CHAPTER 2

LITERATURE REVIEW

This study is *Decision Support System for Research Management in Universities*. It focuses on the problem of research management in universities. Thus in the first part of this chapter research management is reviewed in broad view of how managing research effective and efficiency both in Thailand and others countries. It also includes what are the crucial components in research management. At the end of this topic research management in universities was focused. This was explained current status of research management and what is the current issues, trends, and problems. Now research and researchers in universities are need to be managed in the most suitable way because they are can be counted as intellectual capital of organization. Thus intellectual capital theory was reviewed for finding the optimal method to identify, measure, and manage those intellectual capitals.

For solving the problem found in the first part, research management system should be built. Research management system is used to store detail of researchers, researchers' performance. In its implementing AI techniques is used. Then in the second part of this chapter, AI techniques are focus especially knowledge representation techniques and Ontology, one kinds of knowledge representation used in this study, was explained in depth.

For analyzing research information in ontology, some techniques were employed. These techniques were reviewed in the third part of this chapter. There are three kinds of knowledge engineering techniques: card sorting, social network, and ontology in knowledge management. Card sorting is the technique used for forming new research clusters. The social network is an idea used for analyzing to find out the well known and connection of key researchers in academic world. The last one, ontology in knowledge management, was reviewed the way how to apply ontology in various applications in the real world especially in knowledge management of research management.

2.1 Business Frameworks

2.1.1 Research Management

2.1.1.1 Research

The simple meaning of research is to search or search for again. Because "research" term is composed of two syllables a prefix "re" and a verb "search" (Dictionary.com Unabridged, 2008).

In Encyclopedia Britannica (2008), research is "Diligent and systematic inquiry or investigation into a subject in order to discover or revise facts, theories, applications, etc."

In Webster's third new international dictionary (1961), research is defined as studious inquiry or examination critical and exhaustive investigation or experimentation having for its aim the discovery of new facts and their correct interpretation the revision of accepted conclusions, theories, or laws in the light of

newly discovered facts, or the practical applications of such new or revised conclusions, theories, or laws.

For this study research is systematic inquiry or investigation into a subject in order to discover or revise fact or knowledge. The study focus on universities' research covered three disciplines – Health Science, Science and Technology, and Social Science and Humanities. Those research can divided in two groups, academic research group granting by Thai government and cooperating research group granting by others source both national and international.

2.1.1.2 Roles and Importance of Research

Research is a crucial factor for an organization to survive in these days because it is a process that leads to new knowledge or innovation that are advantageous to the organization and society (Viriyajaree, 2002). Those knowledge and innovations are considered as an Intellectual Capital (IC) of an organization, which enhances the organization's deeds to be effective and efficient, including reduces production cost and results in more benefits. Any organization can do research to find their weakness, thus, they can be evolved and solve any problems successfully. Moreover, research can also be used as a convention in a business decision-making in terms of investment, situation estimation, production development, and plan of strategy and policy (Nation Research council of Thailand. Special activity group, 2003).

Another important role of research is to increase manager's abilities in change management. Since communication system has been improved, the world seems to be smaller and more connected. Far distance is not an important obstacle anymore. Organizations can contact each other in a short time. The communicate system is like

an invisible link, connecting each organization together. When an incident occurs in any part of the world, others will be affected as well. It can be seen that organizations these days are at risk of effects from external factors all the time. Whether they are able to survive or not, the capability of handling the problems promptly, thus, is important. However, there are some organizations that do not only cope with problems well, but also be able to convert that crisis into an opportunity by using knowledge and innovation from research to succeed over competitors and handle a crisis well. It is obvious that research help increase the organizations' capability of coping with external changes. That means the ability of the organizations' competency is escalated (Viriyajaree, 2002).

In many countries, research and development (R&D) is also another crucial aspect that executive realizes as a way to stimulate economics. Many companies allocate a part of their budget to use in R&D spectacularly. Some companies have an R&D department but some distribute fund in doing research to other academic institutions, which makes those companies concern that if the work would be qualified or whether they can control the studies to meet their expectation, including the issue of the copyright belonging (Bowen, 2005).

Although research in Thailand is recognized from every organization, but when compared with developed countries, numbers of research in Thai still lower. Hence organization doing research should be supported and encouraged especially universities, an important source of knowledge and knowledge workers.

2.1.1.3 University Research

Universities have been playing an important role in obligating research effectively and continually since they are great sources for gaining and collecting knowledge through researching by knowledge workers, which in here means lecturers from various fields, including doctorate students (Monsted, 2002; Zarama, Reyes, Aldana, Villalobos, Bohorquez, Calderon, & et al, 2007). This knowledge has not only been used in teaching plans but also be used for widespread benefit to society, nation, and the world.

At present, many universities are forcing many ways to conduct excellent research supported by major national and international funding sources, expand research significantly, achieve a major increase in the level of funding, rise the number of active researchers, including appoint senior research staff in key subject areas (Ahmad, 2009; Higher Ed Services, 2009; Weber, 2004).

Another important part of universities is to distribute research results to communities and society, which several ways can be done (Chang, & Trubow, 1990; Monsted, 2002).

a.) Provide academic services to communities.

b.) Publish research work in academic journals.

c.) Cooperate research with other government and private organizations.

Chiang Mai University (CMU) gives high priority to research as well. The university has made research be the most important of the university's missions and assigns researching to be another task for teachers besides teaching lessons due to the following reasons:

a.) Doing research is another way of enhancing the teacher's abilities after graduated.

b.) The result of research can be reused in the teaching process.

c.) New generation researchers and students can develop their skills by doing research with their lecturers who have great experiences in doing research.

d.) Research results is useful for local, national, and international advantage.

CMU has divided research into 3 fields Health Sciences, Sciences and technology, and Social sciences and Humanities. Under the 3 fields are 17 faculties and 3 research institutions (Kangwanpong, 2006):

a.) Health Sciences discipline.

- Associated Medical Sciences.
- Dentistry.
- Medicine.
- Nursing.
- Veterinary Medicine.
- Research Institute of Health Sciences.

b.) Sciences and technology discipline.

- Agriculture.
- Agro-Industry.
- Architecture.
- Engineering.
- Sciences.

- Institute for Science and Technology Research and Development.

c.) Social Sciences and Humanities discipline.

- Education.
- Economics.
- Fine Arts.
- Humanities.
- Mass Communication.
- Social Sciences.
- Social Research Institute.

Even through universities are concentrated on subjects enforcing research directions of all researchers match to research direction of universities and the country is quite hard. Research management should be applied including new trend of management such as Knowledge Management, change management, project management, and risk management.

2.1.1.4 Research Management

To succeed in doing research, research management is required. Research management is an activity that has to be done continually and concurrently to achieve and give the most benefit to the public (Thailand Research Fund, 2005). In Thailand, there are 4 factors in research management: research manager, management of research unit, management of research support process, and researcher management.

a.) Research manager is an important person in pushing research to achieve and continue the research work. In a university, there are different levels of research managers:

- Vice presidents of research.
- Directors of research institutes.
- Deputy deans of research affair.
- Deputy heads of research.
- Directors of research units/centers.
- Head of research project.
- Etc.

Research manager in each level should consider in several aspects for success in research management:

- Manage the research which is in their responsibility and corresponds to the strategy and policy of the higher level research. For example, research deputy dean should formulate the research strategy which consent to the university's strategy (Thailand Research Fund, 2005). Thus, the country's researches aim for the same direction and truly meet the country's need.

- Understand in multi-discipline research because universities currently aim at doing integrated and multi-discipline research (Conraths, & Sodergard, 2004; Thailand Research Fund, 2005; Weber, 2004). Complicated problems are difficult to solve by only single knowledge discipline. Thus, a combination of different knowledge is required. Generally, there are various experts in various fields in a university, which means a university is a great source of a strength multi-discipline

research team (Coovathanachai, 2007). A research team can be built efficiently by a cooperation of related fields in producing research projects together.

- Avoid hierarchical management which is complex and has many levels such as bureaucratic process because it is a time consuming (Monsted, 2002; Thailand Research Fund, 2005). It may also not be able to respond immediately to a change.
- Encourage to initiate coordination among government organizations, private organizations, and academic institutions, which can be in several ways:
  - Research fund support.
  - Research cooperation.
  - Research usage.
  - Knowledge or staff exchange.

Producing research from coordination can benefit in several ways and for different parts, called "Win-win situation" (Chang & Trubow, 1990; Hacker, n.d.; Numprasertchai, & Igel, 2005; Srihari, & Sammakia, 2002; Zarama et al, 2007).

- Allocate a part of the organizations' income for research activities or search for research fund from outside: government and private organizations (Coovathanachai, 2007).

b.) Management of research unit.

A research unit is a unit that is responsible for doing research, which can have different aims such as academic, commercial, public, and policy research. The research unit must be well organized and has as effective and efficient management as follow (Areekul, n.d.).

- The unit manager should have vision and be able to grasp the whole research process both at national and international levels.
- Make a research plan in both short and long terms.
- Have a measure which can make the research succeed.
- Have sufficient, efficient, and up-to-dated resources. Well management of resources as well.
- Have a precise research process.
- Have an adequate use of database system for researchers.
- Use QC and QA.
- Have a fast and clear communication.

Nowadays, many universities establish many research units called Center of Excellence (COE) involve running research in a specific topic (Conraths, & Sodergard, 2004; Weber, 2004). Thailand also needs more development in particular research to lead the country to be excellent, due to the Ministry of Science and Technology plan (2004-2013) by willing to establish the Center of Excellence (Nation Research council of Thailand. Special activity group, 2003). As the research unit has been more expected by the nation, research unit should: (Areekul, n.d.; National Research Council of Thailand, 2001).

- Create a standard and reliable research .
- Create researchers by using expert's knowledge that meets the country's need.
- Focus research topic on important issue of the country in depth and width.
- Continue research in importance issues.

There are 7 subjects in COE's research (Nation Research council of Thailand. Special activity group, 2003).

- Electronic and computer.
- Metal and materials.
- Biotechnology and military.
- Medical and public health.
- Energy and environment.
- Nanotechnology.
- Industrial development.

From the 2002-2004 survey, there are 145 centers and most of them are in the field of Biotechnology and Agricultures (Nation Research council of Thailand. Special activity group, 2003)

An example center of this field is the Excellence Center of Rice Innovation, established by Kasetsart University. The objective of this center is to (Kasetsart University, n.d.)

- Exchange knowledge and research in the field of rice among the university's researchers and staffs who are expert in rice.
- Specify strategies in specific research on rice to be precise, complete, and beneficial.
- Use the knowledge into teaching lessons and gather together that knowledge to be a complete Rice Knowledge Center.
- Distribute and transfer that knowledge to students, communities, and commercial industries.

In the field of Social Science and Humanities, there are also some efforts in establishing the COE. But, it has not been as obvious as in the Science field. However, in 2006, NRCT conducted an academic conference to specify a way to initiate an Excellent Center specifically in Social Science and Humanities, and Behaviour Science. The purpose of this center is to establish a research institution which is (National Research Council of Thailand, 2001)

- Able to enhance the life quality of Thai people and also the society.
- Capable of solving any problems in individuality, society, economic, and politics.
- A basic institution which is a supporter of the state and private organization.
- A creation of new knowledge and be able to follow changes in every part of Thai society continually and completely.
- Able to apply the research result promptly.

The Higher Education Commission of Thailand grant for COE to government universities as well. This makes CMU research administrators have strong policy to establish some COE. Present, CMU already established many COEs or unit of COE as follow:

1. Center of Excellence on Elderly Care, Faculty of nursing; it was found on 2003. The objectives of this center are to be the hub of researchers and expertise on elderly care and research for developing knowledge on elderly care both for elderly care education and elderly care treatment (Center of Excellence on Elderly Care, 2009).

2. Center of Excellence in Medicine, faculty of medicine: it has medical service for both inside patients and outside patients. It is center of many diseases such as heart disease, cancer, cerebropathy ("CMU establish center of excellence for producing doctors and medical services". 2009).

3. Research Center in Particle Beam and Plasma Physics (PBPP): This is one center of Thailand center of excellence in physics (ThEP). This COE is under board of the Higher education commission (Thailand center of excellence in physics (ThEP), n.d.).

4. Expert Center of Nanotechnology: This center is one center of Center of national nanotechnology. This center found in 2006, which is research in nanomaterials in special section that focuses on nano-oxide, nano carbon, nano zinc oxide, and nano-composite. (Nanotech signed contract with Chiang Mai University, 2549).

5. Institute of Earth Sciences is autonomous organization under CMU governance: It was found for 80th Celebration Birthday King Bhumibol Adulyadej. This institute is going to be center of excellence of earth sciences. The object of this center is research and distributed knowledge in earth sciences i.e. earth and changing of earth. The changing is about geology, geography, pedology, oceanography, etc. (Institute of Earth Sciences, 2009).

Although, CMU recognized an important of COE approach, but how to set up a COE is still difficult. It is because CMU administrators do not know research information of their researchers both potential of researchers and research direction of

researchers. This makes CMU administrators do not know which COE should be established.

c.) Management of research support process.

Research support process is any process that supports research. It is another aspect to be considered to produce qualitative research and also punctual research result. The processes are proposal development, proposal evaluation, research advancement monitor, final report evaluation, seminar attendance, patent registration, and research result application (Coovathanachai, 2007).

In each process, the research manager should consider critically on:

- Proposal development process is critical in producing a research which meets the nation's requirement. Researchers should be encourage to make a research proposal by considering the country and the organization's strategies of researching. Furthermore, participation method can be used to receive various opinions from people who use the research result in their works.
- Proposal evaluation: the research manager should evaluate the proposal by considering the research strategy of the organization.
- Monitoring some research advancement can be done by evaluating the progress from time to time. This can help the researchers know about the current situation. Grantor can also be used to encourage researchers to finish the work punctually.
- Final report estimation is a process that investigate whether the research correspond to the research proposal. This makes research managers be able to control the research direction of the organizations efficiently.

- Conference attendance both national and international should be more encouraged so that researchers have a chance to present their studies publicly, including introducing their works to people who may use them.

- Obtaining a patent should be more registered to protect intellectual property, which can reduce the problem of copyright infringement of research works.

d.) Researcher management.

Researchers are considered as a heart of a research management, who can produce qualitative research continually. At present, every organization desires more high potential researchers (Thailand Research Fund, 2005). Thus, persuading those researchers to work in the organization constantly is important. However, the key to succeed in managing researcher, which research managers should consider, is to make researchers satisfy and want to create excellent works for the organization, which can be done by:

- Create a research environment (Monsted, 2002). Besides interior design or suitable laboratory and efficient equipments, there is also another way, which is to create activities and culture that give a chance for researchers to meet each other and exchange their opinions and works both formally and informally.

- Support researchers to develop their skills constantly such as sending to training, seminars, or a study trip either in the country or aboard. This can make researchers open their mind and be able to follow academic advancement.

- Specify measurement to evaluate the work through researchers, which will stimulate them to improve themselves in the way the organization requires (Kirkland, 2008; Monsted, 2002).

- Allocate sufficient and suitable fund. The fund should not only cover all fields, but should also be allocated due to different natures of the fields. For example, some subjects have to use laboratory, special equipments or expensive materials.
- Acquire standard database that cover all subjects which the organization's researchers have to use in their studies.

In Thailand, it is currently found that research management still faces several problems (Kangwanpong, 2006; Thailand Research Fund, 2005):

- Lack of qualified researchers.
- Lack of research fund.
- Lack of resources.
- Has difficulty in adjusting to changes.
- Lack of systematic connection/link between research processes including monitoring system.
- Has a low spending on R&D, considered by the country's GDP.
- Has a small number of scientists and workers in industrial section.
- Less spending on R&D in private business.
- Lack of interest in science and technology among new generation.

Consequently, Thailand needs to strengthen the research management system in several ways. Also, education system has to focus on research by considering as an important base (Thailand Research Fund, 2005).

The research management needs to apply some management theories such as knowledge management, risk management, change management, project management.

2.1.1.5 University Research Management

Thai universities have managed research by specifying research strategy regarding to the nation's policies. In the past, lecturers would do research as they preferred which made those works do not meet the country's requirement in those days. If the research strategy of the nation and organization are effectively aligned, the nation's research can be more strengthened and have a clear direction.

