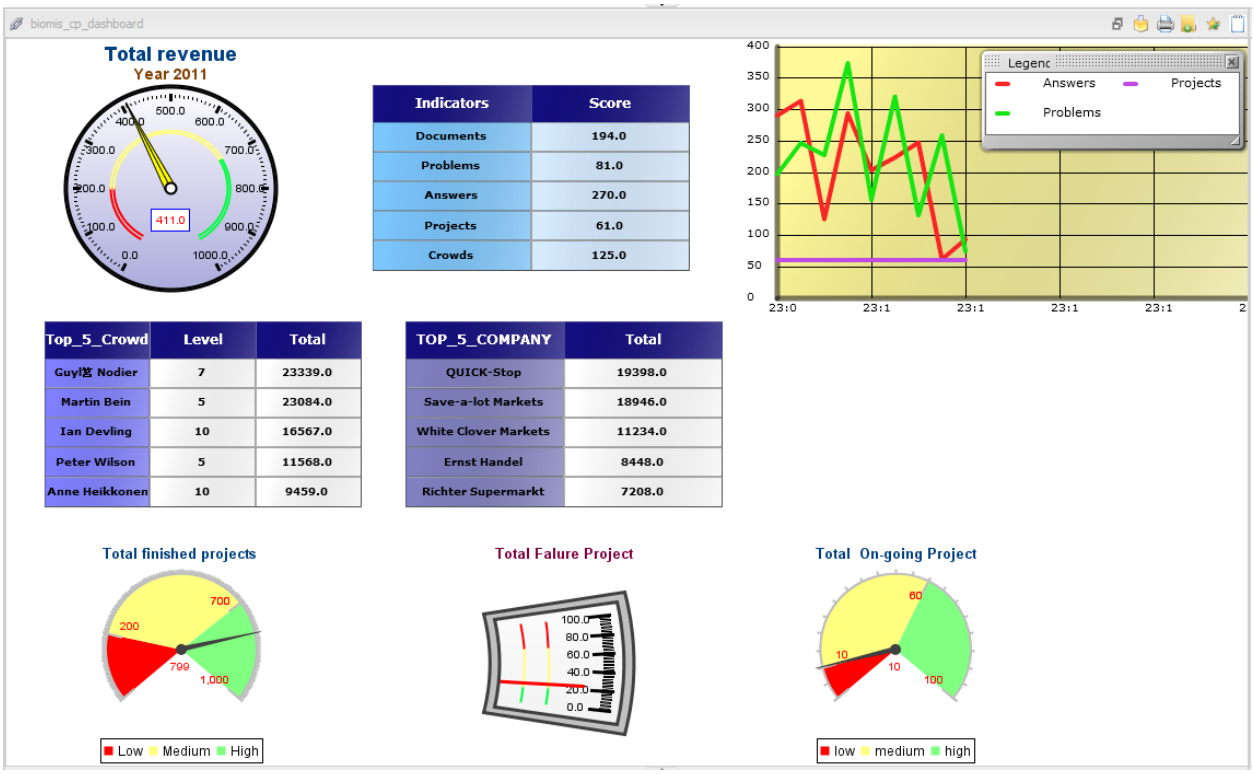


Figure 6.19: The OMIS dynamic dashboard in action

level or medium risk, and the over 700 thousand USD in revenue indicates the low risk with green color. Based on the current value is 411 thousand USD, hence the indicator shows that this current revenue is in medium risk. In contrast, if the pointer lies in the range from 0 to 200 thousand USD, it means that we are in the high risk and managers need to pay attention in improving the low business of CrowdUM such as intensify on failure projects, finding more projects, etc.

- **The CrowdUM status indicator** allows managers can see the total documents in the CrowdUM's OMIS, total current problems, answers, projects, as well as the number of crowd members. Moreover, the value of total problems and answers will be tracked in the next indicator, at which the green line indicates the unresolved problems, the red line indicates the total answers, and the violet line shows the number of on-going projects. Besides, this information will be refreshed after 1 hour (or could be redefined by managers) to help us have a comparison with the past time of this information then we can perceive the increase or decrease of these values. As the current demo, we can see that the total problems are high at first but they are reduced after 2 hours, hence this is such a good signal.

- **The incentive mechanism** relies on two next indicators to show the most dynamic solvers, and also the most valuable seekers in the CrowdUM. In that, the first indicator shows top five crowd solvers with decrease sort order on their total revenue. And the second indicator shows the top five companies (knew as seekers) by their highest payment to CrowdUM. By this way, the CrowdUM's managers could put in place effective incentive mechanisms with the solvers as well as the seekers. The incentive will create a motivation to help crowd members continues to pay attention with our organization and create a target for the newbies to endeavor.

