TACIT KNOWLEDGE EXTRACTION FOR SOFTWARE REQUIREMENT SPECIFICATION (SRS): A PROPOSAL OF RESEARCH METHODOLOGY DESIGN AND EXECUTION FOR KNOWLEDGE VISUALIZATION

Tunç Durmuş Medeni
TURKSAT A.Ş., METU, ANKARA/TURKEY
tuncmedeni@gmail.com
Serbülent Ünsal
Akgun Software A.Ş., ANKARA/TURKEY
serbulentu@gmail.com
Meryem Ayas
Havelsan A.Ş. ANKARA/TURKEY
meryemayas@hotmail.com
İhsan Tolga Medeni
Çankaya University, ANKARA/TURKEY
tolgamedeni@gmail.com

ABSTRACT

Knowledge extraction and visualization is becoming an important research area for the organizations in order to get and share the knowledge. Most important and useful part of the knowledge extraction and visualization is dedicated to tacit knowledge.

There are already known methods to acquire the tacit knowledge. Yet, these methods are mostly general approaches applicable to all the areas in need of tacit knowledge extraction and become too abstract when applied to a specific domain. One such specific domain is the requirement specification process for the software project development. Our own experiences in the area as well as the scientific researches have shown that Software Requirement Specification (SRS) process has field-specific problems that need to be eliminated by using the suitable tacit knowledge extraction techniques. For example, the experts and/or users may not have a clear idea of their requirements. They may also be technically unsophisticated or have different vocabularies than the software developers. Benefiting from the existing body of academic literature in the related fields, as well as co-authors' experience from their domains of practice, this paper aims to find the concrete methods for extracting the tacit knowledge in the area of software project development with specific implications for these academic fields and practice domains, as well as more general suggestions for all related or concerned.

To provide a base for future work, the paper also presents a proposal that aims to develop a tacit knowledge visualization framework to support know-where requirements of the organizational knowledge. With the implementation of our framework in a software application, it is aimed to create a virtual environment, where subject-based knowledge requirements will be answered by the visualized tacit knowledge of individuals and possibly the relations among individual members of the organization.

Tacit Knowledge Extraction For Software Requirement Specification (SRS)

Keywords: Tacit Knowledge extraction, Expert Knowledge, Software Requirement Specification (SRS), Knowledge Engineer (KE), Knowledge visualization, tacit knowledge, research methodology design

INTRODUCTION

Tacit Knowledge is "a non-linguistic, non-numerical form of knowledge that is highly personal and context specific and deeply rooted in individual experiences, ideas, values and emotions" (Gourlay, 2002:p.2). It is most commonly referred to be -what people called- "know-how" or "experience" which makes it hard to verbalize, elicit, write down or extract. On the other hand, tacit knowledge is one of the top demands of the organizations sought for when they hire people or maintain the current projects. Therefore, the extraction and visualization of tacit knowledge is an important field not only for the academic purposes but also for the organizational needs. Accordingly in general terms, Knowledge Visualization can be defined as the visualization of cognitive aspects of a subject for the individual perspective. Then Tacit Knowledge can also be understood as the individual cognitive knowledge, which is hard to explain, and Explicit Knowledge can be understood as codified, documented knowledge, which is easily accessible in organizational base.

In the history, there were different visualization techniques, for example the technique that was given in the study of Dalbello and Spoerri (2006). In nineteenth century, the books provided the basic tools for visualized information and knowledge. The relations and visualized information were done through shapes. It can be admitted that in present there are similar tools, with different perspective of visualization, like given in Boyak et. al. (2002). The visualization was done like a topographic map. More recently, different technological applications to support information and knowledge visualizations are becoming more available. However, in the literature, given knowledge visualization applications are mostly generic and the given systems stay on a level of visualization which tries to implement every kind of relations within a general structure, forcing the researcher away from the real requirements of what is really needed to visualize. Besides, the given systems for visualization are generally based on mostly data mining tools. However, having acknowledged that, these systems still stay for the static datasets. Static datasets can be seen as a good base for explicit knowledge gathering, but they are far away to be a good answer for tacit knowledge utilization.

Accordingly, how the tacit knowledge structure of the organization could be visualized with the existing explicit knowledge warehouse of the organizations is the major research question we have. Whether the visualized tacit knowledge structure could be implacable to a different organization with a different knowledge structure is another research question that complements this major one. Furthermore, the extraction of tacit knowledge especially becomes more challenging, when the software requirement specification (SRS) is in question. Co-authors' different experiences in various domains of the Turkish ICT sector confirms the challenging nature of expert knowledge extraction for SRS. Accordingly, with this work we also aim to shed light on extraction of experts' tacit knowledge for software requirement specification.

With respect to these definitions, explanations, understandings, questions and aims a main purpose of this study is to propose an intelligent framework dynamic and specifiable for visualizing the tacit knowledge structure of the organizations from the perspectives of individual members and their relationship between each other, and explicit subject bases. So first of all, it is needed an integrated concept for tacit and explicit knowledge visualization. Another point that deserves attention is the requirement of displaying a more dynamic structure. Because our concern is related with the cognitive content, a dynamic human related approach, and a system that could understand and guess the tacit knowledge of the organization members is needed.

After this introduction, respectfully, the content of the paper is organized as follows: Firstly, we will present a brief overview of the theory and related studies of the knowledge visualization. The next part we make a brief introduction to the general tacit knowledge extraction methods/approaches that are applicable to the software requirement specification area. Shortly after, we identify the extraction problems specific to the software requirement specification by dividing the problems into two categories, namely User/Expert Based Problems and Engineer/Developer Based Problems. Then we provide a deeper analysis on each problem and suggest new approaches as solutions while finding also the general methods that are applicable to solve each problem. Finally, the information on research design and methodology, as well as the proposed framework design as the result of the possible execution of the methodology will be provided.

BRIEF OVERVIEW OF STUDIES ON KNOWLEDGE VISUALIZATION

Extraction of (experts') tacit knowledge is an issue discussed in the literature. In general terms, Nonaka & Takeuchi (1995) suggest to use metaphors, analogies and models to externalize tacit knowledge into explicit knowledge, distinguishing these two types of knowledge and four types of knowledge conversions (Socialization, Externalization, Combination and Internalization)

In order to support the knowledge extraction, information visualization tends to address organizing and displaying complex information structures for the cognitive understanding of the human beings. (Card, Macinlay & Shneiderman, 1999). Meanwhile, according to Eppler and Burkhard (2004) this concept can be defined as the "use of visual representations to improve the creation and transfer of knowledge between people". The cognitive understanding for information visualization then highlights also these more intangible, tacit aspects as the creation and transfer of knowledge, demanding a combined concept for information, tacit and explicit knowledge visualization.