CMU has focused on research strengthening and how to produce qualified research, which will contribute to welfare, economic, and environmental development of the country. Moreover, the research strategy has also conducted corresponds to the nation's policies and the development strategy of Northern local. CMU also encourages and supports researchers in every level of management (department, faculty, and university), from basic research to applied research, including multi-discipline research establishment and development. The aim of multi-discipline research is to produce research works which support to the nation's direction of research. A connected project between related fields has been established and there is also a continue work.

Research of CMU covers in all disciplines, which can be divided into 4 fields

a.) Health Sciences.

b.) Sciences and Technology.

c.) Agriculture and Industrial agriculture.

d.) Social Sciences and Humanities.

In research management, CMU administrators have set up research supportive fund which is available for supporting research works, developing research, finding fund, encouraging faculties to establish center of excellence, conducting a conference

for presenting research works, forming research teams, and stimulating patent registration. The board of this fund consists of (Chiang Mai University regulations, 2001):

a.) President.

b.) Funding board.

c.) Vice president for Research Academic Services.

d.) Vice President for Planning, Financial and Property Management.

e.) Two deans from each field.

f.) Director of three research institutes.

g.) Resource person.

h.) Director of Financial division.

i.) Director of International Relations Division

The board has the 2-year term and their main duty is to grant research funds especially for academic to acquire necessary equipment in starting research then attract some more grants from external funding agencies. Beyond that, the board has other minor responsibilities in awarding excellent researchers, reporting research results, supporting in publication, finding more income and other roles in supporting research in the university.

CMU administrators also pay attention in providing fund for research, they has arranged money to support research in many ways from (Chiang Mai University regulations, 2001):

a.) Approximately 10% of the university's income (not include collected money).

b.) 1% of other universities' incomes (not include collected money).

c.) 3% of research fund of outside.

d.) Donation.

e.) Income that comes from the research work.

f.) Approve money from the university's congress.

g.) Other support money.

h.) Interest.

At present many universities' administrators convince their universities to become a research world class university (Agadri, 2009), CMU administrators as well. CMU is now recognized as a research university in Thailand. There has been a continual push since 2000. The university is currently trying to shift by producing more efficient research works. The aim is to be internationally well recognized as other universities in Asia such as Singapore, Korea, and Taiwan. In 2006, CMU was officially ranked number three in having qualitative research in Thailand following Mahidol University and Chulalongkorn University respectively (Chiang Mai announces, 2006).

In 2009, CMU was ranked 4th in top-9 ranking of National Research Universities (Thailand) in 2010. This rank announced by National Research University project founded by Office of the Higher Education Commission. The ranking is ordered from Chulalongkorn University and then Mahidol University, Kasetsart University, and Chiang Mai University respectively. All nine universities are received research fund 9,000 million bath in 3 years (3,000 million bath per year) (Chiang Mai University will be the national research university, 2552). National Research Universities criteria are based on the ranking of World Universities 2008

arranged by Times Higher Education - Quacquarelli Symonds (THE-QS). THE-QS ranked CMU in range 401-500. Especially Art & Humanities is ranked at number 270 and Natural sciences is ranked at number 283 (Chiang Mai University, the national research university, ready to be World Class University, 2009; QS Quacquarelli Symonds, 2008).

CMU has been pushing academic and staff to do more research works and there is also an effort to establish a research cluster which the university is specialized. This can be done by pulling academic who has a well-known work and is accepted in their field and also new generation researchers together for purposing innovation research topics and push the works of the university to be as equal as other top ranked universities in Asia (Chiang Mai announces, 2006).

However, it is quite difficult to establish a research cluster because there is not complete information that available to indicate the expertise and the research direction of CMU researchers. This is because all information has been kept separately in each faculty and many faculties have not completed the information yet (Kangwanpong, 2005).

Besides doing research inside the university, CMU also efforts to establish coordinated research and develop human resource with other universities both inside and outside the country, including organizing a learning center and research with universities in the Mekong river region (Kangwanpong, 2006).

For supporting research process, CMU administrators established research management center with responsibility in building the research management system, which is efficiency, above-board, and accountable. CMU research management center

is the main research information sources of CMU. It also support other units in many services as follow (Research management center, 2008).

a.) Support research policies and research directions.

b.) Develop database for research management.

c.) Support intellectual property organization.

d.) Allocate research fund ex. research fund of new research development.

e.) Proceed any work about research fund and research sign contract.

f.) Coordinate between granters outside university and researchers.

g.) Publicize about research result and research fund.

h.) Coordinate between inside and outside units.

i.) Support expense of publishing paper.

j.) Publish research journal of CMU such as CMU Journal and CMU Abstract.

For dissemination of research results to communities and society, CMU administrators planed to (Kangwanpong, 2006)

a.) Set up Intellectual property unit concerning intellectual property organizing (cooperates with Department of Intellectual property).

b.) Establish new units, Chiang Mai University Business Incubator (CMUBI). CMUBI focuses on linking research & development with business requirement. It also organizes intellectual properties and transfer information technology to intellectual properties rights system. New entrepreneurs enable to join with

CMUBI as a Start-up Company using knowledge and new ideas of CMUBI to run their business (CMUBI, n.d.).

c.) Create cooperation with Northern Science Park.

There are many attempts CMU use in research management. However it can not be done easily because there are some obstacles. CMU research information is incomplete and dynamic so decision making and planning about research such as research granting and creating COE cannot do efficiency. This makes CMU do not know research directions and expertise of CMU researchers.

2.1.1.6 University Researcher Management.

Research managers of the university should facilitate researchers by creating a productive environment as same as other research institutions. However, the managers should also consider these things additionally.

a.) Balance lecturers' work load between teaching and researching. It should be careful that they shouldn't pay too much attention in research works or neglect their lessons (Monsted, 2002). In other words, lecturers who prefer teaching than doing research or have lots of teaching work. They might be encouraged to do research that reinforce the lessons in order to benefit both ways.

b.) Doctoral recruitment should be more increased for research assistants. It will also be a good opportunity for students to improve their skills and gain more knowledge from skilful researchers (Chang, & Trubow, 1990; Conraths, & Sodergard, 2004; Monsted, 2002).

c.) Research activities are also necessary for researchers to exchange knowledge and opinions. The activities should be held regularly. For

instance, academic conference should be held at both nationally and internationally, including publicizes research works in academic journals both inside and outside the university.

If researchers have sufficient supports and facilities, they are able to work with satisfaction after that quality results are well achieved.

To motivate researchers to doing higher quality research, CMU research administrators encourages in many ways as follows (Kangwanpong, 2006).

a.) Create a research environment by organizing academic days annually around December.

b.) Create research reinforcement by giving research award named *Gold Elephant award*.

c.) Support expense of publishing academic papers and representing in conferences.

d.) Push research direction according to skill of researcher faculties and trend of international research and development.

e.) Develop research potential by founding center of excellence and finding research niche for each faculty.

f.) Create innovation by founding CMUBI.

g.) Develop research persons by giving research budget for new researchers.

h.) Develop multi-discipline research.

In spite of CMU tried to support doing research in various ways but some researchers, which are high capability in doing research but do not well known by

administrators might be neglected. This might be caused of incomplete and dynamic of research information in database system and MIS of CMU.

2.1.1.7 Research Management Framework in Thailand.

Research management system in Thailand consists of networking and systematic coordination between units, which are researchers, research departments, fund management, and research policy units. There are 4 levels of research policy units in Thailand that involve in research (Thailand Research Fund, 2005).

a.) Policy level: In this level, Nation Research council of Thailand (NRCT) is responsible for development the nation's research policies by law.

b.) Research Fund management level: This level is a unit responsible for allocating fund to research departments. There are many units in this level:

- Thailand Research Fund (TRF) is a unit under the Office of the Prime Minister but it is an autonomous unit. Its duty is to allocate money for every field includes policy research.
- Office of Agriculture Research and Development (OARD) is a private organization, established follow the loan project of Asia Development Bank (ADB). It has a role in supporting research in the Agriculture field, which emphasizes on commercial purpose.
- National Science and Technology Development Agency (NSTDA) is an autonomous unit but under the Ministry of Science and Technology. It supports research funds in the field of Science Technology and

emphasizes on 4 technologies which are Biotechnology, Information Technology, Materials Technology, and Nanotechnology.

- Health System Research Institute (HSRI) is an autonomous unit under the Ministry of Public Health. It supports research funds in the public health system development.
- Research unit levels: This level is any organization has responsibility to do research. This level has a significant role in doing research but variety aims such as academic use, commercial use, public use, and different policies. The example of this level is specific institutions and universities.
- Research user level: This level is important level, which research unit level has to concern. The examples of this level are companies, government, private sectors, and industries.

Those 4 groups have to well define their duties clearly to avoid overlapping. Besides that, they have to be well coordinate to create research works which are practical and beneficial to the nation.

2.1.1.8 Evaluation Research Productivity

Quality evaluation is an important process that a research unit can build its capability. This should be made through researchers, which can be evaluated both directly and indirectly. An indicator can be determined variously depends on the objective of the evaluation and the nature of the researchers' work style in each organization. However, indicators in organizations can be similar to each other. See from the following example.

Example 1: The top hierarchy of the Ministry and the top of the universities is used in evaluating researchers in Denmark. (Armstrong, 1995) The indicators are:

a.) Reputation.

b.) More international research publication.

c.) More external funding.

d.) More and better teaching.

Example 2: This was developed by Armstrong and Sperry (1994), used in the MBA program. The indicators are (Armstrong, 1995):

a.) Academic publication of instructor.

b.) Citations by others of the faculty's work.

c.) Evaluation of instructors by peers.

d.) A prestige rating.

Example 3: KPI of CMU can be divided into 4 aspects (Viriyajaree, 2002):

a.) Create strong research teams at international level and strengthen research quality by coordinating with lecturers in the graduated level.

- The number of Center of Excellence.
- The proportion of lecturers per the total number of lecturers.
- The percentage of research work and creative work, be published, distributed, or being used nationally and internationally per lecturer.

- The percentage of research articles, being cited in refereed journal or in databases both national and international per lecturer.
- The number of patent, which comes from research work and creative work, registered.
- The percentage of lecturers, received research funds from both inside and outside the university per the total number of lecturers and researchers.
- Percentage of lecturer being a thesis advisor outside the university, and a member of academic committee, or professional committee in both national and international level per lecturer.

b.) Produce research works bases on niche and a kind of knowledge support the country's competitiveness.

- The number of research work, which reflects the use of niche and can be competitive.
- The number of nationally and internationally reward-winning research works.

c.) Move forward in encouraging integrated research works by coordinating with public and private sectors to meet the society's needs.

- The percentage of multi-discipline research fund with public and private sectors per total annual government statement of expenditure.

d.) Organize the management system and manage the multi-discipline research to be valuable and efficient.

- The amount of research fund from both inside and outside the university per lecturer.
- The amount of research fund from outside the university per lecturer.

Even though the indicators of each organization are various, depends on the aim of the evaluation, it can be seen that the use of international research publication to evaluate researchers is another popular choice because it can be done simply due to the availability of supportive evidences in commercial databases. However, at the same time, there is a question on what journal should be used for evaluating researchers in each field, organization and university (Monsted, 2002).

KPI of CMU covers and intensives for evaluating researchers and research but evaluating cannot do efficiency because the same problem of research management. That is incomplete and dynamic or search information.

2.1.1.9 Information and Telecommunication Technology (ICT) in Research Management

The application of information technology and communication can make the research process and management easier and be more efficient. This can be done by different ways:

a.) Portal web development is a specific way for researchers to access information resources both inside and outside the organization such as data for studies, information about research fund sources, and research network.

b.) The use of Weblog or blog enables researchers to establish a social network for discussing and exchanging opinions with other researchers in the same field.

c.) The use of electronic databases to find articles in each field to follow advancement in the field (Castellanos & Rodriguez, 2004).

d.) An establishment of research repository or research management system as database system and management information system (MIS) to benefit in various ways:

- Check for historical data.
- Use for referring to research works.
- Use for scholarship application.
- Disseminate research result to others researchers in the organization.
- Be part of a decision-making process in the research management.
- Use for planning the research management in the future

Although Technology brings a lot of benefits; however, if it is used inappropriately, it will not be as beneficial as it should be. To illustrate, if researchers resist the use of IT and do not want to communicate via internet survey or the Management Information System has not been used or updated, including containing incomplete data. That MIS will not be able to be used for making a decision as it should be. Thus, besides choosing proper technology, there should also be a critical plan in the way to make technology usage gives the most benefit.

In CMU, there are many research databases and management information systems. However there are developed separately in each faculty so searching is not

quite efficiency. Further, much of research information do not input cause of many reasons such as researchers do not have time to input or they cannot remember all research data or they reluctant to input data. That makes analysis of research data is bias.

2.1.1.10 Knowledge Management in Research Management

There are lots of organizations using Knowledge Management in managing research works in several ways. The management emphasizes on researchers, considered as a knowledge worker who are skilful and have experiences in doing research. Applying knowledge and innovation from research into organization's tasks in the most benefit is the goal every organization wants to achieve. However, there are several ways of applying knowledge management in research management.

a.) Motivate skilful researchers to share their knowledge to other researchers by establishing research repository and portal, which is a way to share knowledge and exchange opinions among researchers. This makes the whole research works of the organization be developed in the same time (Kidwell, Linde, & Johnson, 2000).

b.) Create research management system to preserve information about research such as research proposal, publication, and research funds. The system might be database or MIS (Management Information System) to be used for accessing retrospective data, making decision, and planning research management (Davey & Tatnall, 2007; Leung & Low,2005; Palomo, Veloso & Schmal, 2007).

c.) Strengthen decision-making ability in research management by searching for the key drivers of key success factor. Those key drivers can be found by

analyzing Skandia model, which is used in classifying intellectual capital (Castellanos, & Rodriguez, 2004).

Knowledge management is a new framework that still ambiguous and has various ways of use, which make it is an unstable form to apply to use in organizations. For the application of knowledge management in research management, it can be prospected from different aspects. However, it is considered as a new trend in organization management which is not focus only on financial capital but also intellectual capital. More than that, ICT is another important thing in managing knowledge and knowledge workers more efficiency. However, knowledge manager should not focus on ICT rather than manage researcher. Otherwise ICT investment will be waste of money.

With use knowledge management in research management, it enables research management done more easily and efficiently. Then universities can handle with big change in risk management. Further, tools and techniques of knowledge management enable to synergize university research activities as well.

2.1.2 Intellectual Capital Theory

2.1.2.1 Important and Role of Intellectual Capital

At present, organizations realize that capitals of organizations are not only financial capital but also intellectual capital. In the past, asset consideration had emphasized on lands, labours, money, and equipment. These assets were considered as hard assets which means physical and monetary. Thus if focus on Microsoft Company, its net asset value and market value are very different. It can be seen that net asset value cannot be practically used for measuring because it estimates just the

physical capital. But, in fact, Microsoft's assets are in other forms such as knowledge, brands, innovation projects, and invisible assets which are considered as an intellectual capital (Castellanos, & Rodriguez, 2004; Roos, Roos, Dragonetti, & Edvisson, 1997)

Intellectual capital is invisible assets which usually occur in the information flow within the company and also between the company and external variance. If there is a vast amount of use, those kinds of intellectual capital can be enhanced. Finally, it is also a longer-lasting source in a competition more than financial asset.

Additionally, intellectual capital is the core competency of the organization. Core competency is a set of skills and technologies of the company which differentiates the company's value from other competitors. This enables the company to enter into new market. Thus, core competency can be considered as a competitive advantage. As can be seen that core competency is not an asset but activities or accumulation of learning (Bueno, Salmador, & Rodrı´guez, 2004).

From the reasons above, many organizations are focusing on intellectual capital to be the real indicator of their asset because it can enhance the value of organization in new point of view.

2.1.2.2 Intellectual Capital Definition

Intellectual capital was used first time in 1997 by Professor Lief Edvinsson. He published the real asset of Skandia in its annual year report in terms of both hard asset and intellectual asset. He is corporate director of intellectual capital of Skandia Company, an insurance and financial service in Stockholm, Sweden.

It is an essential thing that organizations should clarify what intellectual capital is. The meaning of intellectual capital is quite hard to specify because intellectual capital itself has different forms and is taken place in many operational processes. The examples of intellectual capital are as follows: -

a.) Processes.

b.) Innovation.

c.) Company infrastructure.

d.) Knowledge/skill of the employee.

e.) Trademarks.

f.) Patent.

g.) Brands.

h.) Relationships between customers/ partners.

i.) Etc.