- **Managing project's status** is the last feature of the OMIS dynamic dashboard. Via three indicators, the status of project is presented in three types for instances: (1) the total finished project indicates managers how many projects were finished success-

ful, in the same way with other indicators the current pointer shows the high value of total finished projects within the green color area. (2) Besides, the total failure project indicates the unsuccessful risk of CrowdUM's projects, in which the green area represents for the safe value, and in opposite if the pointer lies in the red area, it means that there are existing some urgent problems that need to be fixed right away. (3) Finally, the total on-going project indicates the number of projects that need to resolve soon, hence in this indicator the higher value is better.

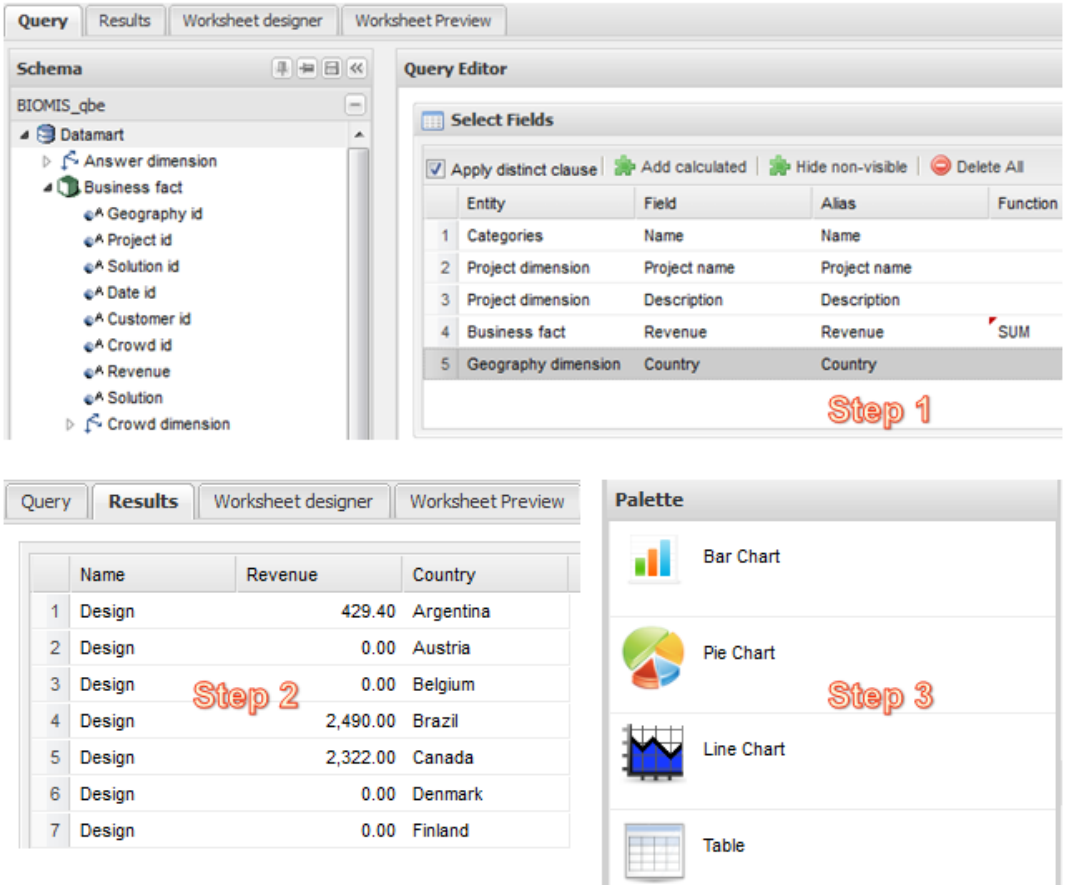
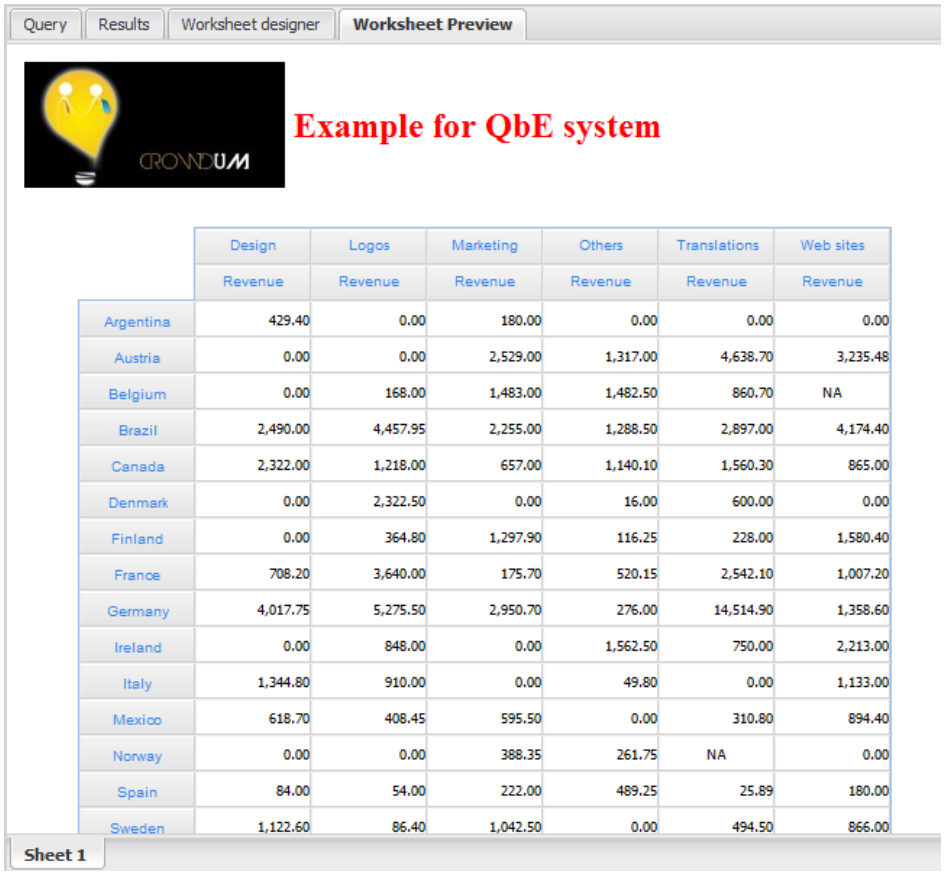


