

Collecting and assessing collaborative requirements

Preben Hansen
Swedish Institute of Computer Science
Box 1263
SE-164 29 Kista
+46 (0)70 488 76 19
preben@sics.se

Guido Lorenzo Granato
Dipartimento di Informatica e
Sistemistica
University of Rome "La Sapienza"
Via Ariosto 25 - 00185 Rome - Italy
guido.granato@gmail.com

Giuseppe Santucci
Dipartimento di Informatica e
Sistemistica
University of Rome "La Sapienza"
Via Ariosto 25 - 00185 Rome - Italy
santucci@dis.uniroma1.it

ABSTRACT

Within evaluation frameworks such as the CLEF evaluation platform, many people are involved in different evaluation tasks such as organizing the task, creating topics, managing document collections, and handling different participants and submission. It also includes the selection of measures and run the final evaluations. In some of these stages of the evaluation process, there may be a need for collaborative information handling procedures. In order to identify some of the needs and requirements for collaboration, we investigated three different domains. This paper presents identified needs for collaborative activities that may occur within an interactive evaluation platform such as in the European Promise (Participative Research labOratory for Multimedia and Multilingual Information Systems) project. Following a user centered design, collected requirements have been assessed using a realistic mock-up scenario.

General Terms

Documentation, Design, Experimentation

Keywords

Collaboration, Information Seeking, User study, Information access. Requirement elicitation, Evaluation

1. INTRODUCTION

As Foster [1] correctly points out, research related to collaborative information seeking and retrieval is an interdisciplinary phenomenon including studies especially from areas such as human-computer interaction (HCI), computer-supported cooperative work (CSCW) and information science. Thus, definitions of collaborative information seeking are developed from the disciplines and circumstances they have been used in. Recent research in information seeking and retrieval (IS&R) extends our knowledge on how people access, retrieve and judge information.

O'Day [2] described four levels of sharing information in collaborative group situations: a) sharing results with other team members; b) self-initiated broadcasting of interesting information;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference '10, Month 1–2, 2010, City, State, Country.
Copyright 2010 ACM 1-58113-000-0/00/0010...\$10.00.

c) handling search requests made by others; and d) archiving potentially useful information into group repositories for others to use.

Karamuftuoglu [3] discusses what he calls *social informatics* and Fidel and colleagues [4] describe a project focusing on collaborative activities of members of a work-team within an organization performing IS&R tasks.

Hansen and Järvelin [5] investigated collaborative IS&R processes performed by patent engineers.

Foster, J. [1] presents a literature review describing current research of collaboration related to seeking, searching and retrieval tasks. The information-seeking task involves both social and collaborative approaches, while the information retrieval task involves collaborative issues such as collaborative filtering and collaborative querying. Foster concludes that research in the field of CIR needs to address and evaluate the conditions that influence development of systems that handle collaborative information activities, such as "...direct and indirect collaboration during information tasks...".

Reddy & Spence [6] conducted an ethnographic field study of a multidisciplinary patient care team in an emergency department. The goal was to identify information needs within a medical team and to identify situations that trigger collaborative information seeking activities.

Empirical studies on collaboration in IR among end-users have been scarce indeed, but one example is SearchTogether system by Morris and Horwitz [7], a prototype that allows a group of users to remotely collaborate when searching information on the Internet.

2. METHODOLOGY

We used the evaluation framework of CLEF¹ evaluation campaign. In these evaluation tasks, a basic set of *participants* are involved:

- organizers,
- experiment participants,
- relevance assessors,
- topic creators,
- site administrators.

Each of these participants is described along a set of activities, or *tasks*, that are performed by these different actors. These are further broken down in more detailed activities, or sub-tasks. In order to collect requirements related to collaborative information

¹ Cross-Language Evaluation Forum (<http://www.clef-campaign.org/>)

handling activities, we used evaluation tasks from the following domains: the patent domain, the radiology domain, and the Cultural heritage domain

Through questionnaires and follow-up communications with key persons of the main organizers of CLEF tracks representing the three different domains, we asked the organizers for each lab-based evaluation task to describe the task in detail and then point out stages or parts of the process that may benefit/need interactive collaborative information handling.

The data from the questionnaires were analyzed and classified. A list of steps was identified for each group of participants describing the task process and a set of different needs for collaboration were extracted and described.