From the example above, intellectual capital is not only pure intellect but also intellectual action. Occasionally, it may be dynamic that can be change to another form. It can say intellectual capital is an aggregated expression of the intangible asset possessed by the organization (Bueno, Salmador, & Rodrı´guez, 2004). However, some intangible asset is not the capital of organization because they might be an asset but they do not generate income when time past. Hence for specific intellectual capital meaning, intellectual capital can be described as a language and set of techniques, which is used in predicting future earning capabilities (Roos, Roos, Dragonetti, & Edvisson, 1997). In other words intellectual capital is the capital of an organization which is priceless but valuable. Intellectual capital is an implicit capital which is embedded in knowledge workers in many forms such as expertise, skill, and work

experience. In this study, intellectual capital means CMU researchers, research competency, researches, and research outcome such as patents, publication, research award, etc.

2.1.2.3 Intellectual Capital Models

Intellectual capital model is a tool using to identify, measure, and manage intellectual capital. Due to the variety of intellectual capital, there should be some techniques to manage those intellectual capitals. However, there is a statement said, *anything could be managed it can be measured and in the other hand, if anything could be measured it can be managed*. Even though management and measurement is important for intellectual capital, but management and measurement will be done if intellectual capital is identified. Thus intellectual capital identification is very important to be clear (Castellanos & Rodriguez, 2004). So there are many intellectual capital models developed for this purpose.

Intellectual capital models can be divided into two groups. The first group is key dollar valuation such as EVA and MVA model, Market-to-Book Value model, and Knowledge Capital Earnings model. This group of intellectual capital model tries to measure intellectual capital in term of money for adding value to organization. However this group is not suitable for managing intellectual capital and it can not guarantee that calculated money value is correct and precise. The second group is called key non-dollar valuation. The famous examples of intellectual capital models in this group are as follow (Tan, Plowman, & Hancock, 2007).

1.) Balance Scorecard, developed by Kaplan and Norton in 1992.

2.) Skandia Model, developed by Edvinsson and Malone in 1994.

3.) Technology Broker method, developed by Brooking in 1996.

4.) Intangible Asset Monitor approach, developed by Sveiby in 1997.

5.) Intelect model, developed by Euroforum in 1998.

6.) Knowledge Management Consortium International Model, developed by Mc Elroy in 2001.

1. Balance Scorecard

The Balanced Scorecard is developed by Kaplan and Norton in 1992. It is tools for executives to understand status of organization in many perspectives not only financial perspective as same as traditional tools. Further executives can use Balanced Scorecard for managing company by balance all four perspectives. This makes formulating organization vision balanced in all perspectives include intangible asset.

The Balanced Scorecard can measure capacities of company in four perspectives; financial perspective, customer perspective, internal business process perspective, and learning and growth perspective as shown in Figure 2.1

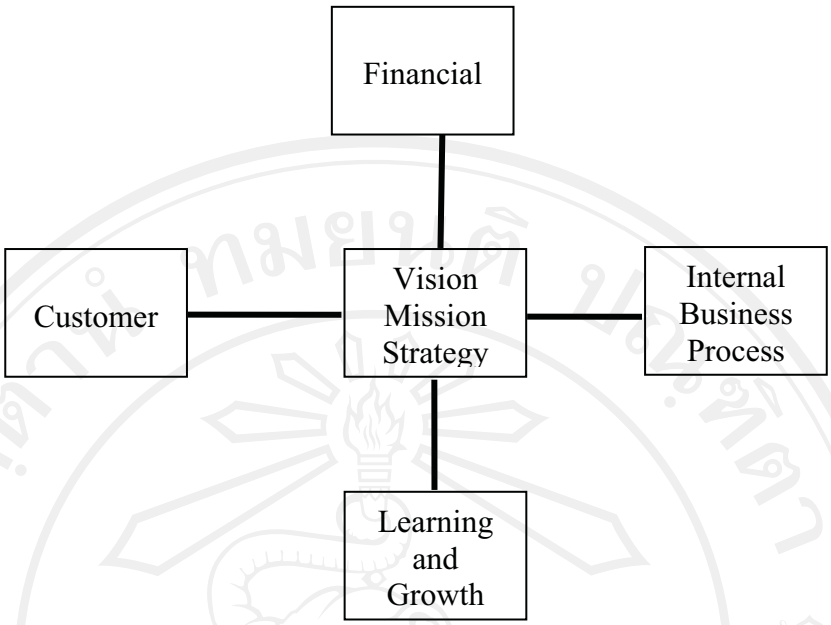

Figure 2.1 Balance Score Card Model

The Balanced Scorecard translates an organization's mission and strategy into a set of performance measures that used for measurement and management system. With using Balanced Scorecard, executives can manage how they enhance internal business process and how invest in people, systems, and procedures, which are necessary to improve future performance. In addition, Balance Scorecard leads executives to emphasize with customer relation that increase the competitiveness of the organization and share more market quota in the future (Kaplan & Norton, 1996).

Even the balanced scorecard try to measure non-financial perspective but it does not focus much more in some kinds of intellectual capital such as patent, innovation, brand, trademarks.

BSC enables to apply in research management in universities for monitoring and managing universities' intellectual capital in all aspects both financial and non-

financial. Finally the value of indicators in BSC refers to the sustainability of universities.

2. Skandia Model

Edvinsson and Malone developed Skandia Model in Skandia Company in 1994. This Company is the pioneer, who identifies intellectual capital by using an intellectual capital model. Then Edvinsson and Malone include Skandia Model in company's annual report. Skandia model was adapted from Balance Score Card (BSC) and well-known in academic area (Roos, Roos, Dragonetti, & Edvisson, 1997).

This model is an extremely valuable guide to how to measure the exactly corporate wealth to reflect the new knowledge-based economy. It also can be used to set the strategic direction for knowledge sharing and then determine its impact (Ives, Ben & Cindy, 2001).

Skandia Model provides five categories of IC: financial capital, human capital, customer capital, process capital, and innovation capital as shown in Figure 2.2.

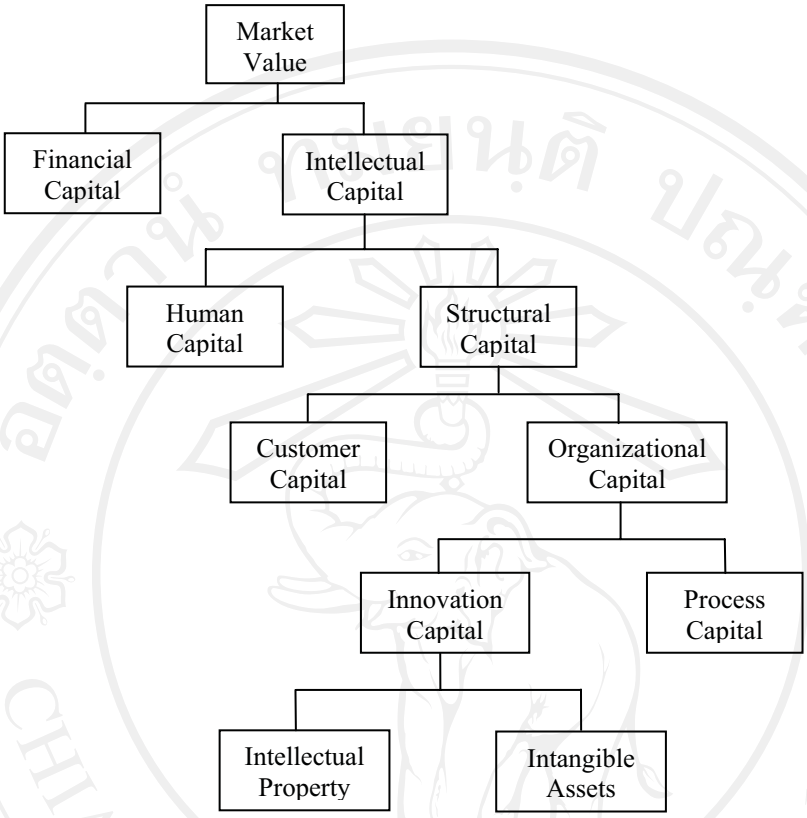

Figure 2.2 The Skandia model (Roos, Roos, Dragonetti, & Edvisson, 1997)

Simultaneous management of intellectual capital and financial capital will complete the management system, which is called holistic management strategy.

The Edvisson and Malone divides the value market into 2 parts (cited in Chen, Zhu, & Xie, 2004; Roos, Roos, Dragonetti, & Edvisson, 1997; Vilanova & Josa, 2003).

1. Financial capital: physical and monetary assets.
2. Intellectual capital: invisible process and assets.

2.1. Human capital: all individual capabilities, knowledge, skill, and experience.

2.2. Structure capital: None thinking part or everything left after people go back home such as brand, trademarks, written procedure for process.

2.2.1 Customer capital: participants or relationship between customers including customer loyalty and customer satisfaction.

2.2.2 Organization capital: variable that support organization routines and processes such as manuals, best practices, intranet resources, project libraries.

2.2.2.1 Process capital: know-how of organizations that enhance the efficiency of manufacturing or the delivery of services.

2.2.2.2 Innovation: anything create profits in the future or renewal capability.

2.2.2.2.1 Intellectual property.

2.2.2.2.2 Intangible assets.

It is noticed that Human capital and structure capital are separated due to the different managements. However, customer capital or relationship were generated by all good relations with other participants in the external environment to the company i.e. customers, suppliers, and alliance partners (Roos, Roos, Dragonetti, & Edvisson, 1997).

However this model still lack of relationship between the company and the social agents (Bueno, Salmador, & Rodrı´guez, 2004), so some kinds of intellectual capital still looked over.

For organizing intellectual capital, Edvinsson (1997) developed and applied a tool called Skandia Navigator. It shows where all types of knowledge flow through the company. The model has shown in Figure 2.3.

Figure 2.3 similars to a house. It is a company visual metaphor. The upper triangle means the company's history, i.e. balance sheet situation. In the middle, there are the walls which symbolize the present activities i.e. all company process and customer relationship. On the base of a house, it is innovation of the company that will create the company's future. However, all things happen in history, today, and tomorrow will be done by human or knowledge workers of the company. These focus areas are the origin of the firm's intellectual capital value inside its competitive framework (Roos, Roos, Dragonetti, & Edvisson, 1997; Vilanova & Josa, 2003).

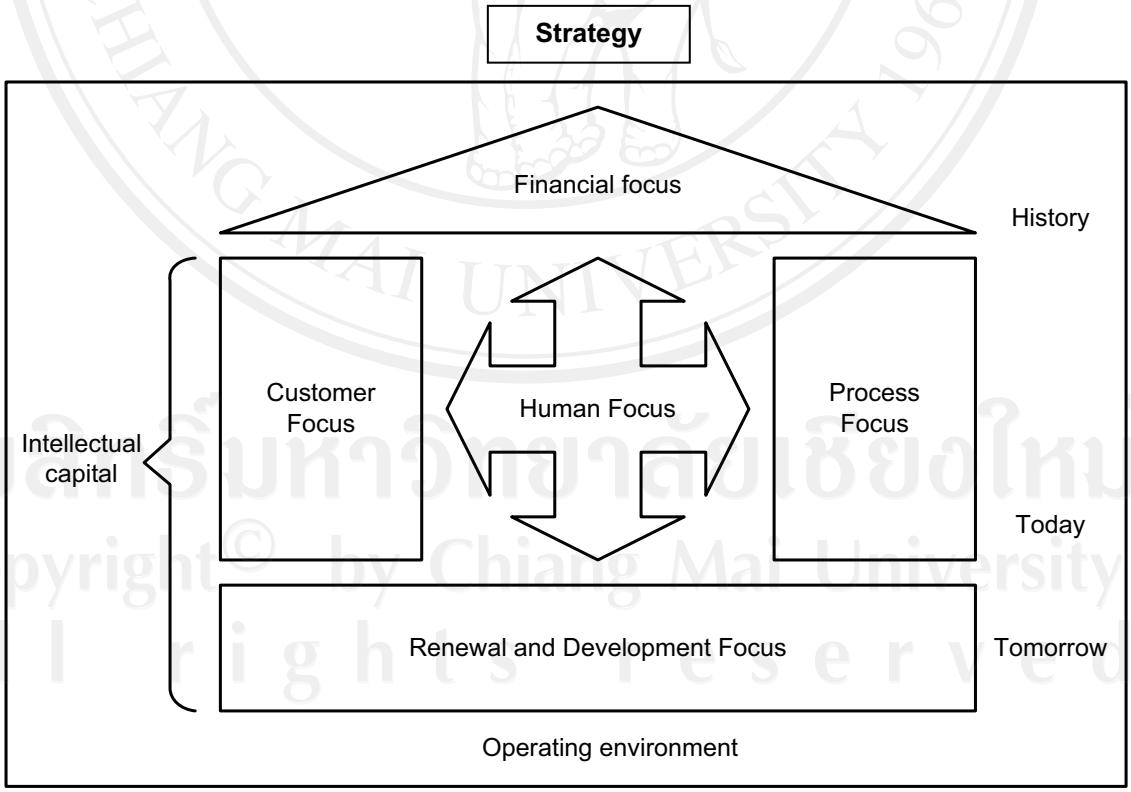

Figure 2.3 Skandia Link Navigator (Roos, Roos, Dragonetti, & Edvisson, 1997)

This study uses Skandia Model in research management of universities as a tool for representing the value perception from stakeholders. With looking at value of indications in each subgroup, stakeholder will be pleasured if all indicators' value responds to their need and their requirement.

3. Technology Broker method

This method was developed by Brooking in 1996. It shows that intellectual capital formed by four groups of asset as follow (Roos, Roos, Dragonetti, & Edvisson, 1997):

a.) Market assets: brands, customers, customer loyalty, distribution channels, backlog, and soon.

b.) Human-centred assets: skill and expertise, problem-solving ability, leadership styles, abilities, and everything that is embodies by the employees.

c.) Intellectual property assets: any intangible which can be protected by copyright.

d.) Infrastructure asset: technologies, methodologies enabling a company to function.

4. Intangible Asset Monitor approach

The model developed by Sveiby in 1997. This approach assumes a three-way distinction between intellectual capitals as follow:

a.) External structure.

b.) Internal structure.

c.) Employee.

For implement these two models in research management in universities, Technology Broker method and Intangible Asset Monitor approach, do not separate between thinking part and non-thinking part such as reputations and grantors. Technology Broker method classifies they in market asset and Intangible Asset Monitor approach classifies they in external structure. Meanwhile, even two intellectual capitals put in the same group, but they need to manage in the different ways. This makes the model is not practicable in a real world.

5. Intelect model

In 1998, Euroforum was developed new intellectual capital model called Intelect model. The model has shown in Figure 2.4. This model is modified from Skandia model. The Relational capital is adjusted from customer capital in Skandia model. The enhanced definition of relational capital is the value of the set of relationships between the firm and its environment. However, main indicators of this capital still refer to customers and other indicator refers to aspects such as reputation, strategic alliances, suppliers, and connection to other agents.

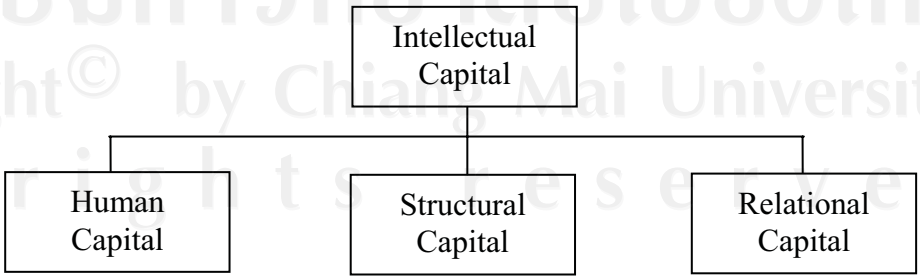

Figure 2.4 Intelect model (Euroforum, 1998)

7. Knowledge Management Consortium International Model

This model was developed by Mc Elroy in 2001. This model is revised relational capital to be social capital as shown in Figure 2.5.