When we consider the knowledge visualization tools, most of them concentrates on Card's definition, visualization on information level (Boyak, Wylie & Davidson, 2002) (Jurasic & et al, 1998) (Wong & et. al,2006) (Wong & et. al, 2009). On the other hand, some visualization applications are directly related with explicit-to-tacit knowledge transformation (Hou&Tsai, 2008).

Various visualization systems have been used for the purpose of decision support like in Mansmann and Vinnik (2006). In this study, the main concern was the intrusion detection in a university gateway. System works on a online analytical processing (OLAP). Most of the OLAP tools are defined under business intelligence (BI) systems which are also categorized as decision support systems (DSS). Using OLAP tool under intrusion detection was an interesting idea, which requires dynamic data analyses, which, in fact, could be applicable in tacit knowledge visualization. Individual knowledge also evolves and declines with time, so following up this type of a dynamic structure is required.

After pre-processing of the explicit knowledge warehouse, how we can visualize system findings can be found, for instance, in the work of Hou and Tsai (2008). In this study, documentation of the computer assembly procedures are visualized with Virtual Reality Modeling Language (VRML). The idea is simple, a person can better learn with visualized aid rather than reading full documentation. With its possible natural language independency, the study underlines the potential for tacit knowledge visualization. However, their proposed system is mostly proper for virtualized information, not visualized ones.

In Dominguez et. al. (2007) study, for the visualization, the system uses a three layer structure: client, host and data-warehouse layers. Artificial Intelligence is created in the database layer which is composed of a database and a logic system. Logic system has a self learning mechanism via artificial neural network that is using Self Organizing Maps (SOM). This is important, because in the database, according to the gathered data-set information, these sets are organized by themselves, and with the relationship in each dataset (in here we can say clusters), they create information maps. The given system was created for a supervision system which is based on procedural knowledge of a production environment. However, the study focuses on explicit knowledge, and the tacit, cognitive part is missing.

SOM can be a basic technique for the construction of the framework. However, there are always minor possibilities where SOM could fail for visualization. For this reason, other data mining techniques could be needed, such as K-Means or DBSCAN. In Gupta et. al. (2010), an unsupervised clustering method is used and for its visualization, especially DBSCAN is used for creating different clusters.

Studies such as Jurisica et. al. (1998), Wong et. al. (2006) and Hou et al. (2008) then provide cases that applicability of created frameworks can be observed in various domains. Sloane (2007) also studied the knowledge that was required for the software projects. An experiment was conducted based on a framework. In this framework, based on the requirements of software projects' each step, there could be different knowledge requirements arising. Actually, this is important because this system is based on the cognitive knowledge of the individuals. Meanwhile, the study of Herrero (2010) gives us a knowledge based system. This system is based on the qualitative data of the organization, and this data is processed through Cooperative Maximum Likelihood Hebbian Learning (CMLHL) model. Also this study highlights the need of multidisciplinary study for this kind of study. In Jeong (2010)’s study, visual analytical systems are included for the tacit and explicit knowledge visualizations perspectives as a significant development for visualization studies and applications.

The dataset itself may cause problem with its structure. A dataset comes from a warehouse could bring redundant data, noise and missing data with itself. When building relationships, this could cause missing relationship, even worse relationship with loops which has no connection between other beginning and end of the visualized structure. In Kang et. al. (2008), a visualization system is used to find and visualize these abnormal and unexpected structure inside. When it is tried to extract tacit knowledge in explicit knowledge warehouses, similar problems could arise, and application of this concept could be helpful to address these problems.

Another important issue is handling the number of relations among entities for the knowledge visualization to address any possibility in real life. In a given problem domain, we can accept n as the total number of dimensions that reflects individuals' total number of tacit knowledge subjects, m as related dimensions, when problem arises, and k as number of individuals in the organizations. Then, an optimistic probability could require k*m (n>m) relationships in a considered subject, meaning the problem domain shifts to a multidimensional domain, number of which could be really high, and needs to be decreased. To address this issue of relation decreasing, for instance, least square projection can be applied. Accordingly, as applied by Paulovich et. al. (2008) it is possible to decrease these relations by dimension reduction.

Other different works also provide insights and information regarding the development of a proper methodology to support tacit knowledge extraction and visualization. For instance, Mercler (2007) reflects upon how the tacit knowledge is diffusing in the organizations. Although a software based approach is not used, the given data collection and analysis methods (critical incident technique, cognitive and reflexive interview, organizational network questioning and participating observations, among others) are insightful for developing a supportive technology and methodology for visualization of tacit knowledge in organizations. This also leads us to discussing further the related issues of extraction methods applicable to SRS process as a specific domain, as well as a general research methodology and design, as in the next sections.

EXTRACTION METHODS APPLICABLE TO SRS PROCESS

To address the challenging nature of tacit knowledge extraction, also a number of methods have been developed so far. Yet the problem is there are not many studies to distinguish between these methods so that they make easier to resolve the area specific problems experienced while a tacit knowledge extraction is in place. Each problem may require a different extraction technique to be applied. Some of these techniques are briefly described in this paper.

First of all, an **interview** is a face to face discussion between the Subject Matter Experts (SMEs) who posses the domain knowledge and Knowledge Engineers (KEs) who ask questions and/or observe the expert solving problems. (Fernandez,González &

Sabherwal,2004) Interviews are simply divided into Unstructured Interviews, Semi-Structured Interviews and Structured Interviews. A basic one-to-one interviewing technique, Output-Input-Middle method, is also suggested to be useful for extracting expert knowledge by KE. (Becerra-Fernandez, et al. 2004, Wu 2007)

- Output: Identify the answers or solutions to the problem under discussion (goals), focusing on understanding subtle differences between goals
- Input: Identify the sources of information that the expert uses to deduce the solution/answer, making sure how these inputs are identified, determined, or generated is known and understood
- Middle: Determine the links between the inputs and outputs that represent the core of the expert's knowledge (Some inputs may not be required initially, but may be requested later after the initial inputs are interpreted and Intermediate goals/hypotheses may be required to complete the connections)

To complement interviews, other elicitation techniques such as observational elicitation and role reversal can also be used (ibid), or Capability Review Sessions and commentary/think aloud problem-solving techniques can be incorporated, when appropriate. Sometimes, the recordings of the interactions between SME and KE could be handy, or the facilitation of a mediator between SME and KE could also be necessary.