The aspects of collaboration as suggested will be described after each actor category. Of course, the summarization of the requirements might be done using other parameters, such as types of collaborative tools or types of activities. After each category of participant, a summary of identified and requested support for collaborative information handling is presented. The summary has two levels: the identified *need* and the suggested *support for collaboration*. Below an example will be described.

3. CATEGORIES OF PARTICIPANTS AND TASKS

In this section we will present one of the categories with related real life evaluation tasks identified through the CLEF evaluation environment. These are followed by a set of identified needs and proposals for tools that support collaborative activities. Section 3.1 describes the main tasks performed. Each main task also has some sub-tasks.

3.1 Organizer

- **Preparation of DATA**
 - Creating copyright forms
 - Data on which tasks are available in a specific year
 - Create metadata forms to collect metadata
- **Add DATA**
 - information about the specific track
 - on tasks and collections
- **View COLLECTION data**
 - collection statistics,
 - documents,
 - metadata,
 - language distributions,
 - citations,
 - classifications,
 - annotations
 - images
 - cases description
 - data usage

- **Create topics/information needs**
 - Access to collection
 - Create information need descriptions and topics
 - Need to have access on user data through survey etc
- **View USER data**
 - login/password,
 - signed copyright forms
 - registrations, submissions
 - data supplied
- **Handle submissions:**
 - define accepted formats
 - check submissions
 - format, (automatically if possible)
 - completeness
 - content
 - prepare for evaluation
- **Run evaluations:**
 - choose measures (usually `trec_eval` with a number of potential measures)
 - define evaluation chains,
 - choose qrels
 - run files,
 - significance tests.
- **Gather and combine the results**
 - visualization of the track results (centralized)
- **Upload results**
 - make results available to participants in the form of the qrels and the lists with the ranked results of participants.

Identified Need #1:

There is an extensive collaboration between the CLEF track organizers during the complete evaluation campaign. On the organizers level, the following has been pointed out as important requirements for collaborative activities:

- The procedure and standards of *preparation of data* normally only includes the organizers, however it is communicated to other track organisers
- During the actual track performance the organizers initiate collaboration with
 - participants,
 - assessors,
 - topic creators and
 - visitors (other researchers)

in order to *discuss* and *prepare performance measures* and *statistical analysis* and *scientific paper publication*.

Support for collaboration:

This call for introducing a common and collaborative discussion area/platform in which discussions could take place and

documents could be shared and where all user groups can collaborate with the organizers or if needed with other users. This platform should be able to allow support for collaborative information handling tasks across both a horizontal as well as a vertical level:

- allow the organizers to collaborate with the other users
 - within a specific track as well as
 - between other organizers in other tracks
 - Tools that allow organisers to define and prepare relevant
 - measures for a task and to be able to communicate this to other actors such as participants, assessors, topic creators and visitors
 - Tools that enable statistical analysis of different data in a collaborative environment. For example, single participants could compare their analysis with other participants.

Identified Need #2:

The organizer would like to be able to view results in different ways and at different levels. It should support collaborative information handling actions for the different people involved an evaluation platform such as: organizers, content providers, track participants, and system builders.

Support for collaboration:

An electronic system should keep the data consistent and limit the amount of manual work and emails necessary.

- Introduce a common working area that allows the organizer to *view*, in a centralized view, the results of the evaluation. This may involve specific results or *comparing results* from several participants. Tools that *adapt* to the type of data needed to be designed.
- Furthermore, different levels of *statistics* could be shared and collaborative tools could then allow for *cooperative manipulation of these statistics*. For example: *data characteristics* (based on metadata), *data usage* and experiments (e.g. number per lab, per task, etc). Different views of these statistics could be displayed.

Identified Need #3:

Also at the participant level, there may be a requirement of having a communication tool (interface) between groups of participants and experts on the organizational level. Amongst others the participants wanted more assistance for newcomers.

Collaborative support:

A Q&A tool may be used to solve this requirement.

Identified Need #4:

It would be really helpful to be able to making different components of a retrieval system available between researchers and participants. Such components can be combined in different ways and everyone can concentrate on their main are.

Support for collaboration:

A tool that a) handle communication for this exchange or discussion on sharing components of retrieval systems or b) a way of pointing to components that are free for use or a common

area/storage area in which these components could be offered to others.