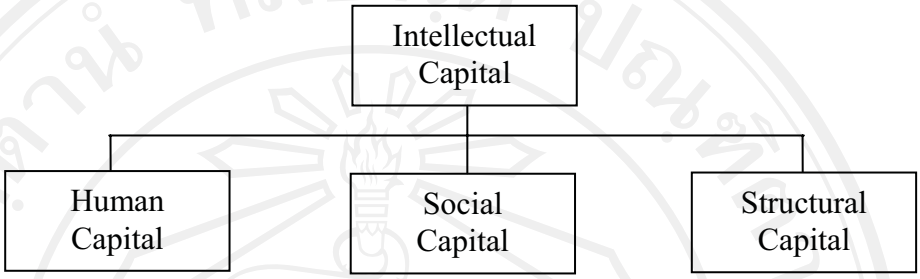

Figure 2.5 Knowledge Management Consortium International Model (Elroy, 2001)

The social capital emphasize on relationship of company with other agents both inside and outside the company. Therefore, this model divides social capital into three categories (Bueno, Salmador, & Rodrı´guez, 2004):

a.) Intra-social capital: relationships between employee in company such as community of practice.

b.) Inter-social capital: basic relationship with clients and shareholders.

c.) Social innovation capital: the set of process and relationships that could be create innovation, which is a key success factors for competition.

Even both Intelect model and Knowledge Management Consortium International Model are adopted from Skandia model but when apply it to research management in universities, it is quite hard. The first reason is structural capital is very complex without subgroups users might cannot identify complete set of intellectual capital. The second reason is relational capital and social capital are very subjective the way to

measure this kind of intellectual capital might be difficult and the output of this might be bias as well.

2.1.2.4 Intellectual Capital Indicators

For intellectual capital measuring and managing, intellectual capital indicators were focus. Intellectual capital indicators can be made up by follow the classification in intellectual capital model. For increasing organization competency, executives should manage intellectual capital based on value of intellectual capital indicators. However if executives try to write all intellectual capital indicators for their company, it becomes a long list. Thus it is not necessary to consider all of them. The identification of intellectual capital indicators should be based on the organization's needs, which might consider from the organization's strategies and mission. The strategies of the organization indicate the aim and direction, which inform the indicators that are used in the organization's performance evaluation. Those indicators signify the intellectual capital forms that the organization should examine and mange. For instance, the goal of CMU is to be a research university, thus, the intellectual capital that should be examine is the number of research works per year. On the other hand, if CMU aims to be a qualitative research university, the number of awarded-winning research per year should be viewed.

Besides, indicators can determine by the characteristic of the organization and its day-to-day operations. Anywhere executives should choose indicators, which impact with market value.

2.1.2.5 Intellectual Capital of Researcher in Knowledge Management for Research Management in CMU

CMU is the first institute of higher education in the north and the first provincial university in Thailand. It was establish since 1954. CMU is aiming at being the excellent university and is going to be the world class university in the future (CMU 45 years, the 5th decade of the excellent university, 2009). Therefore CMU has to more concentrate on research. CMU also has various strong discipline, research areas and well-known researchers. In universities, researches and researchers can be accounted as intellectual capital, which should be considered and measured in order to increase efficiency of university research management. CMU have 17 faculties, 3 research institutes, 1 graduate school, 2,165 lecturers and more than 900 Ph.D. lecturers (Chiang Mai University, 2002). There are 123,679 students graduated from CMU; 472 doctoral degree, 20,513 master degree, 96,346 bachelor degree, and 3,635 diploma degree (CMU the 5 decade, the excellence university, 2009).

CMU continually develops research projects for many years in various disciplines and also received funding from many organizations both public and private sectors (Chiang Mai University, 2002). There are many research publications (more than 2,000 papers) both national and international from CMU lecturers and researchers (CMU the 5 decade, the excellence university, 2009).

Some researches create innovation such as patents and prototypes which are valuable for mankind. Furthermore, CMU has high potential researchers who received awards and reputation (Chiang Mai University, 2002).

In the universities, researchers, especially key researchers are knowledge workers who have done research in many years and have experiences and skills in research which are very valuable for the organization. In order to increase the efficiency of the university's research management, universities need to understand their own expertise and research directions.

In research management, CMU research administrators have to set up ten years road map. However this is difficult because research administrators do not know expertise and research directions of key researchers. Thus, research supplement and encouragement from CMU cannot arrange for enhanced research process efficiency in the right way. Administrators do not know this information because of two reasons as follows.

a.) Research administrators do not know who are key researchers. This kind of researchers is high potential person who will drive CMU research results gloriously and continually. These researchers need to advocate in term of research policies and financial aspect including cooperation of other researchers to set up research team. The way for identifying key researchers need to evaluate from research performance of these researchers in the past. The research performance can be in many forms such as number of research project, number of publication, number of patent, and so on. These entire things can be count as IC, which could be return to be money in the future. However, collecting this data is a problem because information about CMU research is incomplete. The reason is research information was store in many databases separate in each faculty and sometimes pulls in a university database as well. Therefore, research information always redundancy. Besides, research information is not up-to-date. Because researchers are not input data into databases in

many reasons; they do not recognize the benefit of giving their research information, they are reluctant to fill up, and they have no time for this thing.

b.) Research administrators don't know expertise and research directions of CMU researchers. Because research information is dynamic. Many researchers may change their research interest and topics from time to time due to many factors, for examples their supportive fund, their research resources, and new academic trends.

In this study intellectual capital theory could be used for identifying, collecting, and measuring research results, which are intellectual capital of universities. This was increase the efficiency of managing key researchers. The intellectual capital model used in this study is Skandia model because this model divides intellectual capital into subgroups exhaustively and clearly.

2.2 Artificial Intelligence Technique

2.2.1 Artificial Intelligence (AI)

The term AI was introduced by McCathy in 1956. AI is an interdisciplinary field of study that emerged in the 1950s from origins in computing, engineering, psychology, mathematics, and cybenetics. (Nikolopoulos, 1997).

The objective of AI is to build systems, exhibit intelligent behaviour and perform complex tasks with a level of competency that is equivalent to the level currently exhibited by human experts. There are two methods of AI 1) symbol manipulating and 2) the connectionist approach. The symbol manipulating uses logic for representing knowledge, reasoning, and deducting. The expert system was developed from this approach.

In this study many AI techniques, e.g. ontology, are used for handling research information, which is incomplete and very dynamic when time pass. Further some complexity problem in research management, such as the requirement of administrator in universities level and faculties level for identifying key researchers are different, need more efficiency techniques of AI to solved e.g. predicate logic and inference step.

2.2.2 Expert System

Feigenbaum defined an *expert system* as "an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human. The knowledge of an expert system consists of facts and heuristics." (Nikolopoulos, 1997).

However, an expert system is used in some case not for general problem solver. It always uses for solving in specific areas such as expert system for diagnosing infectious blood diseases, and internal medicine problems.

An expert system performs like an expert in a domain and use heuristics rather than the algorithmic approach likes a computer program. Thus an expert system emulates the decision making and problem solving capabilities of a human expert in his/her area of expertise.

There are four architectural components of an expert system as show in Figure 2.6 , detail are Knowledge acquisition module, Knowledge bases, Inference, and Input/output interface (Nikolopoulos, 1997).

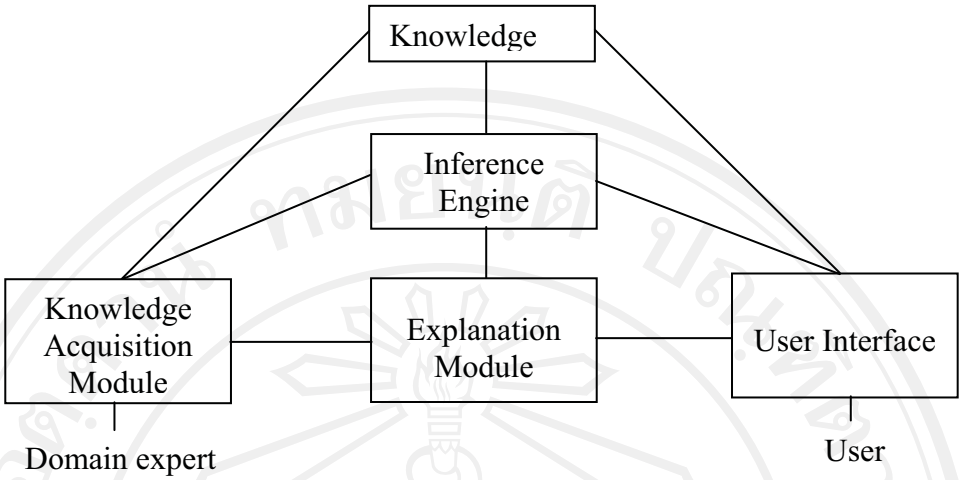

Figure 2.6 The architecture of an expert system (Nikolopoulos, 1997)

a.) Knowledge acquisition module.

Knowledge acquisition composes of knowledge elicitation and knowledge representation. *Knowledge elicitation* is about the acquiring the necessary knowledge and creating a conceptual model of this knowledge in order to solve the problem face with. *Knowledge representation* is the process of choosing an appropriate format for representing knowledge in computer system.

b.) Knowledge base.

Knowledge base contains the specific knowledge and control knowledge, which is used to solve problem in specific domains. In creating knowledge base, expert system can be elicited knowledge from experts directly or it can be learned by the system itself. Knowledge in expert system can be elicited from domain experts and other sources e.g. books, research report, etc. For automating knowledge acquisition, the system itself uses machine learning techniques to directly learn the knowledge from exist knowledge and new knowledge.

The knowledge in knowledge base have to converted in a format suitable for computer manipulation in a process called *knowledge representation*. There are many kinds of knowledge representation. The detail is address in topic 2.2.5.

c.) Inference engine.

The *inference engine* contains general algorithms, which are able to manipulate the knowledge stored in the knowledge base in order to solve problem. There are many inference techniques used depend on the knowledge representation chosen.

The typical inference engine is based on an inference rule and a search strategy. An *inference rule* is a way to deduce new knowledge from the existing knowledge base. Process of inference step has shown in Figure 2.7.

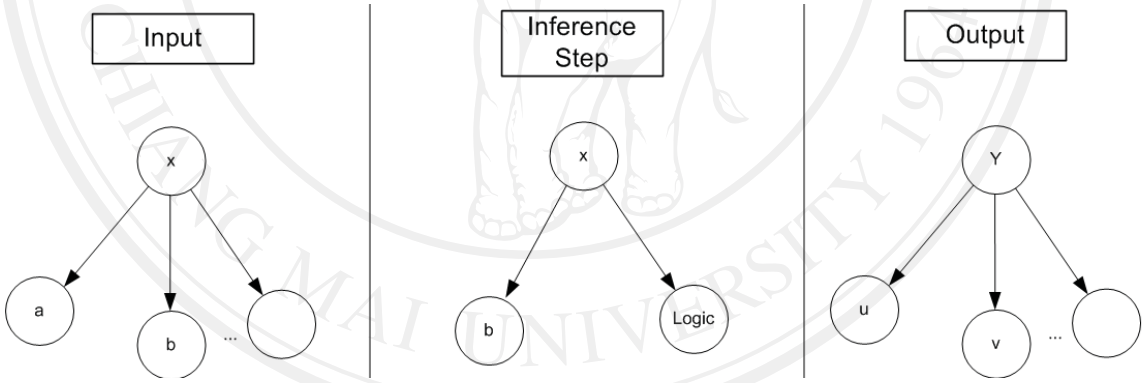

Figure 2.7 Inference process

First of all, a set of information will be entered to expert system. In diagram x is input and a, b, … is subset of input. Information is stored in form of many kinds of knowledge representation format such as predicate logics, semantic networks, ontologies. On the inference step, some technique will be applied for reasoning ontologies. There are many techniques for this kind of work such as forward and

backward chaining, predicate logic, etc. The technique should be chosen depending on characteristic of input. In diagram x is enter as query for inference or reasoning new knowledge from input b by using technique such as predicate logic. The output of inference step is new knowledge (u, v, …), which is reasoned by expert system.

d.) Input/Output interface and explanation.

Input/output interface is the point of interaction or communication between a computer and any other entity, such as a printer or human operator ("Interface", 2002) or the way the expert system interacts with the user. It usually shows in graphical mode. In addition, the interface also provides a way to connect to other systems such as DBMS (Database Management System) and XML (eXtensible Markup Language) (Giarratano and Riley, 1993; Nikolopoulos, 1997).

2.2.3 Theoretical Foundations

Some kind of mathematic theories are used in expert system for knowledge representation (Nikolopoulos, 1997).

2.2.3.1 Propositional Logic

Logic formulas are constructed by combining *propositions* with the basic logical operators. A proposition is a statement having value of either *true* or *false*. The example of statement is

There_are_twenty_four_hours_a_day

The value of this proposition is *true*.

The basic logical operators i.e. negation (~ or ¬), conjunction (∧), and disjunction (∨).

Other logical operators, such as implication ($\rightarrow$), equivalent ($\leftrightarrow$), and logical equivalence ($\equiv$), are reducible to the basic operators.

For example,

$A \rightarrow B$ is logically equivalent to $\neg A \vee B$

$A \equiv B$ is logically equivalent to $(A \rightarrow B)(B \rightarrow A)$ or $(\neg A \vee B) \wedge (\neg B \vee A)$ OR $(\neg A \wedge \neg B) \vee (A \wedge B)$

However, propositional logic is unable to show the concept of an individual object, object properties, and relations between objects. For example, the inference chain is

("all dogs are animal", "there is at least one dog", therefore "there is at least one animal")

This chain cannot be expressed in proposition logic.

2.2.3.2 First order Predicate Calculus and Predicate Logic

a.) First order predicate calculus

First order predicate calculus extends the expressive power of propositional logic. Moreover, it can be used as a formal language to express declarative knowledge.

A predicate calculus formula consists of:-

- Predicates.
- Predicate arguments.
- Logical operators.
- Quantifiers.

Predicate is a relation mapping objects/argument in the universe of discourse into *true* or *false*.

For example:

Mother(Mary, Mane)

This predicate is true if *Mary* is the mother of *Mane*.

Mother here can be defined as a predicate logic of two arguments.

This predicate has two *arities*. An arity is the numbers of argument in predicate. A predicate argument can be *constant* representing objects from the domain knowledge or *variables*. Constant is usually used small letter to be the first letter, where as variable used capital letter.

An atomic formula is a predicate of a number of arguments. Predicates are connected to each other to form logical sentence using the *operators* or *logical connectives*. The operators are and ($\wedge$), or($\vee$), not($\neg$), implies($\rightarrow$), and quantifiers. The quantifier is the way to quantify the variable predicate arguments. There are two types of quantifier. The first one is *existential quantifier*, denoted by $\exists$. The second one is *universal quantifier*, denoted by $\forall$.

The format for using universal quantifier is

$\forall X\ p(X)$

The semantic interpretation of this sentence is “every object X in the domain of discourse, X has the property *p* or makes the logical *p* true”.

For example:

“All human are mammal”

This can be written in predicate calculus as:

$\forall X$ (human(X) $\rightarrow$ mammal(X))

The format for using existential quantifier is

$\exists X$ p(X)

This mean an object X, (at least one), in the universe of discourse, which satisfies the property *p*.

For example:

"There is dog has two legs"

This can be written in predicate calculus as:

$\exists X$ (dog(X) $\wedge$ twolegs(X))

b.) First order predicate logic, functions.

First order predicate logic is as same as first order predicate calculus except adding *functions*.

The function is a relation matching objects from the domain of discourse to other objects. The function can use instead of a predicate argument. For example,

There are two predicate logics

Mother(X, Y)

Wife(X, Y)

and one function

father(X)

The way to use function in logic statement is

$\forall X \forall Y$ (Mother(X, Y) $\rightarrow$ Wife(X, father(Y))

This state that "if X is the mother of Y, then X is wife of the father of Y"

c.) Well formed formulas.

An atomic formula or its negation is called a *literal*.

For example if the atomic formula is

dog(X)

the negation of this formula is

¬dog(X)

Well-formed formula (WFF) is statement connecting of the literals, logical operators, and quantifiers. A well-formed formula is recursively defined as follows:

1. A literal is a WFF
2. If *g* and *h* are two WFF, then *g* <operation> *h* is also a WFF, where <operation> is a logical operator.
3. If *g* is a WFF, then ∀X(g) and ∃X(g) is a WFF.

2.2.3.3 Inference

Inference is the process of analyzing logical statement to derive the new knowledge. In the inference process, each inference step must be based on a logically valid rule of inference.