Figure 6.20: Functions of QbE in the BI-OMIS system

6.6.3 QbE system

As mentioned before, the QbE system is a flexible component that allows users to create specific queries for different purpose. Basically, the QbE proposed three steps to build a query, the first step of QbE enables user building visual queries from the CrowdUM's business cube and document cube as represented in the step 1 of figure 6.20. In this step we can choose relevant dimensions, define measurements (e.g., SUM, AVG, COUNT), and describe query conditions based on their purposes. In the second step, the QbE will list all results rely on our query, consequently, these data enable users to draw various chart types and table as illustrated in the third step of figure 6.20, which can be imported into user's reports later.



	Design	Logos	Marketing	Others	Translations	Web sites
	Revenue	Revenue	Revenue	Revenue	Revenue	Revenue
Argentina	429.40	0.00	180.00	0.00	0.00	0.00
Austria	0.00	0.00	2,529.00	1,317.00	4,638.70	3,235.48
Belgium	0.00	168.00	1,483.00	1,482.50	860.70	NA
Brazil	2,490.00	4,457.95	2,255.00	1,288.50	2,897.00	4,174.40
Canada	2,322.00	1,218.00	657.00	1,140.10	1,560.30	865.00
Denmark	0.00	2,322.50	0.00	16.00	600.00	0.00
Finland	0.00	364.80	1,297.90	116.25	228.00	1,580.40
France	708.20	3,640.00	175.70	520.15	2,542.10	1,007.20
Germany	4,017.75	5,275.50	2,950.70	276.00	14,514.90	1,358.60
Ireland	0.00	848.00	0.00	1,562.50	750.00	2,213.00
Italy	1,344.80	910.00	0.00	49.80	0.00	1,133.00
Mexico	618.70	408.45	595.50	0.00	310.80	894.40
Norway	0.00	0.00	388.35	261.75	NA	0.00
Spain	84.00	54.00	222.00	489.25	25.89	180.00
Sweden	1,122.60	86.40	1,042.50	0.00	494.50	866.00

Figure 6.21: The QbE result in action

Furthermore, the QbE component also allows us to save the final result (e.g., query statements, table results, reports) for later reusing in future, as well as sharing with other CrowdUM's members via the same kind of other documents in the OMIS of CrowdUM (see the figure 6.21). As a result, these features of QbE help crowd members avoid the redundant efforts in repetition of past jobs by searching and reusing available documents and reports that shared via the BI-OMIS system.

Chapter 7

Conclusions

In the final chapter of this dissertation, I briefly intend to highlight the most important points of the BI-OMIS integrated model in supporting the OMIS of CrowdUM, as well as summarizing the achieved results that were outlined in previous chapters. In this context, I am going to discuss the achieved results in this dissertation throughout four sections:

- The first section is going to discuss about the results that were gained throughout adapting and implementing the BI-OMIS integrated model into the crowdsourcing initiative environment.
- The second section is going to propose recommendations to enhance the Organizational Memory via the support from these suitable components of the Business Intelligence system.
- Because the final results in normally do not always fully meet objectives in the proposal, hence the third section mainly focuses on discussing about limitations in the implementation of the BI-OMIS integrated model.
- In the last section, based on the limitations in the previous section, I will present and discuss about the future works in the light of uncompleted objectives to consolidate my thesis in the future.