As a simple tool for asynchronous communication, e-mail can be ranked as an effective and most important method in terms of knowledge conversion from one to another. (Harris, 2008) It is also identified as a topmost socialization tool for extracting and sharing unstructured knowledge. (Dfouni and Croteau, 2004) Email is a preferred since it reduces the time wasted during the interviews and gives the both sides to get prepared before the meetings. E-Discussion boards are also known as asynchronous communication tools which allow its members to post messages, ask or answer the questions online. E-Discussion boards are found to be useful for sharing beliefs and mental models of individuals. (Dfouni and Croteau, 2004) They also refer to bulletin boards or message boards typically.

More systematic tools for visualization can also complement these commonly accepted methods. For instance, a use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted. (Wikipedia,2011, Figure 1)

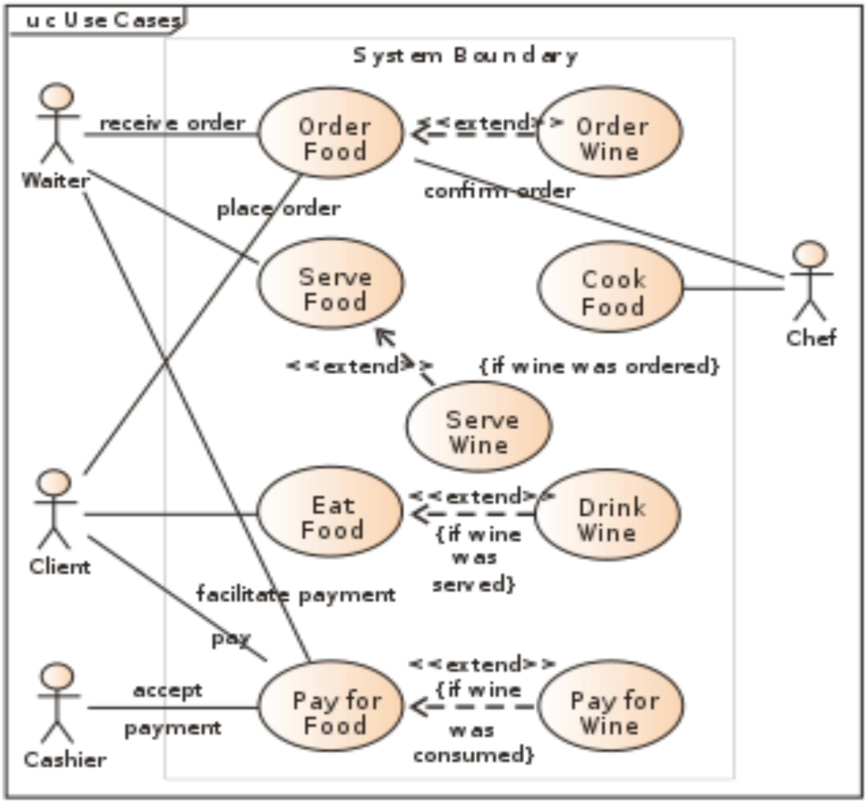

Figure 1: Example Use-Case Diagram (Wikipedia 2011)

Concept maps are also graphical tools for organizing and representing knowledge. They include concepts, usually enclosed in circles or boxes of some type, and relationships between concepts indicated by a connecting line linking two concepts. Words on the line referred to as linking words or linking phrases, specify the relationship between the two concepts. We define concept as a perceived regularity in events or objects, or records of events or objects, designated by a label. The label for most concepts is a word, although sometimes we use symbols such as + or %, and sometimes more than one word is used. Propositions are statements about some object or event in the universe, either naturally occurring or constructed. Propositions contain two or more concepts connected using linking words or phrases to form a meaningful statement. Sometimes these are called semantic units, or units of meaning. Figure shows an example of a concept map that describes the structure of concept maps and illustrates the above characteristics. (Novak & Cañas 2006, Figure 2)

Mind mapping is a quite novel technique for information visualizing. Put more simply, cognitive maps are a method we use to construct and accumulate spatial knowledge, allowing the "mind's eye" to visualize images in order to reduce cognitive load, and enhance recall and learning of information. Also another type called fuzzy mind maps are available. In this type mind maps have weights between relationships. These weights represent power of the relationship. (Kosko 1986, Figure 3)

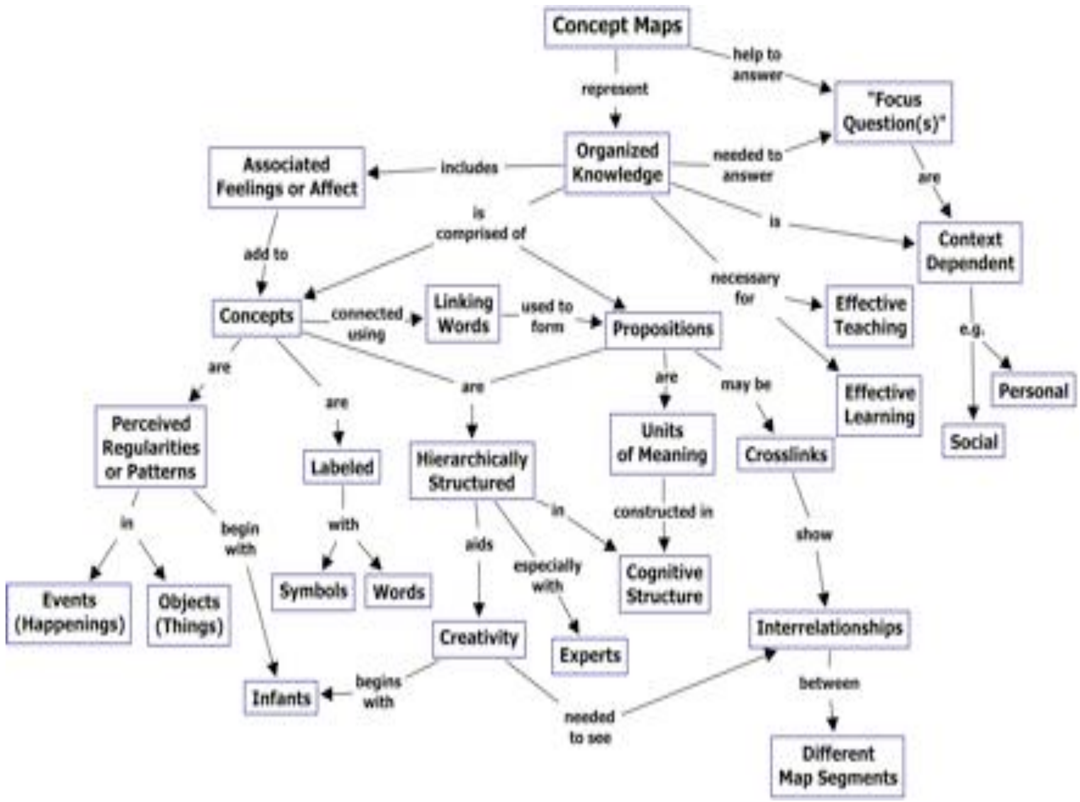

Figure 2: A sample concept map
Source: Institute of Human and Machine Cognition (2008)