The inference process is as follow:

1. Give a set of well-formed logical formula, called the *knowledge base* or the *set of axioms*.

2. The current state of the knowledge base is transformed into a new state with additional WFFs inferred by using a "rule of inference".

Example I if the set of axiom are

$A \rightarrow B$

A

It can be inferred that B is true. Thus the conclusion is B.

Example II: if the set of axiom are

$A \rightarrow B$

$\neg B$

It can be inferred that B is false. Thus the conclusion is $\neg A$.

Example III: if the set of axiom are

$A \vee B$

$\neg B \vee C$

It can be inferred that $A \vee C$ is true.

2.2.3.4 Forward and backward chaining

The way to find the resolution from knowledge base is forward chaining or backward chaining. The characteristic of rules in knowledge base indicates a type of chaining. Consider the next examples:

Example I Knowledge base:

$cow(X) \rightarrow mammal(X)$

$dog(X) \rightarrow mammal(X)$

$crow(X) \rightarrow bird(X)$

$bird(X) \rightarrow animal(X)$

$mammal(X) \rightarrow animal(X)$

dog(Dicky)

Query:

?dog(Dicky)

Prove :

animal(Dicky)

From rule base, if we know that Dicky is dog then we can prove Dicky is animal as well. In this case forward chaining is more efficient as shown in hierarchical tree in Figure 2.8.

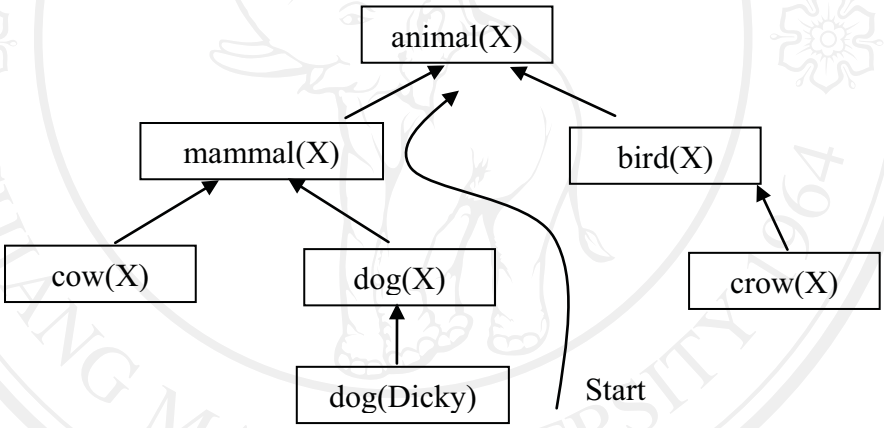

Figure 2.8 The example of forward chaining process

Example II Knowledge base:

dog(X) → domestic_animal(X)

domestic_animal(X) → frienly(X)

dog(X) → mammal(X)

mammal(X) → vertebrate_animals (X)

mammal(X) → animal(X)

dog(Dicky)

Query:

?dog(Dicky)

Prove :

animal(Dicky)

From rule base above proving Dicky is animal, backword chaining is more efficient as shown in hierarchical tree in Figure 2.9.

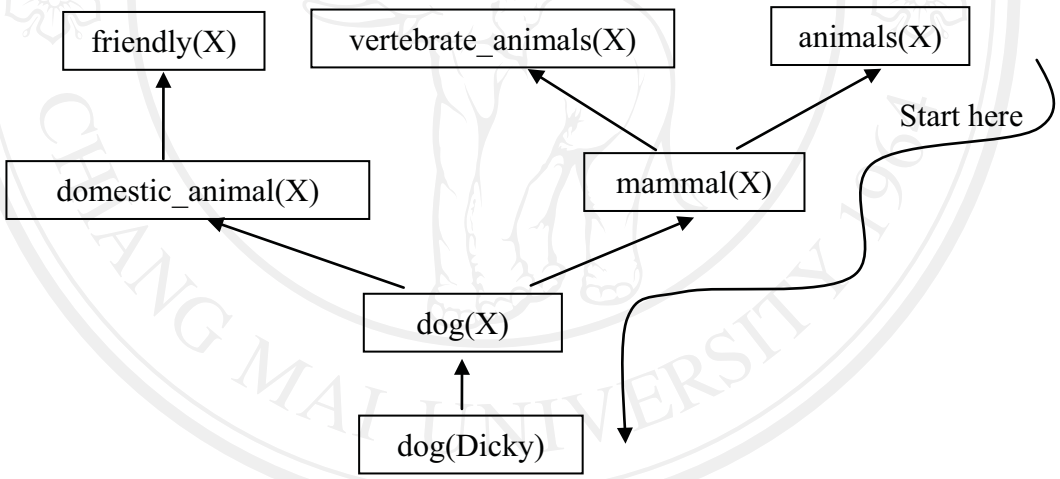

Figure 2.9 The example of backward chaining process

For creating new knowledge from knowledge base, the expert system represents the conditions in the real world in term of first order predicate logic. After that the system will process inference step. This technique will analyze the existing knowledge with those conditions. Finally users will get new valuable knowledge.

2.2.4 Knowledge Representation

Representation of knowledge is one important step of expert system. It claims that the more well organization of representation, the more efficiency of expert system. There are many ways to represent knowledge. In this study, I will give just four example methods as follow.

2.2.4.1 Rules Base

Rules base is based on production rules. This method uses Post machine theory developed by E. L. Post in 1943 (Woowong, 1992). This method use rules to represent the knowledge base in form of IF THEN format. If condition is complete, command behind *THEN* will be proceeded.

IF (condition) THEN (conclusion or action)

For example:

IF Dicky is dog THEN express Dicky is animal

2.2.4.2 Semantic Networks

A semantic network is a labeled directed graph representing concepts, objects, their properties, and their relationships. A semantic network can represent in graphical semantic network. It composes of nodes and arcs. The *nodes* are oval in shape and they represent concepts, objects, or situations. The *arc labels* describe the relationship between two nodes. An arc, point to a node, describes a property of that node Figure 2.10. In the other way, properties are square in rectangle in shape and arc can be used to connect an object to its property as shown in Figure 2.11.

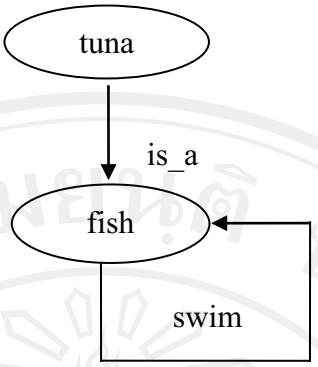

Figure 2.10 The example of semantic network using arc as property

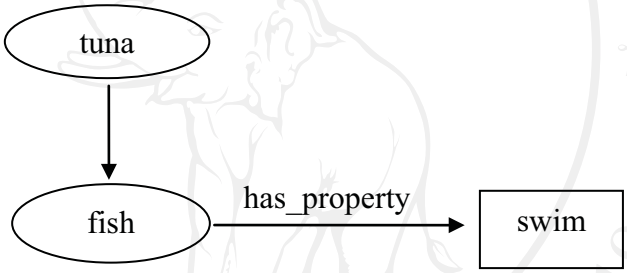

Figure 2.11 The example of semantic network using square as property

The inference algorithms with a semantic network are graph traversal algorithms. For example, statement that "Tuna is fishes and fishes can swim" is expressed as a semantic network in two equivalent ways in Figure 2.10 and 2.11

Semantic networks can use for representing relationship between concepts and especially hierarchies of vocabulary in concept. Therefore with using semantic networks, it is easy to explain fact in nature. In addition, the semantic network also permits *property inheritance*. For example, in Figure 2.10, "tuna is a fish and fishes can swim, so the conclusion is tuna can swim". Here tuna inherit property from fishes so tuna has the same property of fish, swim. As same as the nature, the property of all

things can be override occasionally. For example, in Figure 2.12, the color of tigers are yellow but white tiger, a kind of tiger, has black and ash grey in color. Though, white tiger inherits the color from tiger, but it will override this property with their own color.

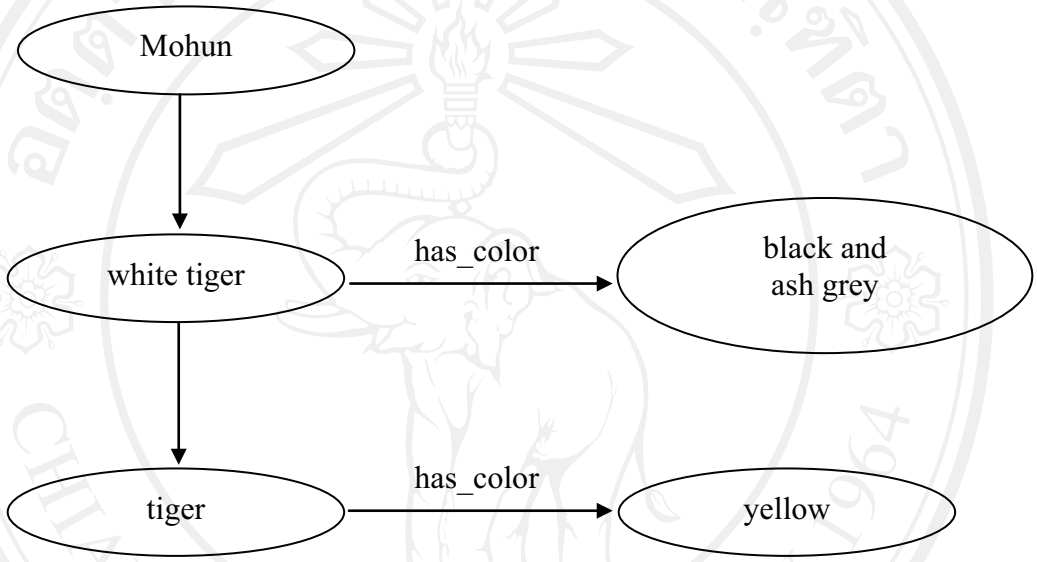

Figure 2.12 The example of inheritance and overriding of color property

2.2.4.3 Frames

A frame is as same as semantic network method, but it is added attributes for each node. These attributes will be called *slot*.

The slot composes of a set of attribute as same as *record* in relational database. However slot is different from record in database because attribute of slot can hold other objects such as a set of objects, algorithms for performing various tasks (called demons), scripts, pointers to other frames, as well as other complex object structures

generated by combining the above elements. The example of frame is shown in Figure 2.13.

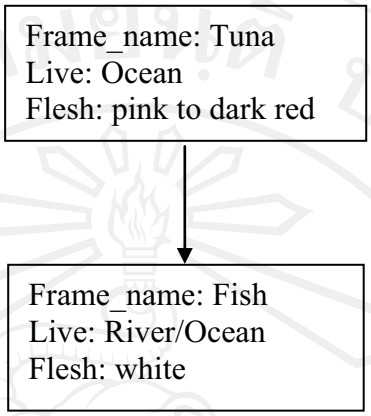

Figure 2.13 The example of frame

2.2.4.4 Ontology

An ontology is a kind of knowledge representation based on frame and predicate logic. Slot in frame is used for representing domain knowledge whereas predicate logic is used as a mathematic basic for reasoning rules in inference process.

An ontology uses for defining a common vocabulary or a shared understanding. The term *ontology* has been borrowed from Philosophy mean "a systematic account of Existence" (Guarino & Giaretta, 1995). It means a systematic account of existence. In expert system, ontology can explain in many ways. Gruber (1993) said ontology is an "explicit specification of a conceptualization". Borst (1997) states ontology is an "explicit formal specification of a shared conceptualization". Uschold and Gruninger (1996) defined ontology is a vocabulary of terms and some specification of their meaning. Ontology is also a representation for the meaning of concepts in a domain (Beck & Pinto, 2002).

Ontology was developed for solving the problem of recognizing the two pieces of data talking about the same thing in the different terminology (Castello & Jacobs, n.d.). Therefore an ontology is use to describe a particular domain. It consists of a set of terms and a set of constraints imposed on the way those terms can be combined (Beck and Pinto, 2002). Those terms are concepts of the domain or properties of the concept. Constraints are the value of the properties (Noy, 2000). Those terms can be called class. For example *Computer Skill*. Each class also has subclasses as well. For example subclasses of Skill are Web technology, Database, etc. The simple ontology composed of class and subclass can called *taxonomy* as shown in Figure 2.14. The complex ontology is taxonomy added relationship among classes as shown in Figure 2.15.

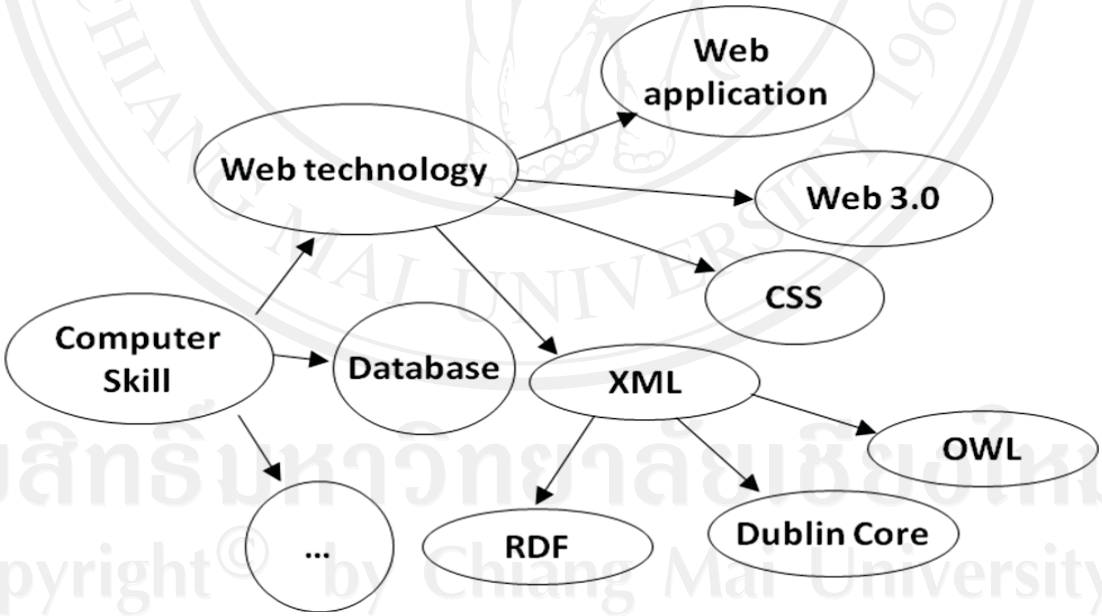

Figure 2.14 The example of taxonomy

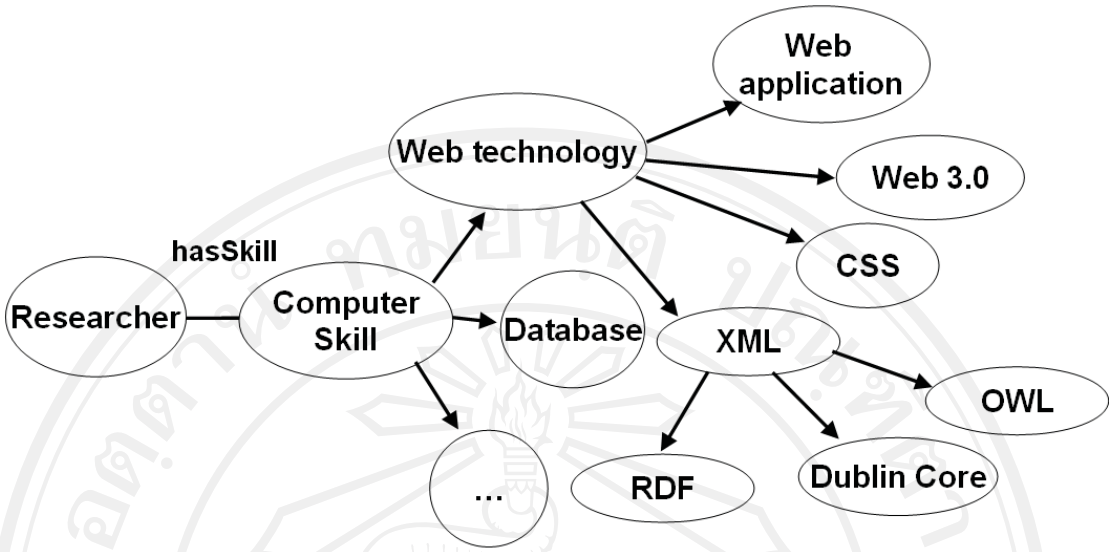

Figure 2.15 Example of ontology in researcher domain

Present, ontologies have been applied in many business and scientific communities as a way to share, reuse and process domain knowledge. Ontologies are now fundamental to many applications such as scientific knowledge portals, information management and integration systems, electronic commerce, and semantic web services (The National Library of Medicine, 2009).

a.) Notation

In low-level syntax of knowledge representation is a *notation*. The low-level syntax is the way to represent the knowledge physically in computer. The recent fashion of a notation is to use XML (eXtension Markup Language). The advantages of using XML for representing knowledge are as follow.