7.1 Conclusion

Based on the usefulness of past knowledge in supporting managers to enhance the decide making and problem solving, the Organizational Memory concept is becoming a key impact factor in large organizations. However, following the main researches on the concept of OMIS since the year 1995, there are only three cases research about the OMIS which intend to develop the real OMIS system (see the table 4.1). In addition, the OMIS researches are still faced with many challenges in practice for instances: challenges about the quality of information proposed from the OMIS, the role of context and meaning in knowledge management, and also the effective in managing a large amount of data in OMIS, etc. These challenges indicate that the OMIS concept is still limited in theory research and lack of methodological and technological tools to be used by practitioners. Therefore, this is an important motivation to promote my research to improve the performance of OMIS in practice.

Moreover, in another aspect to make clear the strength of BI system, the results of this dissertation found significant roles of the BI system in the organizational memory support, which covered a large amount of suitable tools such as decide support system, data mining, data warehouse, etc. As a result, the BI-OMIS integrated model is proposed to enhance the OMIS performance via four substantial layers (see figure 4.5): (1) the acquisition layer plays an importance role in capturing and adapting data from many resources concerns with the BI-OMIS (e.g., ETL, web crawler); (2) the retention layer has a responsibility in managing capability of the organizational repository by using the data warehouse; (3) the maintenance layer includes almost important BI tools to enhance, analysis, and classify the stored data in multidimensional views such as data mining and OLAP Cubes; (4) and the last BI's tools are settled in the search & retrieval layer, which are intensive on the graphical user interface (GUI) to facilitate the user interaction with the BI-OMIS system (e.g., QbE, Spread-Sheet, Report, Dashboard).

Consequently, in order to experience the BI-OMIS integrated model, the proposal

model was implemented in the real context of the crowdsourcing initiative (CrowdUM) with the support from an open source BI system (SpagoBI). In which, the CrowdUM plays a main role in the crowdsourcing model as a brokering service (the intermediary), which allows seekers and solvers to access to a wide range of skills and collective knowledge via the web-based platform (see figure 5.2). In the same way, the SpagoBI is an open source BI platform that gathered a set of suitable BI tools in order to implement the BI-OMIS integrated model (see figure 5.3). Finally, the BI-OMIS system will support the CrowdUM throughout seven steps as indicated in the BI-OMIS work flow model (see figure 6.1) such as: (1) Initiative the crowd via the social network, (2) The seeker registers a project into CrowdUM, (3) The BI-OMIS system maps and spreads relevant projects to appropriate solvers, (4) The BI-OMIS system supports solvers to solve their problems, (5) Enable solvers to submit their solutions and rank other solutions, (6) The seekers use the BI-OMIS system to select the best solutions, (7) The BI-OMIS system enables CrowdUM's managers to inspect the business of CrowdUM via the dynamic indicators in dashboard and various charts.

Finally, the BI-OMIS system is implemented and examined via five main sections of this thesis for instances: 'The BI-OMIS system interacts with the social network' (see chapter 6.2); 'The role of ETL and OLAP components in the BI-OMIS system' (see chapter 6.3); 'The BI-OMIS in supporting CrowdUM's OMIS' (see chapter 6.4); 'The DSS for supporting customer to select a solution' (see chapter 6.5); and 'The BI system in managing the CrowdUM business' (see chapter 6.6). As a result, these overcomes, experiments, and conclusions from the examination demonstrated a significant potential of the BI system in supporting the OM throughout three main functions of the OM structure such as: Information Acquisition, Retention Facilities, and Information Retrieval (proposed by Walsh & Ungson [1991]). Consequently, these possible results will create a first stepping-stone on the path to realize the BI-OMIS system for the CrowdUM system in particular, as well as for the real organizational environment in general.

7.2 Recommendation

After a thorough analysis and develop the BI-OMIS integrated model in the light of practice, the following recommendations are intended to highlight the main suggestions that directly affect to the final result while developing the BI-OMIS system as follows:

1. This research study suggests that the BI is a prominent component to empower the OMIS that enables the crowd members to work together more effectively. With the adaptive BI tools into the CrowdUM's OMIS, this not only shows a good ability in managing a large amount of documents within multidimensional views throughout many years (e.g., ETL, DW, OLAP cubes), but also facilitates the searching and sharing information in the past to other crowd members while solving the new problems/tasks with the support from the OLAP search tool called QbE.

2. Given that the DM holds a high potential in classifying and mapping the suitable information (e.g., documents, projects, comments, feedback) into relevant groups and participants inside the CrowdUM by using the DM tool such as classifier functions, clustering functions. Moreover, the DM's classifier algorithms provide such an intelligent solution for finding and mapping relevant information from the OMIS to suitable users actively, which depend on their personal information and their past behaviors captured from the social network (e.g., Facebook).

3. To provide a DSS for evaluating solutions from the massive submissions, this study recommends an open solution for seekers to ensure the selected solutions are the best ones. This result comes from the using of DSS combined from the comment indicator and the vote (like) indicator. The DSS then allows seekers could inspect a large amount of solutions proposed by crowd quickly and suitably. Moreover, the DSS is also a good method to reduce the most crap solutions from spammers.

