Conceptual Design for Personalized Situation-Aware Information Seeking and Retrieval Services in Data Intensive Domains

Nong Chen, Delft University of Technology, 2628 BX, Delft, The Netherlands; E-mail: nongc@tbm.tudelft.nl
Ajantha Dahanayake, Georgia College and State University, USA; E-mail: ajantha.dahanayake@gcsu.edu

ABSTRACT
Personalized information seeking and retrieval is regarded as the solution to the problem of information overload. Personalization algorithms and techniques are maturing, but their centralized implementation solutions are becoming less efficient for dealing with ever changing user information needs in data intensive, dynamic and distributed environments. In this paper, we present a conceptual design for personalized, situation-aware information seeking and retrieval services, which offers a new way of thinking about the retrieval of personalized information for time critical applications.

Keywords: *personalization, information seeking and retrieval, situation awareness, service-oriented architecture*

1. INTRODUCTION
Today information is distributed globally. Organizations and information seekers can now access information anywhere anytime via wired or wireless networks. The coordination of organizational processes distributed over various independent locations has improved significantly. However, this increase in information availability cannot guarantee that users are able to retrieve and access all the information that specifically meets their needs. Organizations have to operate in increasingly complex environments, where, dynamically changing information structures, and distributed heterogeneous data resources make it increasing difficult to find the 'right information' at the 'right level of details' in the 'right format' , and at the 'right time'. The problem is amplified in domains such as crisis response networks, medical and healthcare networks, where the accuracy of the retrieved information and obtaining it in a time critical manner are extremely important.

Personalized information seeking and retrieval is regarded as the remedy when it comes to relieving the problem of information overload. Research in the field mainly focuses on developing models, methods and tools that can be used to adapt information content and its presentation to the individual user's information search goals and preferences. Today, personalized techniques are able to provide feasible solutions to the problem of inappropriate information overload at technological level ranging from simple user-controlled information personalization to autonomous system-controlled adaptation [Mizzaro & Tasso, 2002]. Many academic and commercial off-the-shelf information search services and tools are available to filter out irrelevant information effectively, and to rank and present information in a user preferred way.

Although these personalization techniques and algorithms are maturing, their centralized implementation solutions are becoming less efficient for supporting the rapidly changing information needs in dynamic and distributed environments. Changes in organizational or personal information needs may lead to a need to redesign a complete application. Therefore, it is necessary to design personalized information seeking and retrieval application in an agile and flexible way, to provide personalized information in a time critical manner, which can be adapted easily to meet changing information needs. Our increasing understanding of component-based design principles, service-oriented architectures and other aspects of complex information retrieval achievements provide the possibility to achieve this objective. The modularization of complex systems into components, or services that interoperate primarily via exchanging standardized messages at interfaces has become IT technology evolution. All these advances in IT have stimulated a new requirement concerning on modeling and designing personalized information search services in a dynamic and distributed environment to satisfy high-level functional requirements for personalization flexibly in one hand, and to use the specific personalization techniques, algorithms and available technology infrastructure to realize these requirements on the other. Building on the advantages of a service-oriented approach, we propose a new conceptual design for personalized, situation-aware information seeking and retrieval services. We believe this approach provides a new way of thinking about the retrieval of time critical, personalized, situation-aware information in data intensive domains.

Our conceptual design is presented in section 2. We tested and evaluated our conceptual design by applying it in a real-life case. The results are presented in section 3. Our conclusions and an outline for future work are presented in section 4

2. CONCEPTUAL DESIGN
2.1 The Concept of Information Seeking and Retrieval
Information seeking and retrieval is a human-IT system interaction activity in the sense of an IT supported environment. According to research into user information behavior, users' information needs are stimulated when they lack the information required to solve a problem [Wilson, 1998]. Problem solving occupies an important place in the research on information seeking and retrieval [Gaslikova, 1998]. [Gaslikova, 1998] summarizes problem solving in the context of information seeking and retrieval in three stages: problem identification, query formulation, and validation of received information. She claims that any information seeking and retrieval system has to provide suitable software tools for realizing each stage of a problem-solving process. Taking this into account, we regard information seeking and retrieval as a user-system interaction process. It comprises three sub-processes shown below.

- *The information seeking process*: This is the process used, generally or precisely, to identify the problem that initializes a user's information acquisition activity. This covers the question: what is the problem?
- *The information searching process:* This is the process used to identify and structure the specific information required to solve the problem, and to find the locations of required information. It concerns on the questions what information is required? Where to find the required information? And how to access to it?
- *The information obtaining process:* This is the process used to obtain the required information from the various resources and to return the information back to the user.