- The knowledge is easy for machines to parse.
- Human can read easily.
- Use less space more than other syntax.

There are many markup ontology languages to encode knowledge, most commonly based on XML. The examples are

- DAML+OIL
- Ontology Inference Layer (OIL)
- Web Ontology Language (OWL)
- Resource Description Framework (RDF)

b.) OWL (Web ontology language)

OWL is one of the popular notation or ontology language endorsed by the W3C (World Wide Web Consortium) in 2007. OWL is adapted from RDF (Resource Description Framework), which is based on XML format. So it is flexible to use OWL in any platform. Otherwise it is considered as one of the fundamental technologies supporting the Semantic Web ("Knowledge representation", 2009). OWL is an ontology language that is primarily designed to describe and define classes. (Drummond, Horridge, and Knublauch, 2005).

c.) Ontology building tools

There are many ontology building tools for developing owl as follow :

- Ontolingua Server, available at

  http://www.ksl-svc.stanford.edu:5915

- Ontosaurus, available at

  http://www.isi.edu/isd/ontosaurus.html

- WebOnto, available at

  http://webonto.open.ac.uk

- Protégé, available at

http://protege.stanford.edu

2.2.5 Ontology Building

Building ontology in this study, Protégé is used. It is a free, open-source platform, developed by Stanford University. Because Protégé gives a rich set of knowledge-modeling structures and actions that support the creation, visualization, and manipulation of ontologies in various representation formats. Protégé also can be customized to provide domain-friendly support for creating knowledge models and entering data as well. Further, Protégé can be extended by way of a plug-in architecture and a Java-based Application Programming Interface (API) for building knowledge-based tools and applications (The National Library of Medicine, 2009). One of a plug-in program is Racer pro using for reasoning process in inference engine ("RacerPro 2.0", 2009). The building ontology processes are as follow:

a.) Create classes and subclass for particular concept.

b.) Create properties both relationship and attribute for each class.

c.) Design appropriate restriction or constrain in particular domain.

d.) Run inference step or reasoning process for generate the new knowledge.

Detail of each step is as follow (Beck & Pinto, 2002; Horridge, Knublauch, Rector, Stevens, & Wroe, 2004).

2.2.5.1 Creating of Classes

Class uses for representing generic concept of particular domain in ontology. For example, domain is *pizza* then all classes in this domain are *Pizza, PizzaBase*, and *Pizza Topping* as shown in Figure 2.16.

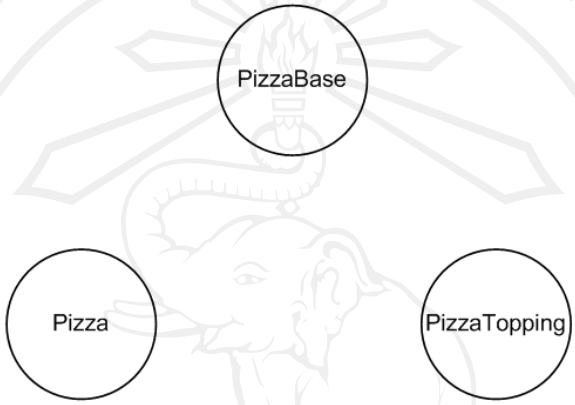

Figure 2.16 The example of classes in pizza domain

When implement those classes in protégé, the result presents in hierarchical tree as shown in Figure 2.17.

Figure 2.17 The hierarchical tree of pizza classes in Protégé

2.2.5.2 Creating of Subclasses

Subclass is a class that is more specific than a particular class. For example shown in Figure 2.18, subclass of *PizzaTopping* class in domain of *pizza* is *VegetableTopping*. *TomatoTopping* is also subclass of *VegetableTopping*. After implement in Protégé, the result has come as Figure 2.19.

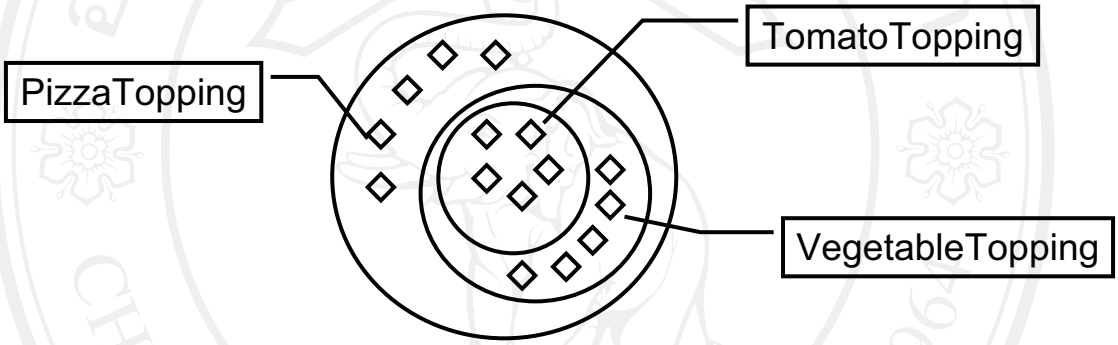

Figure 2.18 The example of subclass of PizzaTopping class (Horridge et al, 2004)

Figure 2.19 The hierarchical tree of subclass in protégé

2.2.5.3 Creating of Properties

Properties are used to describe more about class. There are two kinds of the properties

1. Object properties is a relationship between two classes.
2. Datatype properties is a relationship between class and data value.

According to the example classes in pizza domain, the object properties are relationship between Pizza class and PizzaBase class called hasBase and relationship between *Pizza* and *PizzaTopping* called *hasTopping*. Further, all relationships also have inverse (*isBaseOf* and *isT oppingOf*) respectively as shown in Figure 2.20. Note that all properties has small letter to be the first character of their name.

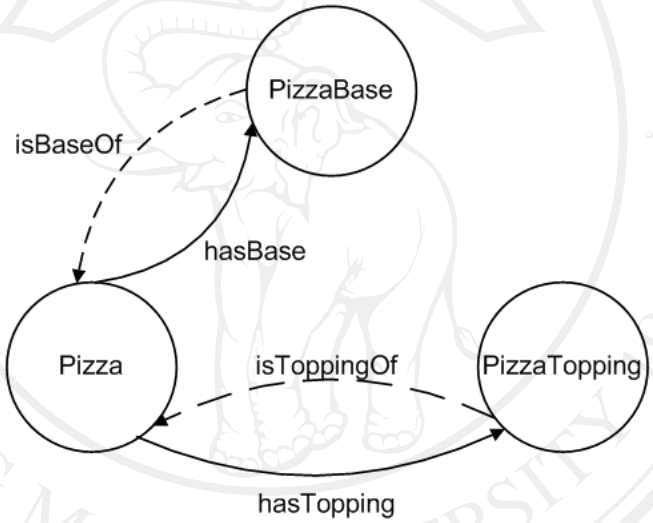

Figure 2.20 The object properties of pizza domain

When implementing properties in Protégé, the result has come as Figure 2.21. Notice that *hasIngredient* and *isIngredientOf* were used to be general properties of the others four properties for grouping in hierarchical form.

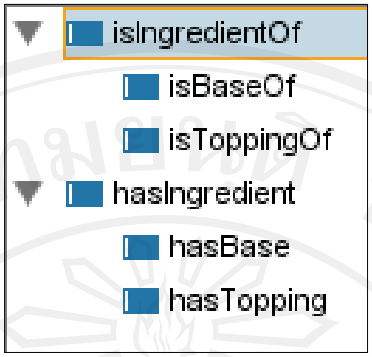

Figure 2.21 The object properties of pizza domain in Protégé

After create object properties, then domain and range of those properties are defined for create relationship between classes/subclasses. For example hasBase is defined domain as Pizza and range as PizzaBase as shown in Figure 2.22.

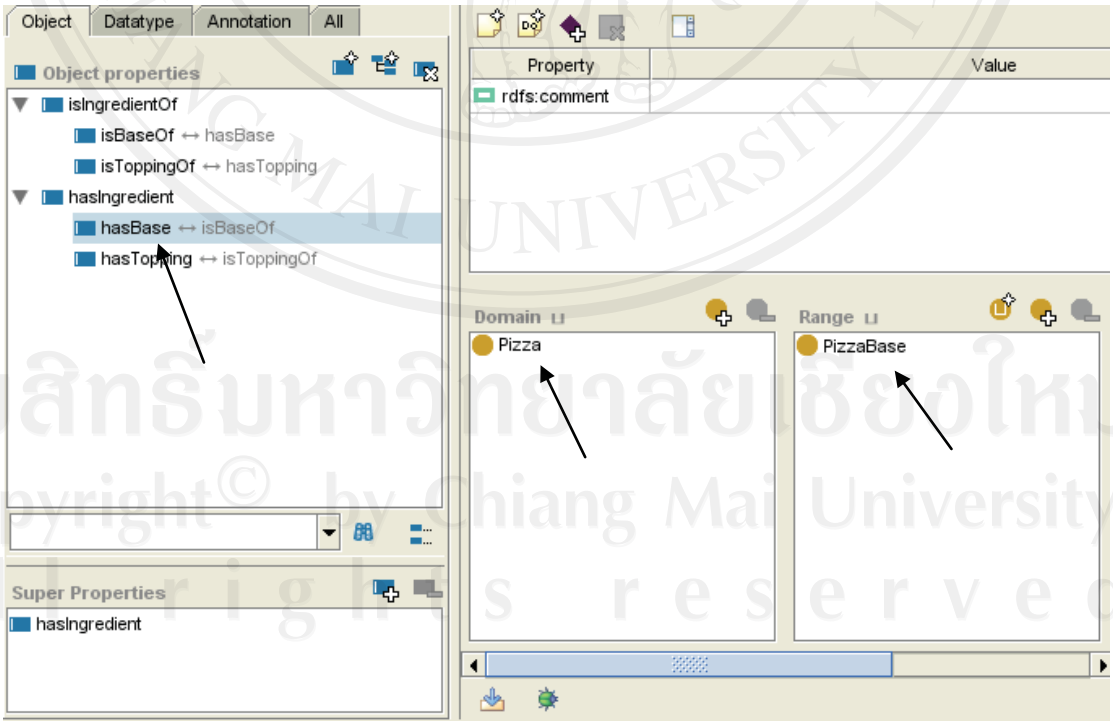

Figure 2.22 Example of domain and range of hasBase

There are 4 types of object properties as follow:

1. Functional : the relationship between two classes/subclass as one to one (1-1). It means if give one individual, there can be at most one individual that is related to the individual via the property.

For example:

Somchai *hasFather* Somboon.

In this example, Somchai has one object properties named *hasFather*, which is refer to his father. Somchai can have only one father. Thus *hasFather* is functional object property.

2. Inverse Functional: the relationship that inverse functional relationship.

For example inverse functional of above example:

Somboon *isFatherOf* SomChai

3. Symmetric: the relationship that can read forward or backward for the same meaning.

For example

Somchai *beFriendOf* Somyeing

This example can read backward for same meaning like this

Somyeing *beFriendOf* Somchai

4. Transitive: the relationship that gives new meaning via transitive.

For example:

Somchat isBrotherOf Somchai

and we already know that

Somchai hasFather Somboon

so we can refer that

Somchat hasFather Somboon

From pizza domain, hasBase can define property type as functional and isBaseOf is inverse functional of hasBase. Then property type of isBaseOf is inverse functional and its inverse property is hasBase as shown in Figure 2.23 and 2.24.

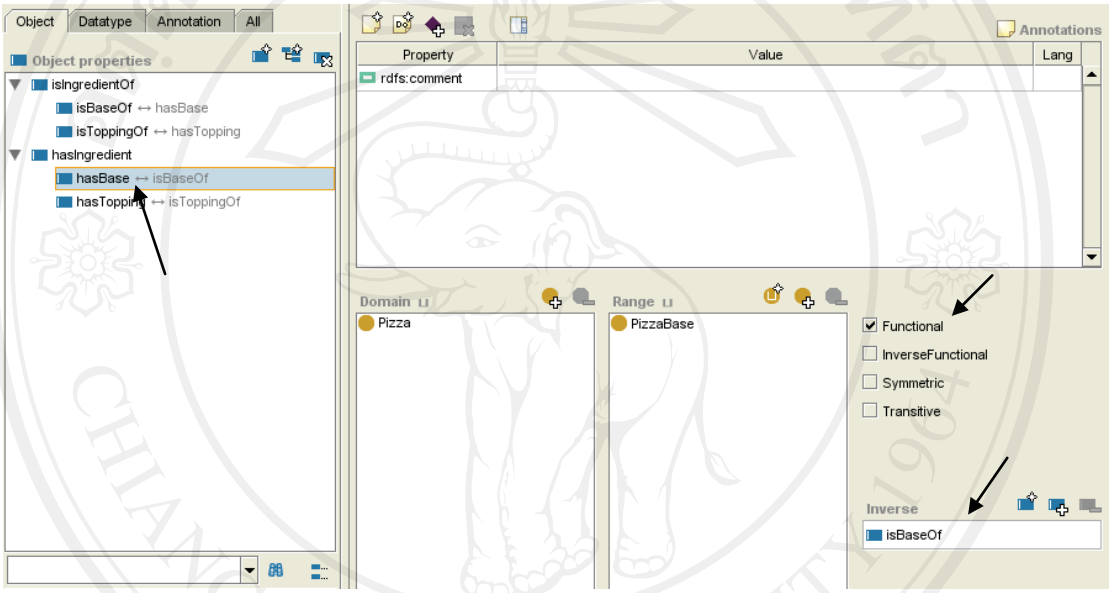

Figure 2.23 Property type of hasBase

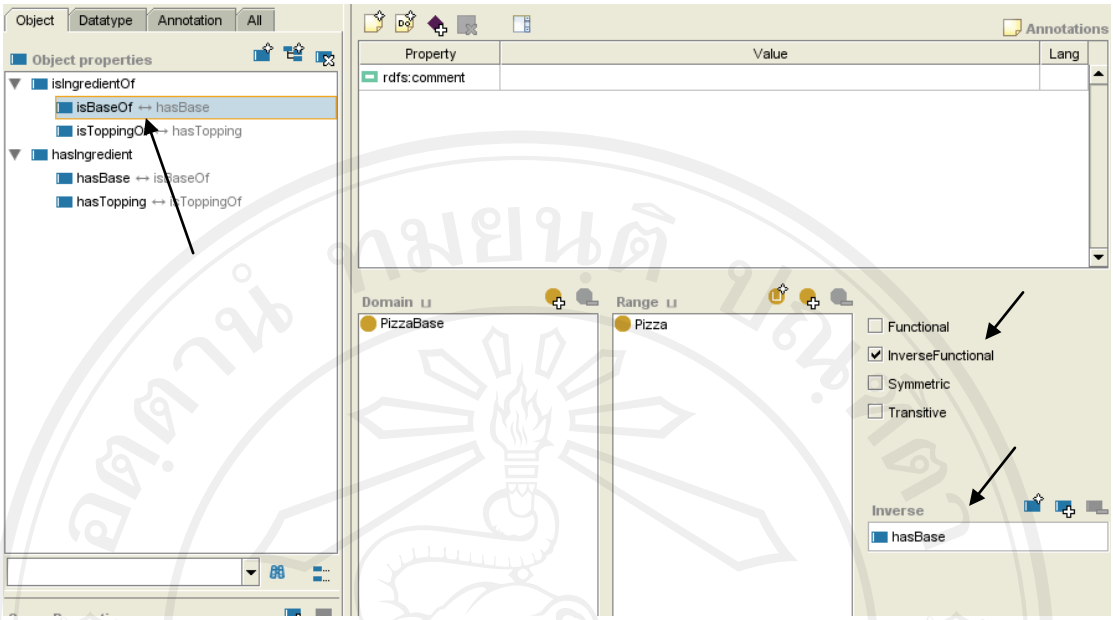

Figure 2.24 Property type of isBaseOf

The other property is *datatype properties*. This kind of properties is attributes of the class or subclass. The example of attributes of *pizza* is *PizzaId*, *Weight*, and *Size* shown in Figure 2.25. The result of implementing in Protégé is shown in Figure 2.26.