Repertory grids can also be used as a list of specific characteristics of a domain that are to be evaluated by an expert

- Mathematically: an attribute-value vector,
- Attributes are also sometimes called elements or labels,
- Values can be binary or a range of values,
- A construct is an attribute-value pair (along with the specification of the range, i.e., set of allowed values) (Becerra-Fernandez, et al. 2004, and Wu 2007, Figure 4)

Automated tools exploit the idea of repertory grids by trying to help elicit what attributes are important for the domain, and what range of values the attributes should have.

There are various other simple of sophisticated methods or techniques that can be directly or indirectly applied for knowledge extraction, However, after this provision of selected general knowledge extraction and visualization methods, we can now analyze and discuss how we can address tacit knowledge extraction problems specific to SRS process. The analysis has benefited from our experience and review of literature (Wikipedia, 2011), as well as participant practitioner feedbacks from a two-week seminar in Knowledge Management and Technologies Graduate Class in 2010 in Informatics Institute, METU, Turkey.

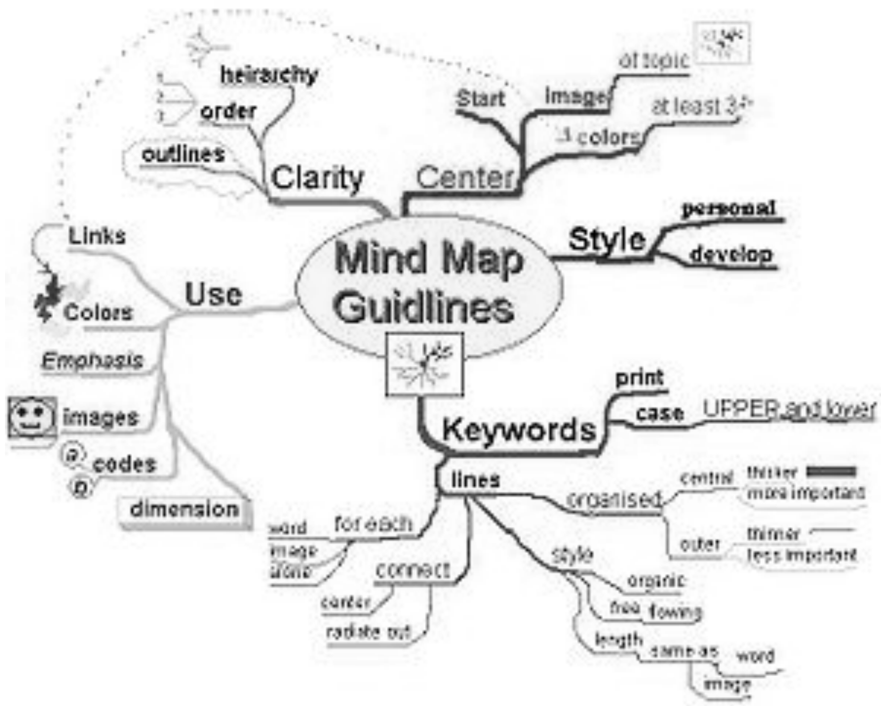

Figure 3: A sample mind map
Source: Study Habits (2011)

ELEMENTS 10, CONSTRUCTS 14, RANGE 1–5
PURPOSE: Staff appraisal

	Staff member No.	1	2	3	4	5	6	7	8	9	10	
1	Intelligent	1	1	4	5	3	3	5	2	3	5	Dim
2	Willing	1	2	4	5	1	1	4	3	1	2	Unwilling
3	New boy	1	2	3	5	4	4	4	1	4	3	Old sweats
4	Little supervision	3	1	4	5	2	1	5	2	2	3	Needs supervis.
5	Motivated	1	1	4	5	2	2	5	3	3	2	Unmotivated
6	Reliable	3	2	2	5	1	1	5	1	2	3	Unreliable
7	Mild	3	4	5	2	2	3	1	5	4	5	Abrasive
8	Idea person	1	1	5	4	2	3	1	3	4	4	Staid
9	Self-starter	2	1	5	5	1	3	5	3	4	5	Needs a push
10	Creative	1	1	5	5	2	3	4	3	4	5	Uncreative
11	Helpful	4	3	4	2	3	5	1	4	5	5	Unhelpful
12	Professional	1	2	3	3	2	1	5	2	4	4	Unprofessional
13	overall rating high	2	1	3	4	1	2	5	2	3	4	overall rating lo
14	Messer	2	2	5	4	3	5	1	5	3	1	Tidy
	Staff member No.	1	2	3	4	5	6	7	8	9	10	

Figure 4. A Sample Repertory Grid
Source: Becerra-Fernandez, et al. (2004), and Dekai Wu (2007)

Tacit Knowledge Extraction Problems Specific to SRS Process

Problems can be analyzed under two main categories which are User/Expert Based Problems and Engineer/Developer Based Problems. User Based problems are the problems faced due to the limitations or incapability of the users/experts (Wikipedia,2011). These problems can be viewed as:

- The users don't have a clear idea of their requirements,
- Some users are technically unsophisticated.
- Communication with the users is slow.
- Users often do not participate in reviews.

On the other hand, there are problems related to Engineers/Developers which makes the SRS Process inefficient or hard to complete. These kinds of problems can be viewed as:

- Technical personnel and end-users may have different vocabularies. Consequently, they may wrongly believe they are in perfect agreement.
- Engineers and developers may try to make the requirements fit an existing system or model, rather than develop a system specific to the needs of the client.
- Analysis may often be carried out by engineers or programmers, rather than personnel with the people skills and the domain knowledge to understand a client's needs properly.