Users validate the received information during the interaction process with the system.

2.2 Personalized Information Needs

Personalized information needs play an essential role in determining the relevance of any delivered information. The literature shows a deepening understanding of the concept information needs and its role in the information seeking and retrieval. [Taylor, 1968] and [Belkin, 1982] argue that users' characteristics, e.g. users' personality, knowledge, personal interest and preferences etc., determine their information needs. [Wilson, 1994] and [Niedźwiedzka, 2003] shows that social placement or a professional role, e.g. users' professional roles connected with their occupied positions, are the most important determinants of users' information needs. [Wilson, 1999]'s problem solving model abstracts the process of problem solving from context.

Since our research is confined to domains such as crisis response and management networks, healthcare and medical networks, etc., we argue that individuals' personal interests and preferences may not strongly influence their information needs. Although different users in these contexts may have different knowledge levels about their professional role, we consider that their knowledge will be inherent to the professional role they play within their work situations. Therefore, we regard the professional roles users adopt in the society to be one of the most important factors influencing their specific information needs although their personalities and/or knowledge may influence their search strategies.

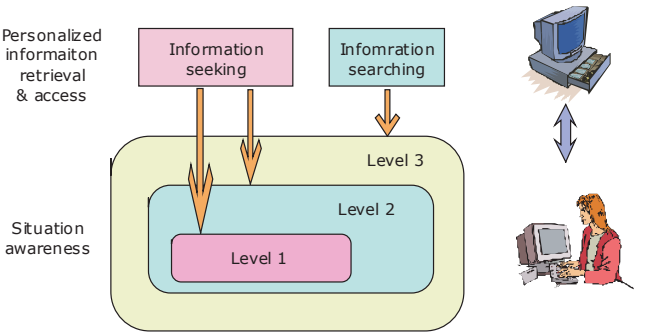
However, users' roles are not sufficient to reflect their actual needs. Information needs change as the users' situation changes, and this directly influences users' judgment regarding information relevance. Role-based information needs are relatively stable and predictable, while an ongoing situation may change these information needs. When users are able to be aware of their situation, in which they are involved, the users are able to identify their problem, and to formulate their information needs. Based on our previous arguments, we believe that personalized information seeking and retrieval is triggered by users' role-based perceptions of a problem they need to solve. We therefore define personalized information seeking and retrieval as an information acquisition process, aimed at providing users, with their role relevant personalized information based on dynamically and automatically perceiving and responding to the users' situations.

2.3 Situation Awareness

2.3.1 Situation Awareness in the Context of Information Seeking and Retrieval

The concept situation awareness (SA) is usually applied to operational situations, especially in the fields of artificial intelligence, agent-based systems, crisis response and management, military planning, etc. [Endsley, Bolte & Jones 2003]. Research in the field of SA focuses mainly on helping persons to be aware of their situations so that they can make informed decision about future actions [Endsley, Bolte & Jones 2003]. [Endsley & Rodgers, 1998] formally defines SA as "the perception of elements in the environment along with a comprehension of their meaning and along with a projection of their status in the near future" [Endsley & Rodgers, 1998]. This definition breaks down into three separate levels [Endsley, Bolte & Jones 2003].