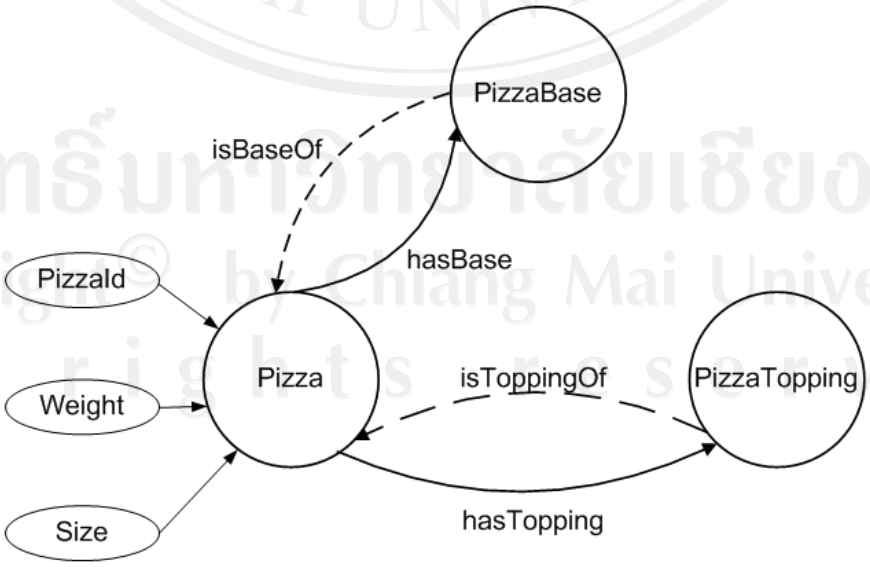

Figure 2.25 The example of datatype properties in pizza domain

After implement datatype properties in Protégé the output becomes as shown in Figure 26

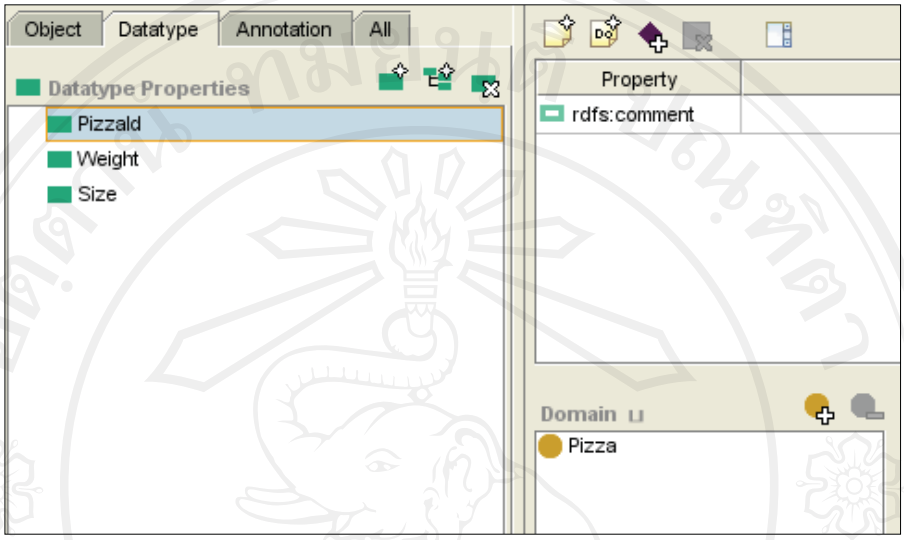

Figure 2.26 The datatype properties of pizza domain in Protégé

2.2.5.4 Designing of Restriction or Constrain

Restriction or *constrain* is a condition or filter using in inference process. Formula of restriction is represented in a *Predicate logic*. Restrictions can be grouped into three main categories:

1. Quantifier restrictions

1.1 Existential restrictions (denote by the symbol ∃, called Existential some) uses for describing the class that have at least one relationship to the others class.

Example: the restriction of Pizza

hasTopping some PizzaTopping

hasTopping is represent a property of pizza and *some PizzaTopping* is a filter.

This formula means every pizza must have at least one topping.

When implement existential restrictions in protégé, the result is become as shown in Figure 2.27.

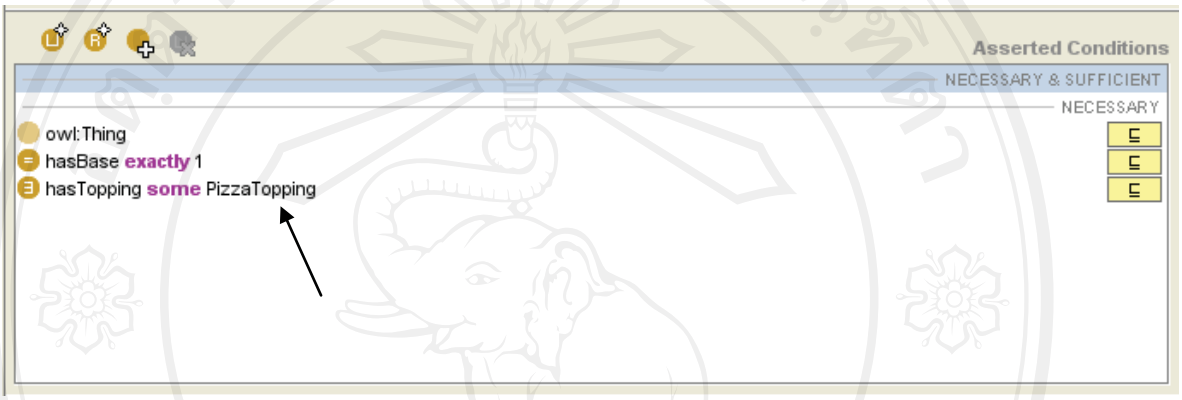

Figure 2.27 The example of quantifier restriction

1.2 Universal restriction (denote by the symbol ∀, called Universal only) uses for describing the class that has at most one relationship to the others class.

Ex. restriction of Americana Pizza: *hasTopping only (MozzarellaTopping or PeperoniTopping or TomatoTopping)*

hasTopping is represent a property of *pizza* and *only (MozzarellaTopping or PeperoniTopping or TomatoTopping)* is a filter.

This formula means every Americana pizza must have only MozzarellaTopping or PeperoniTopping or TomatoTopping.

When implement universal in protégé, the result is become as shown in Figure 2.28.

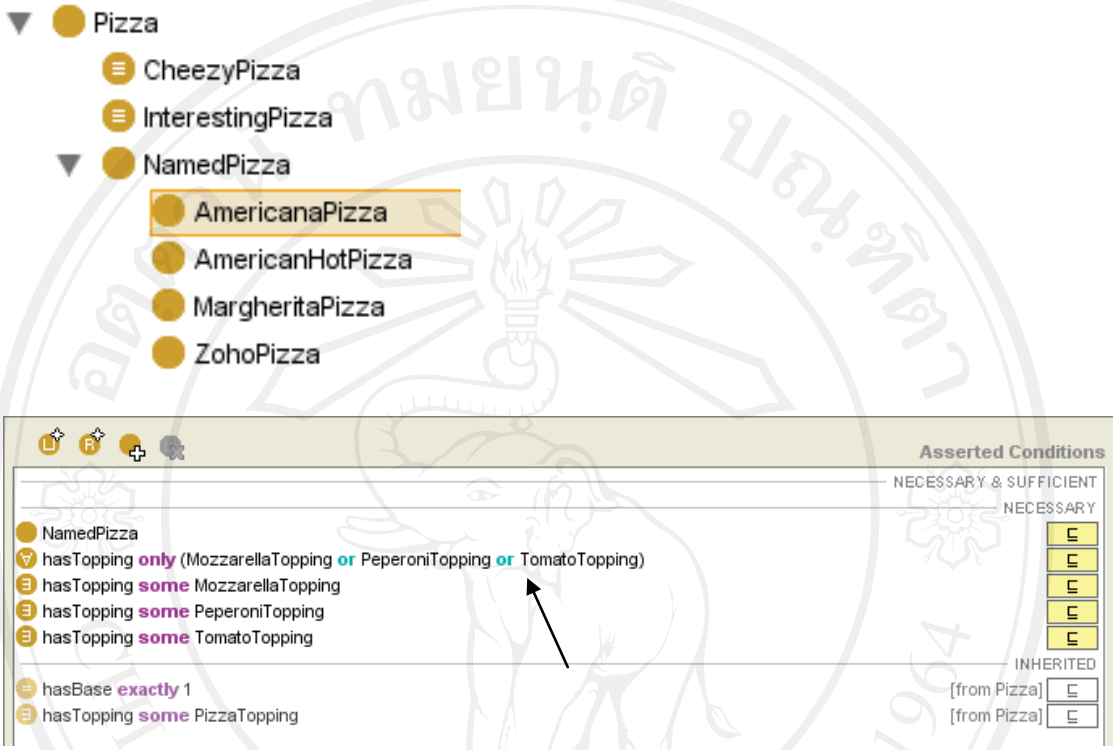

Figure 2.28 The example of universal restriction

2. Cardinality restrictions (>= min, = exactly, <= max)

Cardinality restriction uses for describing the class that has exactly number of relationship to the others class. For defining number of topping equal three, the formula should be *hasTopping exaxtly 3* as shown in Figure 2.29.

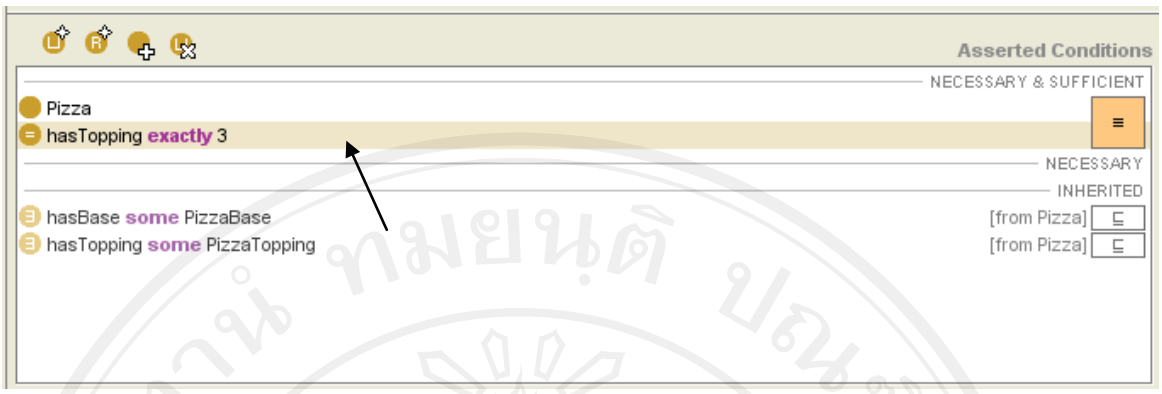

Figure 2.29 The example of cardinality restriction

3. hasValue restrictions (denote by the symbol ∋, called *has*)

hasValue restriction uses for describing the set of individual that have at least one relationship along a specified property to a specified individual. After implementing in Protégé, the result has come like Figure 2.30. For example, if we know that normally CountryPizza makes by some ItalianPerson, the formula should be *hasPizzaMaker some ItalianPerson.* If specific the person, one new condition is added. Suppose that only ItalianPizza is made by chef named Yok. The new restriction became *hasPizzaMaker has Yok* as shown in Figure 2.30.

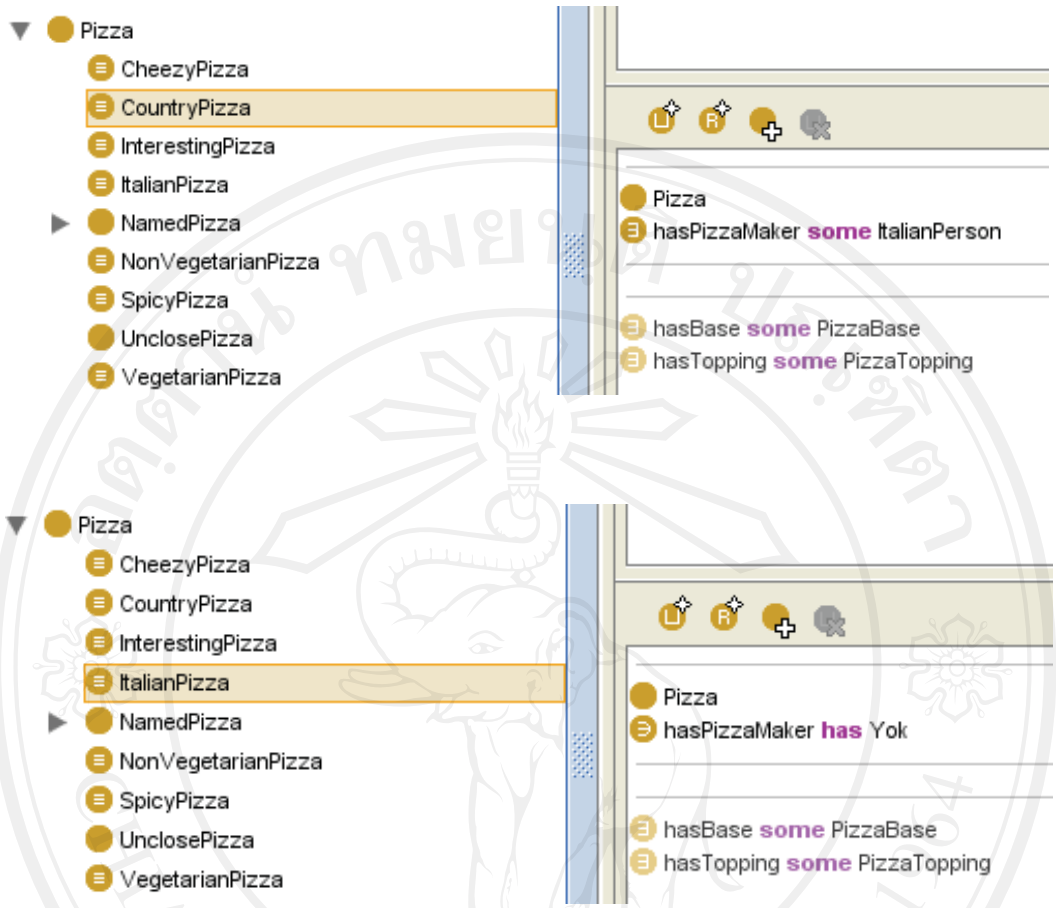

Figure 2.30 The example of hasValue restriction

2.2.5.5 Complex classes

Complex classes can be built up using simpler classes that are combined together with logical operators. In particular:

1. AND (∩) , or called intersection used for combine two classes by choosing all elements of the first class that also belong to the second class.
2. OR (∪) , or called union used for combine two classes by choosing all distinct elements in those two classes.

2.2.5.6 Running of Inference Engine

Inference engine of Protégé has the process to reasoning the particular concepts (classes/subclasses) with restriction or constrain. At the end, the new knowledge is generated. Running inference step in Protégé has to install plug-in, named *RacerPro*. This software has been produced by Racer Systems GmbH & Co. KG, the commercial company. RacerPro software is the semantic middleware for industrial projects based on the W3C standards RDF/OWL ("RacerPro", 2009).

In this study ontology and inference engine can apply for identifying key researchers of CMU. This starts from storing all information of candidate key researchers' performance in term of ontology and representing key researcher criteria in term of predicate logic. After processing inference step, the system will give the list of key researchers in CMU.

2.2.6 Searching Problem

Besides ontology is used in expert system, it is applied in search problem as well. Because of the enormous amount and inaccuracy of information on the web, it is necessary to increase the effectiveness of searching on the Internet. It is obvious that the Internet is an excellent source for searching information. However, many times users found that the result of a search, especially from search engines like Google, resulted in an enormous number of hits. Not only an unreadable number of search results but also many of the search results do not accurately relate. Even though the user takes only a few seconds to find information, they have to spend a lot of time to screen through huge amounts of inaccurate information to find what he/she needs.