The problems and suggestions addressing those problems are discussed in detail, as below:

1. User/Expert Based Problems

1.1 The users don't have a clear idea of their requirements

A common problem with the users is the degree of their awareness of what they want. Most of the users do not know what they really need and cannot express their thoughts. Sometimes, they do not even know that they know. In such cases, more visualization and face to face sessions are needed. In order to overcome this problem interview method must be applied, supported with the visual methods such as Concept Maps, Mind Maps and Use Case Diagrams. In the first meeting (the Kick-off Meeting) the expert(s) and the engineer(s) will have the chance to explore the domain and they will establish some ground rules and mutual understanding of expectations. Putting a Capability Review Session between the first meeting and the second meeting might be a good idea in order to make the expert to understand the limitations and capabilities of the system. This capability review session should be supported by use case diagrams and other diagram/visualization based techniques to make the users have more concrete understanding of the requirements and the system.

1.2 Some users are technically unsophisticated

Technically unsophisticated users may be a nightmare for the developers/engineers in some cases. This causes both sides not to understand each other and waste time having arguments. Since the user is unable to understand what can/ cannot be done, they may be requesting impossible, or hard and too late to develop characteristics or requirements. In this case, having a "mediator" who will act as an interface or bridge between the two sides is a must. The mediator will find a way to establish or improve the dialog between the parties. The efficiency of the agreement depends on the skills and the experience of the mediator. However, having a mediator only will not totally address the problem and a technical review between the meetings might be required to make the user familiar with the technical side of the task.

1.3 Communication with the users is slow

Another common problem is the rate of the communication between two sides. Sometimes, especially in large-scale projects, a variety of meetings might be needed. This may be either due to the scale of the modules to be covered or due to the unresolved parts in the previous meetings. In order to handle this problem e-mailing will be good solution. E-mail makes it easier to understand the unclear parts, to get prepared before coming to the meeting and to have an individual time to work on the parts and respond to the others.

1.4 Users do not participate in reviews

Not having all the necessary people in the meetings is an important problem for SRS process. The experts are mostly busy people and they usually have too many meetings to join. This will end up with the absence of the expert in the meeting and will probably cause to arrange some other meeting in some other time, wasting other participants' time. Another obstacle for the expert to join the meeting may be developer/hired company may be in a different city. In both cases having the required people might not be possible. To resolve this problem E-mailing, E-Discussion Boards, and/or Tele/Video Conferencing need to be used. Using any of these methods will reduce time wasted by the participants.

2 Engineer/Developer Based Problems

2.1 Technical personnel and end-users may have different vocabularies

It is a common problem that users and technicians using different languages. In most of the cases, users don't understand technicians' vocabulary and technicians don't know users' domain. There are some methods to solve this issue. First of all a mediator who have information about both sides would be very effective. Also in semi structured interviews, technical team and users learn each other's languages. For a productive interview, technical team should make a domain research before meeting. The main problem of research is distinguishing which term/knowledge is important and which is not in the domain. We suggest using tag clouds (Wikipedia,2011) as a special mapping/visualization tool to handle this issue. A tag cloud of the domain will represent important terms of the domain.

Collaborative mind mapping would be another method for mediation of user and technician. When they try to create a mind map together they will understand each other's point of view. Teach-back methodology (Epistemics, 2003) can also be our final suggestion for this problem. In this method domain expert (end-user) comments on what the engineer describes about the learning of previous sessions. Thus engineer could correct misunderstandings.

2.2 Engineers and developers may tend to fit to an existing model rather than a new specific product

To summarize this problem, "if you have a hammer everything starts to seem like a nail" could be said. Sometimes assimilating one problem to another solved one might be helpful. But if they are not really similar, lots of time will be wasted for wrong analysis. Commentary/think aloud problem-solving technique could be a solution for this case. When engineer thinks aloud user/domain expert could correct his/her assumption about the problem. One drawback of this solution is cognitive overload on user. To overcome, off-line reporting using a recordable media (video, etc.) could be used.

Also observation technique can be used to understand real process in domain. In observation technique engineer observes the expert and take notes while they perform the tasks. This technique could also be supported with video recordings.

RESEARCH METHODOLOGY AND DESIGN FOR DEVELOPING FRAMEWORK PROPOSAL

While we exemplify the importance of extracting experts' knowledge with respect to SRS domain above, we aim to develop a general framework that can be used for tacit knowledge extraction from the explicit knowledge structure of the organizations. The requirements for the framework development are determined with the methodology. Two basic types of knowledge, tacit and explicit, are then important for our research methodology.

Basically tacit knowledge is considered as the knowledge of individuals who are not sure what knowledge they posses (Polanyi, 1966). Ernst and Kim (2002), identifies four types of tacit knowledge in addition to the tacit and explicit knowledge types by Polanyi (1962).

1. Embodied knowledge refers the type of tacit knowledge that may become part of the human body as skills.
2. Embrained knowledge refers the type of tacit knowledge that may become part of human being as coginitive capacity.
3. Embeded knowledge refers the type of tacit knowledge that is routinized in organizational practice.
4. Encultured knowledge refers the type of tacit knowledge that is inculcated in the organization as basic assumptions as beliefs and norms. (cited from Durukan & Pamukçu 2011)

Also, knowing where or who could posses the required knowledge lead another tacit knowledge of other individuals. In here it is needed to show individual's relationships with other individuals in the organization. This leads to a qualitative approach for the research (Cresswell, 2009). On the other hand, when organizational hierarchy and procedures related with the internal organizational explicit knowledge are considered to understand numerical descriptions, trends, attributes of individual members and generalize findings to the whole organization it, a quantitative approach will be required (Cresswell, 2009).

The qualitative and quantitative part of this research leads to a mixed methodology. Accordingly, triangulation design can be applied (Cresswell, Clark, & et. al, 2003) to obtain different perspectives of the organization and use these perspectives to gain complementary data (Morse, 1991). Under the triangulation design, the convergence model can be applied for supporting the flow of qualitative and quantitative parts together. Accordingly, the below figure (Figure 5) illustrates the flow of this study.