4. For the suitable interaction between the users and the OMIS, the BI-OMIS supports a web-based platform that allows a large scale of crowd members from various

places and time could work together. Moreover, the BI system also provides the dashboard, report, and spread sheet components to help managers to control the business status of CrowdUM immediately throughout the BI-OMIS system. This plays an important role to notice managers the weaknesses of the OM that need to give high attention, for example: the dashboard component will indicate the risk of OM system via notifying the current number of problems, answers, and failure projects dynamically as indicated in the figure 6.19.

Throughout these recommendations for the OMIS in general and the CrowdUM's OMIS in particular, we can see these significant roles of the BI system in supporting a large scale of the OMIS mnemonic functions (see the figure 4.3), which spread from the Acquisition, Retention, Maintenance functions to the Search and Retrieval functions.

7.3 Limitation

Despite of the carefully prepared, as well as reached results that exhibited in the framework of proposed work (see the chapter 2.3), there were some inevitable limitations concerning with my research.

First of all, because of the research was examined in the real context of the crowdsourcing initiative (CrowdUM), which was deployed online just since the middle of year 2011, hence the real data are limited in a little bit projects and some other documents developed with them. This limitation restricted the suitable validation while implementing the BI-OMIS system in CrowdUM context. Moreover, as a result from the lack of real data, my research has to generate junk data for testing the performance of the BI-OMIS system in the CrowdUM context. For that reason, in the process of validating the BI-OMIS system, the real data does not offer enough comprehensive to accurately evaluate all cases in this dissertation. Thus, this research is limited information on the BI-OMIS overall validation.

Second, the personal information that collected from the social network is such a good resource for the OMIS to comprehend the human being by their hobbies. However, there is a limitation on the mapping method between the website users click like and the accurate hobbies users really like. For example, if user clicks like the page designcrowd[1], it is still a hard issue to understand which kind of category belongs to their hobbies. In this study, the simple solution here is to define a mapping array that held the most common sites with definitive hobbies and allows the user to adjust their really information. Anyway, there is still a plentiful resource to help the crowdsourcing comprehend about their crowd members.

Finally, the third limitation in this dissertation depends on the technical aspect. As we know, the CrowdUM was built by the Joomla - an open source Content Management System (CMS) - written in the PHP programming language. On the contrary, the SpagoBI - an entirely open source Business Intelligence suite - was written in the Java programming language. Therefore, there is a little bit incompatible communication between the CrowdUM system and the SpagoBI system. The current solution is to connect two separate systems through the ETL component; hence it still exists a lag time to synchronize the data between them.

7.4 Future researches

In the previous sections, I presented the main results, recommendations, and limitations throughout this dissertation. In addition, regarding the future work, no doubt that the development of the BI-OMIS integrated model is just a first step in using the BI system to support the OMIS in practice. Thus, it really needs more adjustments and refinements for future researches as proposed in three directions follows:

- As indicated in the limitation section, the first future work is to enlarge the real data

[1]http://www.designcrowd.com

of the CrowdUM system. This will help this study complete the validation process. Consequently, the results of the validation process will highly empower the entire roles of BI-OMIS integrated model in supporting the OM in real context.

- Second, with the closed relationship between the crowdsourcing and the social network, there is a new direction that leads this research to deeper comprehend the crowd hobbies throughout a plentiful resource from the social network. For this purpose, the mapping method that used in this research needs to be enhanced for deeper usage of the personal information. Certainly, this will enable the OMIS to interact with unique group in a more elaborate manner.

- Finally, based on the power of BI, this research used the DW to save documents' information in multidimensional cubes, meanwhile the DM was used to support the OMIS to classify and link relevant documents together via the classify functions. In order that, for the future research, I propose an effective mechanism for combining the semantic web concept with the DW and DM to empower the retention and retrieval features of OMIS. In which, the semantic web enables the OMIS to create a large network space of linked documents that keeps track the relationship between documents throughout the multidimensional views of DW (e.g, geographic, key-terms, categories, or time dimension). By this purpose, the OMIS allows user to easily expand the search results with nearby documents on different dimensions, which is adapted from the visual graphic interface of semantic web tool. Besides, the documents' correlation coefficient is analyzed by the DM functions to decide the relationship between the current document and other documents.

References

Abecker, A., & Decker, S. (1999). Organizational memory: Knowledge acquisition, integration, and retrieval issues. *XPS-99: Knowledge-Based Systems. Survey and Future Directions*, 113–124.

Abel, M.-H. (2008). Competencies management and learning organizational memory. *J. Knowledge Management*, *12*(6), 15-30.

Ackerman, M. S. (1994). Definitional and contextual issues in organizational and group memories. In *Hicss (4)* (p. 191-200).