Figure 1. Information seeking and retrieval to support SA process



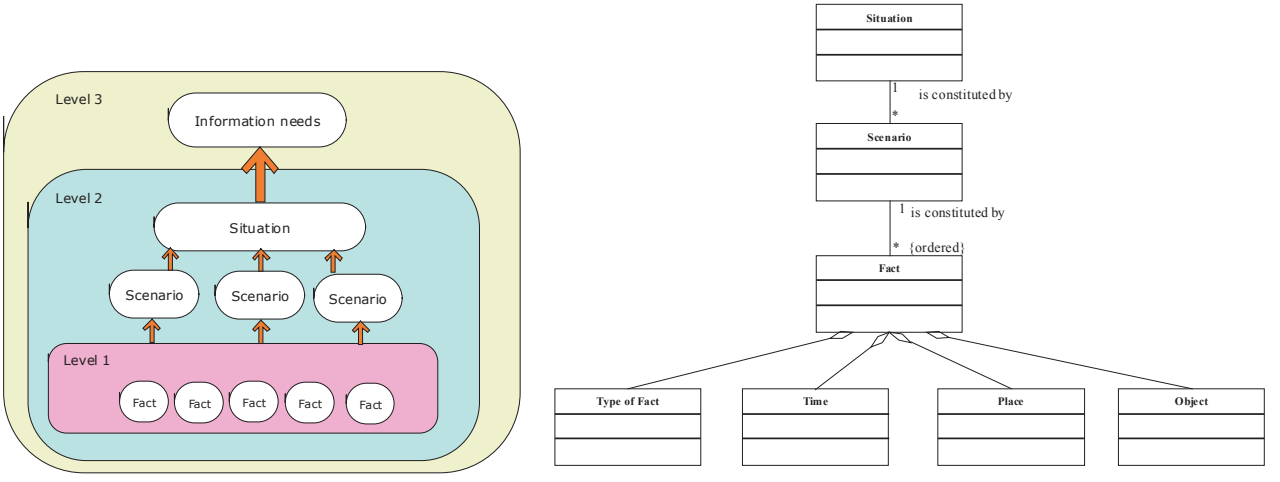
- Level 1—perception of the elements in the environment
- Level 2—comprehension of the current situation, and
- Level 3—projection of future status.

These three levels reflect the process of how people are aware of their situations mentally. Although today's advanced IT technology can replace a huge amount of information processing work, until now, it cannot replace a human's mental information processing process. We regard the information seeking and retrieval provided in our approach as a means to provide support for the users' SA process. We claim that our definition of information seeking and retrieval can be used to provide suitable support for all three steps of SA defined by [Endsley, Bolte & Jones 2003] shown in Figure 1. Levels 1 and 2 of [Endsley, Bolte & Jones 2003] are supported by the information seeking process, aimed at identifying the problem. Level 3 is supported by the information searching process, aiming at defining what information is needed.

2.3.2 Situation Model

Users perceive the problem they need to solve when their situation is described clearly. We choose to use the term situation since it can imply dynamic changes in users' surroundings, the influence of changes on the users, and the users' stable or permanent surroundings. Situation is defined in the [Merriam-Webster] as "the relative position or combination of circumstances at a certain moment". In the [Dictionary], situation is defined as "the combination of circumstances at a given

Figure 2. SA process



(a) (b)

moment; a state of affairs". From these definitions, we can see that, a situation consists of a concrete story describing where what happened, when, who are involved, and their relationships.

It is not feasible to specify all possible situations in the context of information seeking and retrieval for any domain. Historical usage data is required to detect situations. This argument is also used in [Endsley, Bolte & Jones 2003]'s three levels of SA model, where an unknown situation is derived from known information. The question, what historical data or information is required to be collected at different level of the SA processes for the users to realize their situation, becomes important; and can be reformulated as what information can be used to describe and model the situation. Based on the [Endsley, Bolte & Jones 2003] 's three levels of SA model, we need to perceive the elements in the environment (Level 1 in the SA) as the information that can be used to comprehend current situation (Level 2 in the SA), in order to project future status (Level 3 in the SA).

The first step in the SA process is to perceive the elements in an environment. The information elements that can be directly perceived describe the things that are known to have happened or to exist, i.e. the concept fact. Therefore, the information describing those things that have happened or existing things can be abstracted and conceptualized as a combination of 'type of fact', 'time', 'place' and 'involved objects'. The description of a fact can be made by combining 'type of fact' and any or all other three concepts.

Perceived facts are only direct observations made in the environment. They cannot provide narrative descriptions of users' situation. Therefore, facts do not supply sufficient information to understand a situation fully. To support the second level of the users' SA, we use the concept scenario in our research. We define *scenario* as a *short story reflecting a situation*. In this scenario, known outcomes are described, in conjunction with the casual relationships of a group of detected facts. Unknown scenarios can be derived by combining known facts, or from known scenarios. The SA process in our research is presented in Figure 2 (a), and the situation model is presented in Figure 2 (b).

2.4 Service-Oriented Approach for Personalized Information Seeking and Retrieval

A service-oriented approach provides a design principle for the handling of complex, dynamic and distributed information systems. Personalization can be provided by composing existing services, which are implemented on the basis of well-defined service behaviors and interfaces. This design principle is suitable for a distributed, dynamic and heterogeneous environment. According to the service-oriented system design principle, we assume that a solution for the detected information needs in users' specific situations can be quickly reconfigured by using a composition of encapsulated, replaceable and reusable services [Stojanovic, Dahanayake & Sol, 2004]. Based on sub-sections 2.2 and 2.3, we can now formulate the complete framework for personalized, situation-aware information seeking and retrieval, as shown below in Figure 3.