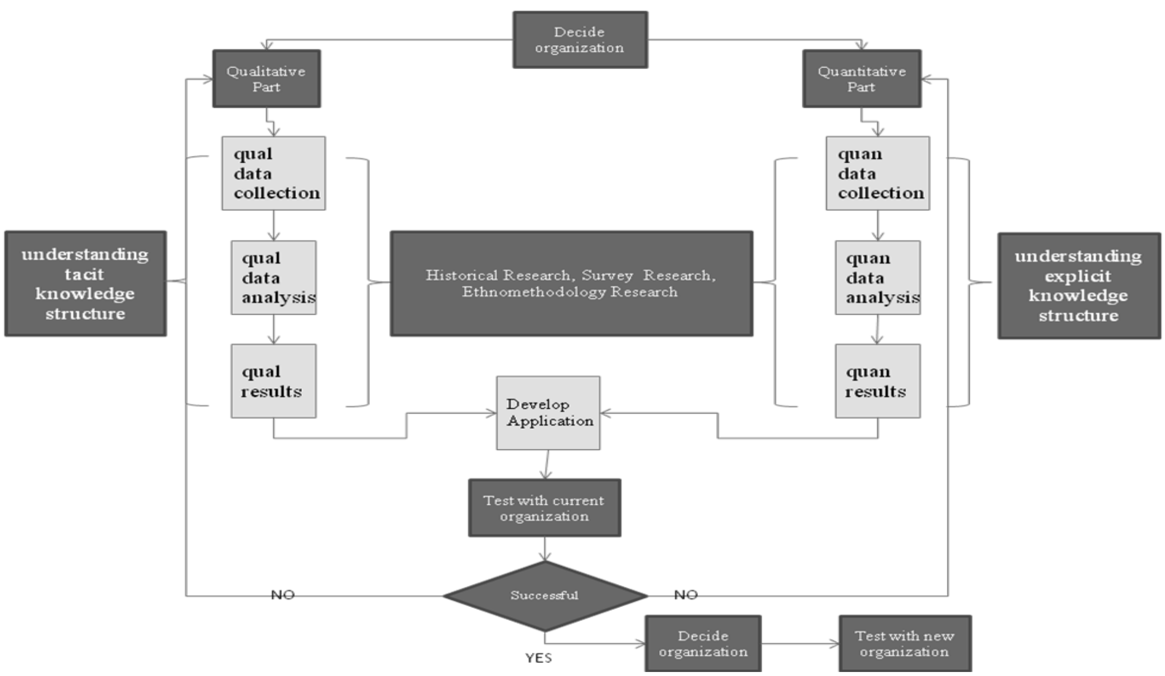

Figure 5. Flow of Study

This figure could actually be divided into three phases. First Phase will be started from deciding the organization and will be end with the development of an application. Second Phase will include developing application and testing it with current organization. Third Phase is to continue further tests with other organizations.

Phase 1: Decide and Collect

This study will be started with the selection of an organization. In the defined organization at least five primary functions must be established inside. These functions are human resources, accounting and/or finance, strategy decision, information systems, and main function as defined by mission and vision statements of the organization.

After deciding the organization, to understand tacit and explicit knowledge structure, a sequential flow will be required. Accordingly, first of all, "Historical Research" will be conducted on organization and department-related documentations. This research will give us the possible source of procedural and hierarchical structure and possible explicit knowledge warehouses of the organization. This will lead to the comprehension of the departmental structure of the organization that will help to decide main departments. For collecting quantitative data of the defined departments, "Survey Research" to the management will be conducted. Then a sample of the department will be selected and "Ethnomethodology Research" will be conducted on them. With the result of "Historical Research" and "Ethnomethodology Research", the data will be brought together, giving the qualitative data. For this phase the suggested instruments could be given as the following table (Table 1)

Table-1 **Research Type and Basic Instrument**

Research Type	Main Instruments	Explanation
Historical Research	Documents	The documents related with the organization will give the explicit and tacit information related with organization. For the documentation, the internet documentation related with organization, inside organizational documentation and internal information system that is used for knowledge storage, sharing and creation will be included in this research.
Ethnomethodology Research	Interview and questionnaire and observation	To understand the individuals' understanding of the explicit knowledge of the organization and how they include their tacit knowledge to this structure will be examined based on their daily activities. Interviews and specifically prepared questionnaires will be implemented.
Survey Research	Interview and questionnaire	To understand what the related department do and bring main explicit knowledge subject into the ground interviews and questionnaires will be implemented to the management personnel of the department.

The collected data will define the structure of the framework. And this framework will be implemented in the application. The details are given under the second phase.

Phase 2: Build the System and Check Internal Validity and Reliability

This phase corresponds two sub-phases, building the system, and checking its internal validity and reliability. A candidate framework will be visualized in a software base, as illustrated in Figure 6.

After the data analysis are concluded, the results will provide the first artificial neural network structure which is called explicit subject extraction/inclusion engine. To train this engine, the explicit knowledge structure of the organization will be used.

The dynamic tacit knowledge structure will be created with the information of individuals such as education, family and previous work experiences. The tacit knowledge extraction engine will be used to create and structure subject-based or individual-based tacit knowledge clusters. Here, because of the possible delay/decay of the individual knowledge, the tacit knowledge structure requires to be updated by this engine.

Tacit knowledge visualization will use the data created by the tacit knowledge extraction engine, and will be displayed based on subject or individual relational base. Formal organizational knowledge visualization will depend on explicit subject extraction/inclusion engine. This will be used to visualize formal structure of the organization, and compare the explicit structure with the tacit structure.