Ackerman, M. S., & Halverson, C. (2004). Organizational memory as objects, processes, and trajectories: An examination of organizational memory in use. *Computer Supported Cooperative Work*, *13*(2), 155-189.

Ackerman, M. S., & Mandel, E. (1995). Memory in the small: an application to provide task-based organizational memory for a scientific community. In *Hicss (4)* (p. 323-332).

Amidon, D. M. (1991). Origins of the knowledge-based firm. *Management Systems Research*. Available from `http://www.entovation.com/gkp/origins.htm`

Atwood, M. E. (2002). Organizational memory systems: Challenges for information technology. In *Hicss* (p. 104).

Basaruddin, S., & Haron, H. (2011). Knowledge taxonomy for developing organizational memory system (oms) for public institutions of higher learning (ihl) in malaysia. In J. M. Zain, W. M. b. Wan Mohd, & E. El-Qawasmeh (Eds.), *Software engineering and computer systems* (Vol. 180, p. 43-57). Springer Berlin Heidelberg.

Bichindaritz, I. (2006). Memory organization as the missing link between case-based reasoning and information retrieval in biomedicine. *Computational Intelligence*, *22*(3-4), 148–160.

Bichindaritz, I. (2008). Memory structures and organization in case-based reasoning. *Case-Based Reasoning on Images and Signals*, 175–194.

Blue, J., Andoh-Baidoo, F., & Osatuyi, B. (2011). An organizational memory and knowledge system (omks): Building modern decision support systems. *International Journal of Data Engineering (IJDE)*, *2*(2), 27.

Borrajo, M., Baruque, B., Corchado, E., Bajo, J., & Corchado, J. (2011). Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises. *International Journal of Neural Systems*, *21*(4), 277–296.

Burstein, F., Linger, H., Zaslavsky, A. B., & Crofts, N. (1997). Towards an information systems framework for dynamic organisational memory. In *Hicss (2)* (p. 262-270).

Burstein, F., Smith, H., & Fung, S. (1998). Experimental evaluation of the efficiency of a case-based organisational memory information system used as a decision aid. In *System sciences, 1998., proceedings of the thirty-first hawaii international conference on* (Vol. 1, pp. 209–217).

Chapman, S., Lanfranchi, V., & Bhagdev, R. (2009). K-tools: Towards semantic knowledge management. *The Semantic Web: Research and Applications*, 806–810.

Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, *54*(8), 88–98.

Cheng, H., Lu, Y.-C., & Sheu, C. (2009). An ontology-based business intelligence application in a financial knowledge management system. *Expert Syst. Appl.*, *36*(2), 3614-3622.

Chou, T., & Cheng, S. (2006, APR). Design and implementation of a semantic image classification and retrieval of organizational memory information systems using analytical hierarchy process. *Omega-International Journal Of Management Science*, *34*(2), 125-134.

Chung, W., Chen, H., & Jr, J. F. N. (2005). A visual framework for knowledge discovery on the web: An empirical study of business intelligence exploration. , *21*, 57–84.

Cody, W. F., Kreulen, J. T., Krishna, V., & Spangler, W. S. (2002). The integration of business intelligence and knowledge management. *IBM Systems Journal*, *41*(4), 697-713.

Croasdell, D. T., & Paradice, D. B. (2002). Tooms: In pursuit of designing an organizational memory system for category manager. In *Hicss* (p. 107).

Davis, R., Shrobe, H., & Szolovits, P. (1993). What is a knowledge representation? *AI magazine*, *14*(1), 17.

DeLone, W. H., & McLean, E. R. (2003). The delone and mclean model of information systems success: A ten-year update. *J. of Management Information Systems*, *19*(4), 9-30.

Delone, W. H., & McLean, E. R. (2003, April). The delone and mclean model of information systems success: A ten-year update. *J. Manage. Inf. Syst.*, *19*, 9–30.

Dieng, R., Corby, O., Giboin, A., Ribière, M., & Lucioles, R. D. (1998). *Methods and tools for corporate knowledge management*.

Dunham, A. H., & Burt, C. D. B. (2011). Organizational memory and empowerment. *J. Knowledge Management*, *15*(5), 851-868.

Fan, W., Gordon, M., Pathak, P., & PATHAK, P. (2005). Genetic programming-based discovery of ranking functions for effective web search. *Journal of Management Information Systems*, *21*(4), 37–56.

Farquad, M. A. H., Ravi, V., & Raju, S. B. (2009). *Data mining using rules extracted from svm: An application to churn prediction in bank credit cards*.

Fortier, J., & Kassel, G. (2003). Building adaptive information retrieval systems for organizational memories: a case study. In Satyadas, A and Dascalu, S (Ed.), *Proceedings Of The Isca 12th International Conference Intelligent And Adaptive Systems And Software Engineering* (p. 50-53).

Goesmann, T. (2001). Kontextnavigator: A workflow-integrated organizational memory information system to support knowledge-intensive processes. In *Inap* (p. 393-403).