2.4.1 The Concept Service

According to the literature [Papazoglou, 2003, Stojanovic, Dahanayake & Sol, 2004, Douglas, 2003. etc], in the initial phase of our research, we informally defined the term *service* as "*A service is planned and designed in such a way that it has a specific functionality and it is very simple, but together they perform relatively complex tasks*". This informal definition gives the basic requirements for defining a service in the service-oriented approach. A service must have a specific functionality. As mentioned before, the solution for a detected problem is constituted of a service or a combination of services. Therefore, the specific functionality a service must have in our design is that it provides information. As a result, we simply define the services that consume information and provide information as information services. We assume that the solution to satisfy user's information needs is constituted by an information service or a group of information services. The information provided by a group of services is the collective outcome of all involved information services instead of a simple combination of outcomes of each service. Information services can be assembled and composed by smaller information services. At the level of a simple information service, its required operation and output is realized by grouping a specific collection of information search software components. We present the concept of service in Figure 4.

Figure 3: A framework of personalized situation-aware information seeking and retrieval

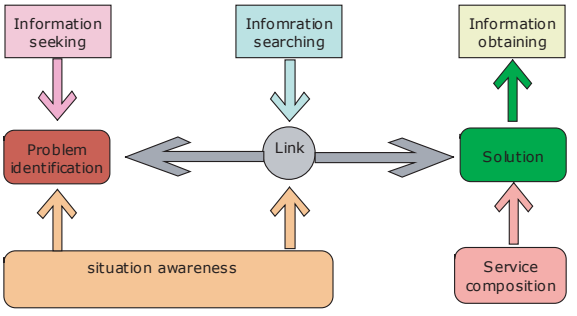
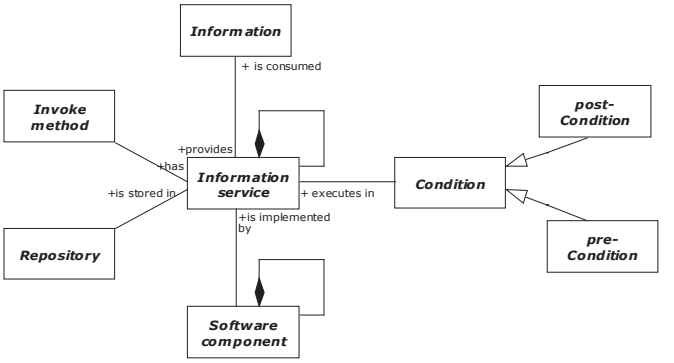


Figure 4. The concept of information service



2.4.2 Task Model

After the appearance of task-oriented information seek and retrieval in the 1990s [Vakkari, 2003], the concept task has been diligently used in the context of information search as [Byström & Hansen, 2005]. The concept task has gained increasing attention as it provides an important cue to help us to understand why people seek information, what type of information they need, and how they are going to use the information [Byström & Hansen, 2005; Taylor, 1991]. Our definition of a task is mainly based on the research of [Vakkari, 2003; Byström & Hansen, 2005; Vakkari, 2003]. We regard *a task to be a specific piece of work, in which a person or a group of persons undertakes a series of actions in a situation*. Defining a task as a piece of work indicates that it has a performer, a meaningful purpose, and an undertaken context [Hackman, 1969; Byström & Hansen, 2005]. This is a definition, which emphasizes the conceptualizing of tasks more from the point of view of the actors and the social context of the task performance [Checkland & Holwell, 1998, Suchman, 1995].

We define *the organizations or organizational units as actors*. Each actor has a list of professional roles. A role is defined in [Merriam-Webster] as "a function or part performed especially in a particular operation or process". Obviously, role is a function relevant concept. *The professional roles of an actor are defined in terms of functions an actor must provide.* Therefore, actors are exclusive, and based on the functions they provide, i.e. their professional roles. A task is performed when an actor adopts one of its professional roles. A task can be composed of smaller tasks. At the level of a simple task, it is constituted of a series of interconnected actions. A task is undertaken in a context, i.e. the situation, where an actor is required to adopt one of its professional roles. According to the definition of task in our research, tasks are required to be identified from an actor's professional roles, i.e. from the functions an actor can provide. However situation is a dynamic concept, as perceived by the SA process defined in section 2.3.2. It s not feasible to define all tasks corresponding to a specific situation. Instead, the tasks can be defined in facts, which are more concrete and tangible. Since a task can be composed of smaller tasks, required tasks in a detected situation can be composed of sub-tasks identified using the relevant facts. Tasks are undertaken in a process to formulate the solution for an existing fact.