The traditional information retrieval using in search engines such as Google and Yahoo is fulltext search. The statistical techniques are used based on the frequency of keywords appearing in a document. All keywords found in web documents are index, which are put in search engine database for using in searching process (Beck & Pinto, 2002). For example, the keyword "wiki" would appear in this database, and would point to every document containing the word "wiki". A user searching for "wiki" is shown a list of all these documents, along with a ranking based on the number of times the word "wiki" appeared in each document. It most likely the more times the word appears in a document, the more relevant that document is.

However there are so many documents on the web. Thus the overwhelming number of these documents is not relevant to the user's interest. It is possible that the important documents to the user do exist because different words were used in the document that did not match the search terms directly. These two types of errors are formally measured as precision and recall (Beck & Pinto, 2002).

$$\text{Precision} = \frac{\text{Number of relevant documents identified}}{\text{Total number of documents identified}}$$

$$\text{Recall} = \frac{\text{Number of relevant documents identified}}{\text{Number of relevant document in the collection}}$$

A search engine with perfect precision and recall would find all and only the documents relevant to the user's interest. Precision is a measure of how many of the documents identified as a result of a search are relevant to the user requirement.

Recall is a measure of how well the search engine did in positioning relevant documents (did it find all the relevant documents in the collection). Normally, with fulltext search in most search engine is not able to do the good number of both precision number and recall number.

Furthermore, fulltext search is not capable of processing a question statement, such as “list of top five highest mountains in Thailand”, since it can simply generate a list of documents containing these terms. For processing this question, query language would be required to database.

Fulltext search have no technique to understand the meaning of terms being used. The meaning or words are very complicate for interpreting in many reasons as follow.

a.) Words can have many senses (wiki=quickly versus wiki=open source software).

b.) Synonymous terms are used in different situations (mouse versus rat).

c.) Words have a wide variety of different associations and interrelationships (cat is a kind of animal, tail is a part of cat).

d.) Terms appear in different languages (cat in English, แมว meaw in Thai).

Ontologies have been developed to correct these kinds of problems. Thus it can be used to better organize information resources and assist users in retrieving related information (Beck & Pinto, 2002).

2.2.6.1 Thesaurus

The thesaurus can be considered as a simple kind of ontology. The thesaurus has been made by librarians. It is used for building subject classifications and cataloging documents within subject headings. The thesaurus contains lists of words grouped together according to similarity of meaning (containing synonyms and sometimes antonyms) ("thesaurus", n.d.).

A thesaurus attempts to categorize words using a three of simple abstractions (Beck & Pinto, 2002).

a.) Broader Term (BT) - A particular term is more general than another term ("cat" is broader than "Siamese Cat").

b.) Narrower Term (NT) – A particular term is more specific than another ("cat" is narrower than "mammal").

c.) Related Term (RT) – Two terms are associated each other("cat" is related to "flea").

d.) Use For (UF) – A particular term is the preferred term among a set of synonymous terms (use "feline" for "cat")

Thus, a thesaurus entry for the term "cat" might look like:

Cat

BT: mammal

NT: siamese cat

RT: flea (หมัด)

UF: feline

The advantages gained by this approach are that user can search either search terms or relevant topics. Thus thesaurus-based search can have a high rate of precision and recall.

The more complication of thesaurus is *taxonomies*, which demonstrate the BT/NT relationships in a hierarchical tree. Taxonomies with added object properties or datatype properties are ontologies (Heflin, 2004).

2.2.6.2 Advantage Search

With use ontology, a representation containing interconnections between related terms, the search engine has information about the meaning of term. All information directly related to this concept is directly accessible via the relevant terms. Further, it can use these relations to examine similar or related concepts. In this way, the user gets access to all and only the information available on a particular concept (Beck & Pinto, 2002).

2.2.6.3 Semantic Web

The web using ontology to be the solution for searching is called a Semantic Web. Berners-Lee (2001) cited in W3C as "This kind of web is an extension of the current web in which information is given a well-defined meaning, better enabling computers and people to work in cooperation". When searching for information, search system will look not only information itself but also the meaning of that information. This technique enables the outcome of a search to be more precise to the direct meaning of the information. For example a term *warehouse* is a place for storing old or unused stuff, whereas in the computer science it is a place for keeping

old, routine data. This example shows that one term possibly has more than one meaning. Therefore searching such this kind of information in general Search Engine, the output becomes enormous number. Because the output comes up with warehouse of both meaning even users need only *warehouse* in meaning of computer science.

Whereas in Semantic Web, data mining, which is "the practice of automatically searching large stores of data for patterns", is applied to the searching system ("Data mining", 2006). Data mining enables the searching system to learn the user's searching pattern. After that, system can know what kinds of outputs those are hit users' requirement. Therefore the output of the second time in searching, system will select only output which match those kinds. For example, University of Miami web site uses this technique to improve their search engine which searches information in University of Miami web. Users found that the more frequently they search information, the more accurate output they receive it (Finin, Ding, & Joshi, 2005). Data mining will keep and learn the user searching pattern and give search results which are more number of hits. Another example is a digital library, which uses semantic web to help user searching their information. After using semantic web, user can access the right information in the short time (Warren, 2005).

The other example of semantic web is *Bing*, which is a new search engine launched in 2009 by Microsoft. Formal name of Bing are Live Search, Windows Live Search, or MSN Search. It was advertised as a *decision engine* because it uses semantic model for storing all data. Thus Bing has gained market share in its first few months of release. Bing does a better job countering "search overload," offering more-relevant, higher-quality results ("Bing", 2009).

Research management of universities also faces of the same problem. Administrators want to find out expertise and research direction of their researchers. But sometime they find that search output is not hit to their requirement. It causes of many vocabularies refer to the same research so the search output do not cover all relevant information. For example if administrators want to find research about *Nanotechnology* but keywords or vocabularies using in these researches may not use term *Nanotechnology*. Thus administrators cannot find these researches in the research management system due to there are many words refer to *Nanotechnology* as well such as *Nanomaterial*, *Nanocomposites*, or *Ceramic*. The ontology, one tool of AI, should be used for representing this kind of information, which is captured by knowledge engineering technique first. This AI technique makes search results are more precise and accurate. Addition each university has its own search space related to expertise and research direction of their researchers so search space of each university is different. The ontology developed specific for supporting search space of each university can be called *ontology commitment*. For this study, *CMU ontology commitment* is developed. With search this ontology by semantic web technology, research management system enable to synergize between administrators and researchers including synergize between new researchers and high potential researchers as well.

2.3 Knowledge Engineering Techniques

2.3.1 Knowledge Engineering

The term *Knowledge engineering* is used to describe the expert system development process. A *knowledge engineer* is the individual engaged in knowledge engineering (Nikolopoulos, 1997).

The responsibility of knowledge engineer is about

1. Development knowledge acquisition.
2. Development user interface design.
3. Selection of software, hardware, and implementation.
4. Verification and validation of part of the expert system

2.3.2 Ontology in KM

Ontologies serve to provide a convenient representation for the semantics of some particular domain (Finin, Ding, & Joshi, 2005). There are many attempts try to apply ontologies in knowledge management such as employing an ontology development method to describe the relations between technology, KM and strategy (Saito, Umemoto, & Ikeda, 2007).

If apply ontologies in semantic web, it will improve web searching in enormous information (Berners-Lee, Hendler, & Lassila, 2001). One research tries to explore the role of ontologies in the construction project delivery process, particularly with respect to information and KM. It found that there is considerable merit in ontology-based approaches to information and KM, but that significant technical challenges remain. Middleware applications, such as semantic web-based information

management system, are contributing in this regard but more needs to be done particularly on integrating or merging ontologies (Anumba, Issa, Pan, & Mutis, 2008). Sometimes ontologies were used to share a common understanding. This research seeks to develop an ontological approach, in order to make it possible to share a common understanding of accounting theory. In this case, accounting theory is the specific structure of the profit and loss account among people or software agents. On implementing, database management system was used as tool (Chou, Vassar, & Lin, 2008).

Additional, ontology was used to improve the searching in knowledge management system (KMS) as well. This will result in increasing knowledge sharing among researchers (Wang, Yang, Kong, & Gay, 2003). With using ontology users can search not only for words but also semantic of the words.

One research combines P2P technology with explicit ontologies and uses it in inter-organizational knowledge management. The knowledge management can benefit from the combination of two technologies, semantic web technologies and p2p information sharing. This using was supported creating virtual organizations (global access and local control). The semantic web technologies provide rich knowledge structures that can be used to efficiently search and integrate information across system boundaries. In the other hand, P2P networks provide an adequate technical infrastructure for knowledge management applications that guarantee the independence and autonomy of individual information sources. When setting up a p2p network to support virtual organizations, there are a number of design decisions that

have to be made. This research design decisions in terms of the different aspects of a semantic p2p information sharing solution (Stuckenschmidt, Siberski, & Nejdl, 2005).

Further for research management, there is a research used semantic web to apply for research information in an office environment to improve search results. The research ontology was developed for serving office job for storing stake-holders and activities which participate in research e.g. the researchers, their publications, research interests, conferences and journals, and subjects. The research ontology was developed for describing the area that plan to conduct research, the problems that want to address, and the methods, systems and approaches which have been described in the literature (Carr, Miles-Board, Wills, Wouken, & Hall, 2004).

In Universiti Putra Malaysia (UPM), there is project building knowledge management portal providing lecturers' profile and knowledge assets. The lecturer profile contains: lecturer teaching load, research, publication, etc. The structure of lecturer profile is based on taxonomy and some algorithm was written for analysis taxonomy for measure lecturer workload (teaching, research, and services). This structure called the algorithmic taxonomy. It found that this method contributes better evaluation of teaching workload for lecturers of UPM (Hashim, Hamid, Selamat, Ibrahim, Abdullah, & Mohayidin, 2006).

In this study ontology was built for the sake of research management. All candidate key researchers' information and their research results, count as intellectual capital, was populated in CMU research ontology for the sake of identifying key researchers, forming research clusters, setting research clusters' priority, and formulating research strategies and policies.

2.3.3 Social Network

Social network theory is a field in social science, which is used for a wide range of human organizations, from small groups of people to entire nations (Ethier, n.d.). Social networks are explicit representations of the relationships between individuals and groups in a community (Finin, Ding, & Joshi, 2005). It can represent a social structure ("Social Network", 2008) by nodes and links among nodes. Node is the actor within the network and refers to people or groups of people. The arc is the relationship between the actors and refers to relationship between people or groups. For example, a network might consist of researchers and a mapping from those researchers to each of their colleagues and relatives (Ethier, n.d.; "Social Network", 2008).

With use social networks, sociologists can identify primary groups and can be especially construct a person's reputation. Present there are many research interested the power within social networks and stability of networks. They found that a balance of the power within a social network is necessary for the stability of the network. Since people are more likely to stay in a group wherever he/she can share equal power with his/her peers. Social network can use in many research such as used to better understand how and why researchers interact with each other and show how technology can change this interaction (Ethier, n.d.).

To visual represent and analyze social network, network analytic tools are used. They are computer software providing mathematical functions. These tools makes user understand network data easily. Further they are used to change the layout, colors, size, and advanced properties of the network representation ("Social Network", 2008).

One of network analytic tools is *Cytoscape*. It is an open source software platform for visualizing network in form of nodes and arcs. Further it can integrate each node with annotation. Cytoscape always uses in bioinformatics for visualizing molecular interaction networks and biological pathways (Cytoscape Consortium, 2008). In this study we used Cytoscape version 2.2 as a tool for exploring social network of CMU researchers, which embedded cooperate research among CMU researchers and others researcher. We expected that the more relations researcher has in the social network, the more power, knowledge, and influence the researcher have and provide for university.

Cytoscape was used in this study for generating social networks of candidate key researchers via analysis list of authors and co-authors in papers published in academic journal. These networks can be helped CMU research administrators to predict, which are new research clusters or COE suitable to be founded in the future.

2.3.4 Card Sorting

2.3.4.1 Definition of Card Sorting

Card sorting is a simple technique or process involving grouping and sorting a series of cards. On process, each card is labeled with a piece of content, which is used to group and sort making sense to users or participants. The results of card sorting can represent mental model of these users (Spencer & Warfel, 2004, “Card Sorting”, 2008). So it can be claimed that card sorting are a knowledge elicitation technique (Userfocus, 2008).

2.3.4.2 Objective of Card Sorting

Card sorting is used in many purposes such as organizing content, gathering user-centred labeling, creating top level of concept, and thinking validation (Boulton, 2007). With implementing card sorting technique to computer system, the ability of the system to find information will be enhanced. Then the system is easier to use (Spencer & Warfel, 2004). Card sorting can lead to build category tree, folksonomy, or taxonomy (a simple ontology). It is a useful approach for designing workflows, menu structure, or web site navigation paths ("Card Sorting", 2008).

2.3.4.3 Type of Card Sorting

There are two types of card sorting: closed card sorting and open card sorting. Closed card sorting is the way to pre-define category labels before sorting. This type helps to identify where new things belong in a structure. This type is useful for validating the new cards.

Open card sorting type is the way to build structure from individual items of content upwards. This type allows user can generate new labels on running process. This type is useful for the beginning of creation (Boulton, 2007; Spencer & Warfel, 2004).

2.3.4.4 Selecting Participants

Card sorting can performed either individually or in groups. The most important thing of selecting participants is they are come from and representative of target group. A benefit of group sorts is that they typically provide richer data than individual sorts. The number of groups needed may depend upon the size and complexity of the site or product.

2.3.4.5 Process of Card Sorting

Before sorting card, some steps have to prepare (Information & Design, 2006).

1. Ensure that each term is as clear and unambiguous as possible.
2. Included all the need items to categorize.
3. Shuffle or randomize cards prior to each participant session.
4. Script a set of instructions so that all participants have the same understanding of the process.
5. Leave participants alone while they are sorting the cards to avoid placing them under unnecessary time pressure.
6. Provide additional blank cards for people to write group names

After prepare all thing, card sorting will be conducted. It can be conducted in a variety of circumstances using various means: one-on-one, during workshops, by mail, or electronically. The following is the basic process.

1. Names of items to be categorized are printed on individual cards.
2. Cards should be large enough to contain the names in a font that participants can read easily when spread out on a desk or table-at least 14 point.
3. Participants group items in a way that makes sense to them.
4. If use closed card sorting, participants have to name the resulting groups.
5. The results of card sorting have to be analyzed with some technique.

However on proceeding, users should consider on some aspects (Boulton, 2007; Spencer & Warfel, 2004).

1. Number of card should between 30 and 100 cards for working well.
2. The labels need to be as clear as possible.
3. Get a digital camera to record each sort.
4. Get observation. At least one person to take notes on the session.

2.3.4.6 The Advantage of Card Sorting

There are many advantages of card sorting as follow (Boulton, 2007; Spencer & Warfel, 2004).

1. Simple: Card sorts are used easily.
2. Cheap: All expenditure of card sorting depends on only cost of building card, printing labels, a pen, and time.
3. Quick to execute: Several sorts can be performed in a short period of time but it will give significant amount of data.
4. Established: Card sorting technique has been used for over 10 years, by many designers.
5. Involves users: Card sorting is user-centered method so it is based on real user input, not designer opinions.
6. Involves multiple users: Card sorting can identify patterns of thinking across multiple users. Thus it is efficiency tool for communicate in working group.
7. Provides a good foundation: It provides a good foundation for the structure of any concept.

2.3.4.7 The Disadvantage of Card Sorting

Although card sorting has many advantage but it still has some disadvantage that users have to consider with (Boulton, 2007; Spencer & Warfel, 2004).

1. Does not consider users' tasks: Card sorting may lead to an information structure that is not usable when users are attempting real tasks. Therefore the results of card sorting needs analysis to ensure that the content being sorted meets user needs and the result leads users to achieve tasks.
2. Results may vary: The card sort may provide vary of output depend on users' idea.
3. Analysis can be time consuming: The sorting is quick, but the analysis of the card sorting output is difficult and time consuming especially if there is little consistency between users.
4. May capture "surface" contain only – Sometime users do not understand the content of terminology labeled on cards in-depth. They may just sort it by the surface of content.

In this study card sorting was used for sorting and grouping vocabularies/Keywords, which refer to expertise and research direction of candidate key researchers of CMU. The output of this process is CMU research clusters, which can be used to plan for establishing laboratory or managing research fund in the future.