Hackbarth, G., & Grover, V. (1999). The knowledge repository: Organizational memory information systems. *IS Management*, 21-30.

Halverson, C., & Ackerman, M. S. (2003). "yeah, the rush ain't here yet - take a break": Creation and use of an artifact as organizational memory. In *Hicss* (p. 113).

Hamidi, & Kamaruzaman. (2009). The characteristic and success factors of an organizational memory information system. *Computer and Information Science*, *2*(1), 142-151.

Hamidi, S., & Abdul Rahman, A. (2006). Organizational memory information system: Case study in faculty of computer science & information system, utm.

Herskovic, V., Ochoa, S., & Pino, J. (2006b). A model to incorporate privacy in organizational memory systems. In *Computer supported cooperative work in design, 2006. cscwd'06. 10th international conference on* (pp. 1–6).

Herskovic, V., Ochoa, S. F., & Pino, J. A. (2006a). A model to incorporate privacy in organizational memory systems. In *Cscwd* (p. 989-994).

Howe, J. (2006). The rise of crowdsourcing. , *32*(1). Available from `http://www.wired.com/wired/archive/14.06/crowds.html`

Hwang, W., & Salvendy, G. (2005). Effects of an ontology display with history representation on organizational memory information systems. *Ergonomics*, *48*(7), 838–858.

Jennex, M., & Olfman, L. (1997a). The impact of organizational memory information systems on organizational learning. In Gupta, JND (Ed.), (p. 985-987).

Jennex, M., & Olfman, L. (1997b). Training and intranet recommendations from a case study of organizational memory. In Gupta, JND (Ed.), *Association for information systems proceeding of the americas conference on information systems* (p. 524-526).

Jennex, M., Olfman, L., Panthawi, P., & Park, Y. T. (1998). An organizational memory information systems

success model: an extension of delone and mclean's i/s success model. *Proceedings of the ThirtyFirst Hawaii International Conference on System Sciences*, *I*(C), 157–165.

Ji, Y., & Salvendy, G. (2002a). Development and validation of user-adaptive navigation and information retrieval tools for an intranet portal organizational memory information system. *Behaviour & Information Technology*, *21*(2), 145–154.

Ji, Y., & Salvendy, G. (2002b, APR). A metadata filter for intranet portal organizational memory information systems. *International Journal Of Human-Computer Studies*, *56*(5), 525-537.

Ji, Y., & Salvendy, G. (2004, DEC). Interface methods for using intranet portal organizational memory information system. *Ergonomics*, *47*(15), 1585-1597.

Kuhlman, A., & Deiters, W. (2000). The organizational memory information system prince. In *Coopis* (p. 236-250).

Lang, K. R., & Schmidt, M. (2002). Workflow-supported organizational memory systems: An industrial application. In *Hicss* (p. 208).

Lehner, F. (2000). Organisational memory. konzepte und sys-teme für das organisatorische lernen und das wis-sensmanagement. *Munich, Vienna*.

Linger, H., Burstein, F., Zaslavsky, A., & Crofts, N. (1999). A framework for a dynamic organizational memory information system. *Journal Of Organizational Computing And Electronic Commerce*, *9*(2-3), 189-203.

Machlup, F. (1983). Semantic quirks in studies of information. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 641–672). New York, NY: Wiley.

Mahapatra, R. K. (1997). Case-based reasoning: Extending the frontiers of knowledge-based systems. *Journal of Computer Information Systems*, *38*(2), 87–94.

Miller, L., Nilakanta, S., Song, Y., Zhu, L., & Hua, M. (2008). Managing knowledge in organizational memory using topics maps1. *IJKM*, *4*(1), 17-28.

Mokhtar, S., Bakar, Z., & Husin, W. (2008). A framework for faculty memory information system. *Innovation and knowledge management in business globalization: theory & practice, vols 1 and 2*, *6*, 1241-1244.

Morrison, J. (1997). Organizational memory information systems: Characteristics and development strategies. In *System sciences, 1997, proceedings of the thirtieth hawaii international conference on* (Vol. 2, pp. 300–309).

Nasukawa, T., & Nagano, T. (2001). Text analysis and knowledge mining system. *Ibm Systems Journal*, *40*, 967–984.

Nemati, H. R., Steiger, D. M., Iyer, L. S., & Herschel, R. T. (2002). Knowledge warehouse: an architectural

integration of knowledge management, decision support, artificial intelligence and data warehousing. *Decision Support Systems*, *33*, 143–161.

Nevo, D., & Wand, Y. (2005). Organizational memory information systems: a transactive memory approach. *Decision Support Systems*, *39*(4), 549–562.

Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company. *Oxford University Press*.

Ochoa, S. F., Herskovic, V., Pineda, E., & Pino, J. A. (2009a). A transformational model for organizational memory systems management with privacy concerns. *Information Sciences*, *179*(15), 2643 - 2655.