Based on the concepts defined in the previous subsections, we formulate a meta-model of personalized information seeking and retrieval on the basis of service-oriented system design principle. The meta-model is shown in Appendix 1.

3. CASE STUDY

We tested the conceptual design by applying it in a real life case. The case study presented in this section was taken from an ongoing EU research project. The harbor described in this case study is one of the largest and most important harbor infrastructures in the world. Global collaborations for natural disasters recovery and anti-terrorism cooperation have encouraged management at this harbor to take put in global information networking, designed to share information with other crisis experts and professionals. The harbor management intends to build a more flexible and extendable information platform to provide all authorized and involved actors with access to their role-based personalized, situation-aware information to facilitate the performance of their tasks during a crisis response.

Since the development of a complete system is not yet feasible, due to the difficulty of building trust between the various crisis relief/response organizations, and getting them to share their information, we built an early prototype demonstration to show that it would be possible to build such a service-oriented platform to provide personalized, situation-aware information seeking and retrieval services for crisis response situation.

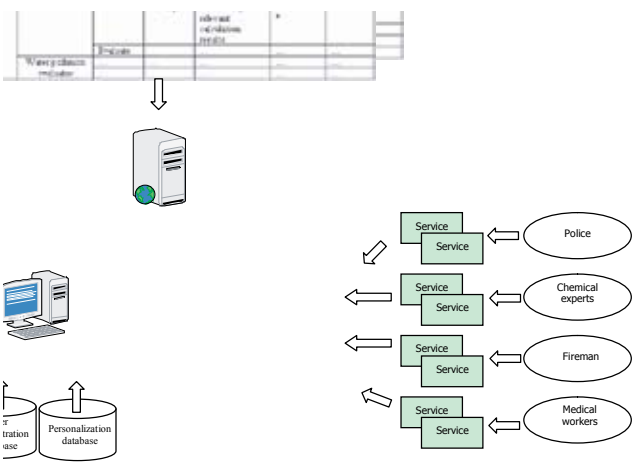
We utilized three computers in our prototype implementation shown in Figure 5, representing the service consumer, service provider, and service broker in a SOA respectively. The information provided by the actors involved was encapsulated as information services. We believe there are four main actors in crisis response: police, medical experts, fireman, and chemical experts. We implemented several information services for these 4 actors as jini services in our prototype. These jini services were required to be registered on a jini lookup server. The requested registration information is shown in the table in Figure 5. We used a simple example from a chemical expert to show what, and how, the required information on information services is registered in the jini lookup server. In the client PC, we use Liferay 4.0[1] as the portal software, and embedded Tomcat 5.0[2] as the web server to build the 'crisis response and management portal'. We built two databases, a user administration database and a personalization database, which were used to support personalized information seeking and retrieval applications running on this portal. User's role-based profiles, stored in the user administration database, were used to control their information access. The personalization database was built based on the meta-model presented in Appendix 1. Previous existing crisis situations, their constituting scenarios, scenarios' constituting facts, facts' solutions, etc, were stored in the tables of situations, scenarios, facts, solutions and tasks as the historical information in the personalization database. The personalization database was implemented in MySQL[3].

Tasks are keys that are used to search for the required information services provided by the different organizations. When a task is selected, it generates a service search template, which is constituted of three attributes, actor, role and task. A selected task generates this service search template by filling in the actor name, role name and task name. The service search template will then be sent to the jini lookup server to look for the relevant information services. All the information services of a selected task are returned to the user. Information service name, information service description, service status and location are also returned to the user. It is the users' work to figure out their required information services based on the returned service information.

4. CONCLUSION & FUTURE WORK

In this paper, we presented a new way of thinking of retrieving personalized information based on service-oriented design principle. Our situation model is capable of reflecting and inferring the unpredictable and dynamic situation users are facing. Our task model is capable of personalizing users' information needs in a dynamic situation according to their professional role(s). The combination of the situation model and task model allows users' role-based personalized, situation-aware information needs to be sufficiently inferred, and well structured in a meaningful way. Simultaneously, applying service-oriented design principle in our conceptual design allows us to realize independent service implementation and service modeling, and quickly to configure information acquisition applications

Figure 5. Prototype architecture



to satisfy users' dynamic information needs by choosing the required services. We believe that our conceptual design provides a possible solution to building a bridge between high-level functional requirements and low-level technology availability.