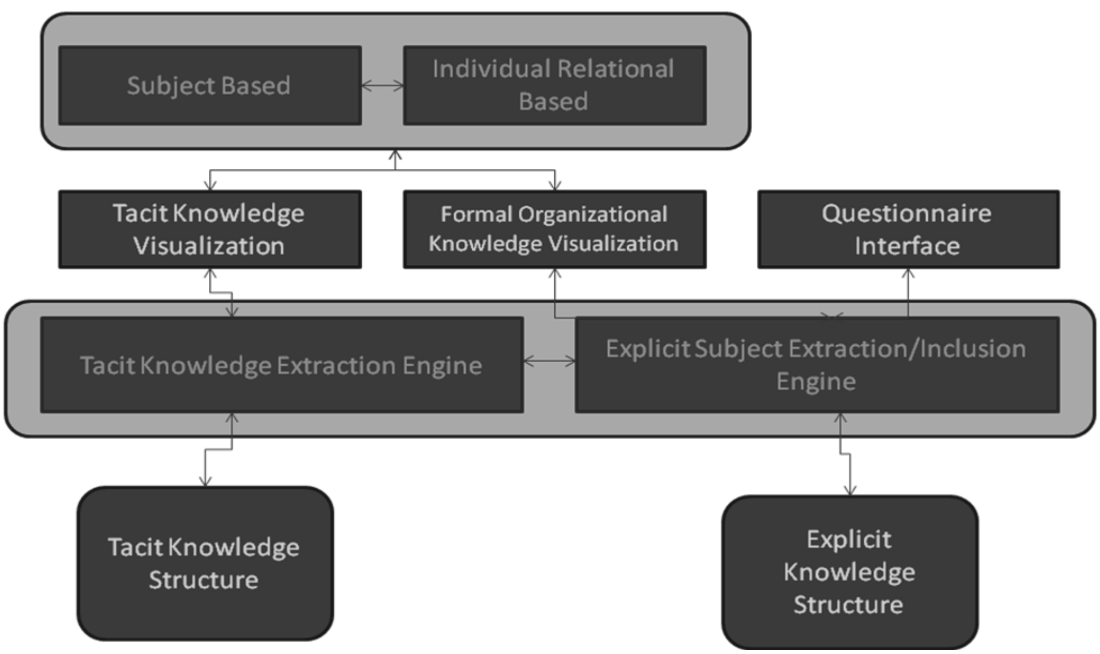

Figure 6. Tacit Knowledge Visualization Framework

Table-2 Framework Components and Main Instruments

Component	Main Instruments	Explanation
Subject and Individual Relational Based	Least Square Projection*	The number of dimensions could be more than 3, so to visualize the related tacit knowledge or tacit knowledge with the explicit knowledge relationship it will require projection
Tacit Know. Visualization	Self Organized Maps	Clustering tacit knowledge according to given subject or individual knowledge could require self organized maps
Tacit Know. Extraction Engine	Data Extraction/ Inclusion/Update	From the subject extraction/inclusion, data structure of the tacit knowledge structure will be extracted/included and updated according to requirement
Formal Organizational Knowledge Visualization	K-Means Clustering	According to distances between explicit subject distances to the related subjects and organizational structure, the explicit knowledge and subject will be visualized
Explicit Subject Extraction/Inclusion Engine	Data Extraction/ Inclusion/Update	From the given questionnaire results, explicit knowledge structure will be able to extract, include or update, when it is needed. Also the main subjects are determined inside of this structure.
Questionnaire Interface	Online Questionnaire	After first collection of the survey and interview, the updated version of the questionnaire will be uploaded and individuals will be able to access.

*For tacit knowledge visualization, a preparation stage will be required so that this least square projection method could be developed and used as one of the system module.

For the internal validity and reliability testing of the system and the validity of the questionnaires, an external interface will be created for the remaining members of the organization. (The initial sample of members was used in the first phase to collect qualitative and quantitative data). Reliability will be measured through re-sampling of the first group who will be asked to fill the questionnaires again from the computer interface. Using the whole first sample or just a part of it will be defined, after testing of the system. For this phase, the instruments for the each component could be given as Table 2.

Phase 3: Decide to Go on or Return to the Beginning

If the framework is visualized as expected, the test of the external validity will start with the existing application. If not, it is required to update the questionnaires and test it until an internally valid and reliable framework can be built.

Within this challenging mixed methodology, and after the completion of these phases of research, we can summarize the research questions and possible answers that will come with the proposed framework as the Table 3. In order to address these questions, this paper has first presented a brief overview of the theory and related studies of the knowledge visualization. Then, the information on research design and methodology, as well as the proposed framework design as the result of the possible execution of the methodology has been provided. As a result, an intelligent framework dynamic and specifiable for visualizing the tacit knowledge structure of the organizations from the perspectives of individual members and their relationship between each other and explicit subject bases has been proposed.

Table 3. Research Questions and Possible Answers' Sources

Research Question	Possible Research Answer Source
How the tacit knowledge structure of the organization could be visualized with the existing explicit knowledge warehouse of the organizations?	The historical data on the organizational documents and the given Information System infrastructure of the organization will reveal the explicit knowledge structure. With the inclusion of Ethnomethodology Research data the existing tacit structure inside the explicit structure will be visualized or mapped
How reliable the predetermined formal organizational structures are, when the informal structure of the organization is considered?	If the main subjects of the organizations are properly visualized with the existing tacit structure, it will give the reliability value of the organization's formal structure.

CONCLUSION

The content of the paper has been organized as follows: First, we have made a brief introduction to the general knowledge visualization, as well as tacit knowledge extraction methods/approaches that are applicable to the software requirement specification area.

Then, we have identified the extraction problems specific to the software requirement specification by dividing the problems into two categories, namely User/Expert Based Problems and Engineer/Developer Based Problems, providing a deeper analysis on and suggest solutions to each problem. While we have tried to identify new approaches as solutions to existing problems, we have also found useful certain general methods as applicable to solve specific problems in different domains. However, these methods and problems are not mutually exclusive, or suggested methods and experienced problems do not necessarily match. It should then be acknowledged that sometimes just common sense and simple tools could work, and it is always good to allow certain buffer space and time for sorting out misunderstandings and conflicts that could arise between SMEs and KEs, among whom committed and trusted interaction is important. Accordingly, most of the discussed methods and problems, as well as other new ones can also be found in different domains' of literature, which can be revealed by future research on reviewing academic and practitioner literature.

We should also note that this paper mainly reflects co-authors' subjective experience and analysis. We aim to apply the results of this paper to our own practice. We also believe these results and findings would be useful for others that study practitioner experience or academic literature on tacit knowledge extraction of experts for SRS or related topics.

Finally, the information on research design and methodology, as well as the proposed framework design as the result of the possible execution of the methodology has been provided. For future work, questions such as how reliable the predetermined formal organizational structures are, when it is considered the informal structure of the organization, as well as whether the informal structure of the organization could be visualized with the social network hierarchy could also be considered. An application into a suitable and feasible small or medium enterprise could also be suggested for an initial test of the framework.

ACKNOWLEDGEMENT

Parts of this paper is adapted from a proposal submitted to METU Informatics Institute, as a partial fulfillment of class work for the Research Methodology course instructed by Soner Yıldırım in Summer 2010.

REFERENCES

Bart Kosko, *Fuzzy Cognitive Maps*, International Journal of Man-Machine Studies, 24(1986) 65-75 (first introduction of FCMs)

Becerra-Fernandez, et al (2004), Dekai Wu (2007), Knowledge Management www.cs.ust.hk/~dekai/600G/notes/KM_Slides_Ch10.pdf [Accessed 12.02.2011]

Boyack K.W., Wylie B.N., & Davidson G.S. (2002). Domain visualization using VxInsight (R) for science and technology management. *Journal of the American Society for Information Science and Technology,* 53(9), 764-774. doi: 10.1002/asi.10066

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). Information visualization. In S. K Card, J. D. Mackinlay, & B. Shneiderman (Eds.), *Information visualization. Using vision to think* (pp. 1-34). San Francisco: Morgan Kaufmann.

Cresswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). SAGE.