Ochoa, S. F., Herskovic, V., Pineda, E., & Pino, J. A. (2009b). A transformational model for organizational memory systems management with privacy concerns. *Inf. Sci.*, *179*(15), 2643-2655.

Ochoa, S. F., Herskovic, V., Pineda, E., & Pino, J. A. (2009c). A transformational model for organizational memory systems management with privacy concerns. *Information Sciences*, *179*(15), 2643–2655.

Pemberton, J. D., & Stonehouse, G. H. (2000). Organisational learning and knowledge assets – an essential partnership. *The Learning Organization*, *7*(4), 184–194.

Pirttimaki, V., Lonnqvist, A., & Karjaluoto, A. (2005). Measurement of business intelligence in a finnish telecommunications company. In *European conference on knowledge management* (pp. 451–458).

Ramos, I. (2011). Crowdsourcing creative tasks: the various models.

Rao, V. S., & Goldman-Segall, R. (1995). Capturing stories in organizational memory systems: the role of multimedia. In *Hicss (4)* (p. 333-341).

Schwabe, G. (1994). Providing for organizational memory in computer-supported meetings. In *Hicss (4)* (p. 171-180).

Schwartz, D. G. (1998). Shared semantics and the use of organizational memories for e-mail communications. *Journal of Internet Research*, *8*(5), 434-441.

Sciarrone, F., Starace, P., & Federicit, T. (2009). A business intelligence process to support information retrieval in an ontology-based environment. In *Intelligent systems design and applications, 2009. isda'09. ninth international conference on* (pp. 896–901).

Shepherd, M. M., & Martz, W. B. (2006). Organizational memory: Three experiments on the quality of information. *Cybernetics and Systems*, *37*(2-3), 151-169.

Sowa, J. F. (1999). *Knowledge representation: Logical, philosophical, and computational foundations*. Brooks Cole Publishing Co., Pacific Grove, CA.

Spek, R. v., & Spijkervet, A. (1997). Knowledge management: Dealing intelligently with knowledge. *CIBIT*.

Stein, E. W., & Zwass, V. (1995). Actualizing organizational memory with information systems. *Information Systems Research*, *6*(2), 85-117.

Stein, E. W., Zwass, V., Vandenbosch, B., Robey, D., & Sandoe, K. (1994). It-supported organizatinal memory: How will it affect organizational learning? In *Icis* (p. 489-490).

Stenmark, D. (2002). Information vs. knowledge: The role of intranets in knowledge management. *Hawaii International Conference on System Sciences*, 928-937.

Tsoukas, H., & Vladimirou, E. (2001). What is organizational knowledge? *Journal of Management Studies*, *38*(7).

Turban, E., Sharda, R., Aronson, J., & King, D. (2008). *Business intelligence: a managerial approach*. Pearson Prentice Hall.

Turban, E., Sharda, R., & Delen, D. (2010). *Decision support and business intelligence systems, 9/e*. Pearson Prentice Hall.

Vasconcelos, J., Kimble, C., & Rocha, A. (2003). Organisational memory information systems an example of a group memory system for the management of group competencies. *Journal Of Universal Computer Science*, *9*(12), 1410-1427.

Walsh, J. P., & Ungson, G. R. (1991). Organizational memory. *Academy of management review*, 57–91.

Wargitsch, C., Wewers, T., & Theisinger, F. (1998). An organizational-memory-based approach for an evolutionary workflow management system - concepts and implementation. In *Hicss (1)* (p. 174-183).

Watson, R. (1995). *Data management: an organizational perspective*. John Wiley & Sons, Inc.

Watson, R. T. (1998). Data management - databases and organizations (2. ed.). *New York*.

Wei, S., & Qing-pu, Z. (2007). Study on knowledge mining of the business intelligence system. In *Wireless communications, networking and mobile computing, 2007. wicom 2007. international conference on* (pp. 5435–5438).

Wiig, K. M. (1994). Knowledge management foundations: Thinking about thinking - how people and organizations represent, create, and use knowledge. *Schema Press*.

Wijnhoven, F., & Slooten, K. van. (2000). Three challenges of organizational memory information systems development. In Khosrowpour, M (Ed.), *Challenges Of Information Technology Management In The 21st Century* (p. 6-10).

Wijnhoven, F., & Slooten, K. van. (2002a). Three problems of organizational memory information systems development. *Knowledge mapping and management*, 37.

Wijnhoven, F., & Slooten, K. V. (2002b). Three problems of organizational memory information systems development. In *Knowledge mapping and management* (p. 37-48). IGI Global.

Yang, K.-J., & Chen, Y.-M. (2006). Ontology-based knowledge retrieval in organizational memory. *Innovative Computing ,Information and Control, International Conference on*, *1*, 566-569.

Zhan, J., Loh, H., & Liu, Y. (2009). Gather customer concerns from online product reviews-a text summarization approach. *Expert Systems with Applications*, *36*(2), 2107–2115.