The next step of our research is to focus on how to reduce the time to action in a time critical situation using an effective information coordination service for situation aware process orchestration [González-Rivera, 2006].

REFERENCES

[Byström & Hansen, 2005] *Conceptual Framework for Tasks in Information Studies*, Katriina Byström, Swedish School of Library and Information Science, Göteborg University and Högskolan I Borås, Preben Hansen, Swedish Institute of Computer Science (SICS), JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 56(10):1050–1061, 2005, http://www3.interscience.wiley.com/cgibin/fulltext/110497382/PDFSTART

[Checkland & Holwell, 1998] *Information, Systems and information systems: Making sense of the field. Checkland, P. & Holwell S., Publisher: John Wiley & Sons; 1 edition (November 27, 1997), ISBN: 0471958204*

[Dictionary] http://dictionary.reference.com/search?q=environment&x=42&y=13

[Douglas, 2003] *Web Services and Service-Oriented Architectures: The Savvy Manager's Guide (The Savvy Manager's Guides)***,** Douglas K. Barry, Publisher: Morgan Kaufmann (April 1, 2003) ISBN: 1558609067

[Endsley, Bolte & Jones 2003] *Designing for situation awareness: an approach to user-centered design***,** Endsley, Mica, R., Bolte, B., and Jones, Debra G., British Library Cataloging in Publication Data, ISBN: 0-7484-0966-1 (hbk) ISBN:07484-0967-X(pbk)

[Gaslikova, 1998] Gaslikova, Irina *"Information Seeking in Context*" and the development of information systems." Information Research, 5(1) Available at: http://informationr.net/ir/5-1/paper67.html

[González-Rivera, 2006] Information Coordination Service for Situation Aware Process Orchestration: improving time to action in emergency response, Rafael González-Rivera, draft proposal, Delft University of Technology, the Netherlands, Feb, 2006

[Hackman, 1969] Hackman, J.R. (1969). T*owards understanding the role of tasks in behavioral research*. Acta Psychologica, 31, 97–128.

[Merriam-Webster] http://www.m-w.com/dictionary/environment

[Mizzaro & Tasso, 2002] Stefano Mizzaro and Carlo Tasso, Personalization Techniques in Electronic Publishing on the Web: Trends and Perspectives, Universidad de Málaga, Departamento de Lenguajes y Ciencias de la Computación, Proceedings of the AH'2002 Workshop on Personalization Techniques in Electronic Publishing Málaga, Spain, May 2002 Selected papers, http://www.dimi.uniud.it/mizzaro/AH2002/proceedings/PerElPub.pdf#search=%22user-controlled%20information%20personalization%20to%20autonomous%20system-controlled%20adaptation%22

[Niedźwiedzka, 2003] *"A proposed general model of information behavior*" Niedźwiedzka, B. (2003) *Information Research*, **9**(1) paper 164 http://InformationR.net/ir/9-1/paper164.html

[Stojanovic, Dahanayake & Sol, 2004] *An Approach to Component Based and Service-Oriented System Architecture Design*", Stojanovic, Z., and Dahanayake, A. and , Sol, H.G., (2004):Faculty of Technology, Policy and Management, Delft University of Technology Jaffalaan 5, 2628. BX Delft, The Netherlands, {Z.Stojanovic, A.Dahanayake, H.G.Sol}@tbm.tudelft.nl

[Vakkari, 2003] *Task-Based Information Searching*., Vakkari, Pertti, Annual Review of Information Science and Technology (ARIST), v37 p413-64 2003,

[Wilson, 1994] Prof. T.D. Wilson, INFORMATION NEEDS AND USES: FIFTY YEARS OF PROGRESS? Department of Information Studies University of Sheffield, In Vickery, B.C. (Ed.), *Fifty years of information progress: a* Journal of Documentation *Review*, (pp. 15- 51) London: Aslib, 1994.

[Wilson, 1999] *"Exploring models of information behaviour: the 'uncertainty' project,* Wilson, T.D, Information Processing and Management, Volume 35, Number 6, November 1999, pp. 839-849(11)

ENDNOTE

1 Information about liferay can be found from http://www.liferay.com/web/guest/home

2 Information about Tomcat can be found from http://tomcat.apache.org/

3 Information about MySQL can be found from www.mysql.com

APPENDIX 1