Cresswell, J. W., Plano C., & et al.,(2003). Advanced mixed methods designs. In A. Tashakkori & C. Teddlie (eds). *Handbook of mixed methods research in the social and behavioral sciences* (pp. 209-140). Thousand of Oaks, CA: Sage.

Dalbello M., & Spoerri A. (2006). Statistical representations from popular texts for the ordinary citizen, 1889-1914. *Library & Information Science Research,* 28(1), 83-109. doi:10.1016/j.lisr.2005.11.015.

Dominguez M., Fuertes J.J., Reguera P., & et al. (2007). Internet-based remote supervision of industrial processes using self-organizing maps. *Engineering Applications of Artificial Intelligence*. doi:10.1016/j.engappai.2006.11.017.

Durukan, C. & Pamukçu, T. (2011). Global Production Networks and Knowledge Transfer Mechanisms. ICEBEG Conference, İzmir, Turkey

Epistemics (2003), Knowledge Acquisition, http://www.epistemics.co.uk/Notes/63-0-0.htm [Accessed 7.01.2011]

Eppler, M., & Burkhard, R. (2004). Knowledge Visualization: Towards a New Discipline and its Fields of Applications. Working Paper, Wirtschaftsinformatik. Retrieved from http://www.wirtschaftsinformatik.de/wi_arbeitsberichte.php?op=anzeigearbeitsbericht&abid=142.

Ernst, D. and L. Kim (2002), "Global production networks, knowledge diffusion and local capability formation", *Research Policy*, 34, 1417-1429.

Fernandez,I.B.,González, A.J.,Sabherwal. (2004). Knowledge management: Challenges, Solutions, and Technologies.,Pearson/Prentice Hall.

Gourlay, Stephen (2002), "Tacit knowledge, tacit knowing or behaving?", Kingston Business School, Kingston upon Thames,UK

Gupta G., Liu A., & Ghosh J. (2010). Automated Hierarchical Density Shaving: A Robust Automated Clustering and Visualization Framework for Large Biological Data Sets. *IEEE-ACM Transactions on Computational Biology and Bioinformatics*, 7(2), 223-237.

Herrero A., Corchado E., Saiz L., & et al.(2010).DIPKIP:A connectionist Knowledge Management System to Identify Knowledge Deficits in Practical Cases. *Computational Intelligence*, 26(1). 26-56.

Hou J. L., & Tsai A.W.J. (2008), Knowledge reuse enhancement with motional visual representation. *IEEE Transactions on Knowledge and Data Engineering*, 20(10),1424-1439.

Institute for Human and Machine Cognition (2006), The Theory Underlying Concept Maps and How to Construct and Use Them, http://cmap.ihmc.us/publications/researchpapers/theorycmaps/theoryunderlyingconceptmaps.htm [Accessed 23.02.2011]

Irick, Michael L. (2007), "Managing Tacit Knowledge In Organizations", *Journal of Knowledge Management Practice*, Vol. 8, No. 3, pp.1-5.

Jeong, D. H. (2010). *Knowledge Visualization: From Theory to Practice* (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3404838)

Jurisica I., Mylopoulos J., Glasgow J, & et al .(1998). Case-based reasoning in IVF: prediction and knowledge mining. *Artificial Intelligence in Medicine*, 12(1), 1-24. doi: 10.1016/S0933-3657(97)00037-7.

Kang HM, Getoor L, Shneiderman B, & et al.(2008). Interactive entity resolution in relational data: A visual analytic tool and its evaluation. *IEEE Transactions on Visualization and Computer Graphics*,14(5),999-1014. doi:10.1109/TVCG.2008.55.

Koskinen K.U., & Makinen S. (2009). Role of boundary objects in negotiations of project contracts. International Journal of Project Management, 27(1), 31-38. doi:10.1016/j.ijproman.2007.10.006.

Makovoz, M. (2008). *Utilization of Semantic Linking and Visualization Techniques to Facilitate Knowledge Creation from Textual Software Artifacts* (Doctoral dissertation). Available from ProQuest Dissertations & Theses database. (UMI No. 3297657).

Mansmann, F., & Vinnik S. (2006). Interactive exploration of data traffic with hierarchical network maps. *IEEE Transactions on Visualization and Computer Graphics*, 12(6), 1440-1449.

Memon, Nasrullah , Knowledge Elicitation – Converting Tacit Knowledge to Explicit, http://www.mip.sdu.dk/~memon/10-2.pdf [Accessed 7.01.2011]

Mercier, D. (2007). *Le transfert informel des connaissances tacites chez les gestionnaires municipaux* (Doctoral dissertation). Available from ProQuest Dissertations & Thesis database.

Morse, J. M. (1991, March/April). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research,* 40(1), 120-123.

Nonaka, Ikujiro; Takeuchi, Hirotaka (1995), The knowledge creating company: how Japanese companies create the dynamics of innovation, New York: Oxford University Press

Paulovich F.V., Nonato L.G., Minghim R, & et al.(2008). Least square projection: A fast high-precision multidimensional projection technique and its application to document mapping. *IEEE Transactions on Visualization and Computer Graphics* , 14(3) , 564-575. doi: 10.1109/TVCG.2007.70443.

Polanyi, M. (1966). *The Tacit Dimension*, Double Day.

Polanyi, M. (1962) *Personal Knowledge: Towards a Post-Critical Philosophy.* University of Chicago Press, Chicago

Sloane, J. J. (2007). *Showing the Design/Business Relationship Visual Tool to Moderate Perceived Risk* (Master Degree Project). Available from ProQuest Dissertations & Theses database.

Study Habits (2011), How To Mind Map http://www.study-habits.com/how-to-mind-map [Accessed 12.02.2011]

Tagger, Ben, An Enquiry into the Extraction of Tacit Knowledge, http://www.cs.ucl.ac.uk/staff/B.Tagger/FinalIMR.pdf [Accessed 7.01.2011]

Wikipedia (2011), Requirement Analysis, http://en.wikipedia.org/wiki/Requirements_analysis [Accessed 7.01.2011]

Wikipedia (2011),Use case diagram, http://en.wikipedia.org/wiki/Use_case_diagram [Accessed 07.01.2011]

Wikipedia (2011),Tag Cloud, http://en.wikipedia.org/wiki/Tag_cloud [Accessed 07.01.2011]

Wong P. C., Foote H., Mackey P.,& et al. (2006). Generating graphs for visual analytics through interactive sketching. *IEEE Transactions on Visualization and Computer Graphics*,12(6), 1386-1398.

Wong P.C., Schneider K., Mackey P., & et al.(2009). A Novel Visualization Technique for Electric Power Grid Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 15(3),410-423.