4. DISCUSSION

This exploratory study descried a way of capturing requirements for collaborative information situations in an IR evaluation setting. The study describes the different participants involved and the tasks and sub-tasks performed for the different participants. More specifically, we reported on the participant category of organizer. 4 different needs for collaborative support was identified and how these needs could be supported was proposed.

Within the Promise project, and according to the user feedback, several collaboration tools have been selected, to accomplish the requirements described on Section 3 (see Figure 1).

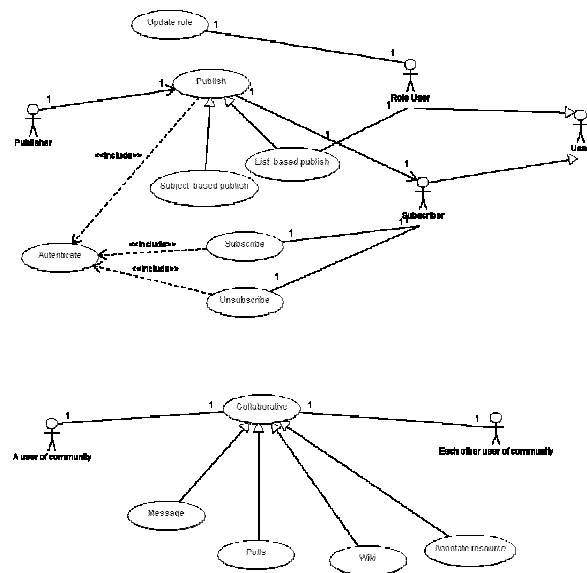


Figure 1: UML use cases for the PROMISE collaborative requirements. Note that the actor Role User encompasses the roles described on Section 2.

In particular, we foresee:

- an annotation system;
- a set of basic interaction tools, like news, wiki, messages, polls, and chat;
- a publish/subscribe mechanism strictly linked with the annotation system and the above interaction tools. This architectural pattern involves two kinds of users, publishers and subscribers, and allows subscribers for expressing their interest for particular events, in order to receive relevant notifications. Moreover, in Promise users are automatically added to one or more lists by the system, according to their role(s) (see Section 2). When a user creates or modifies a system object, or start a collaborative activity (e.g., annotates a PROMISE artifact) all the users that are in the relevant lists will receive a notification about the event.

The feedback has been collected using mock-up, showing in details all the phases of the user interaction with the system. See Figure 2.

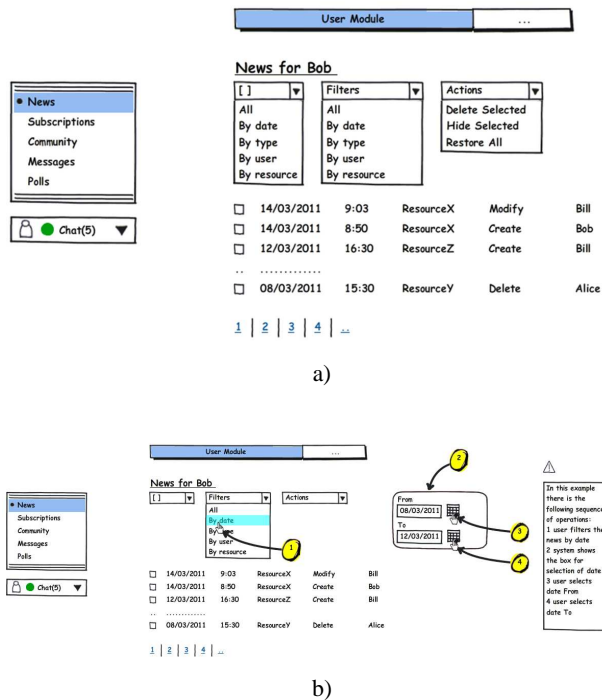


Figure 2. Assessing the collected user requirements with a mock-up. Figure 2.a shows the News page, while Figure 2.b shows the associated filtering mechanism.

We can conclude that it is feasible to extract requirements from real life situations and propose support for interactive collaborative information handling.

Moreover we can report a very positive feedback about the mock-up usage: people were engaged in the discussion and reported that the high number of details showed in the mock-up increase the feeling of realism and allows for a better understanding of the foreseen system functionalities.

5. ACKNOWLEDGMENTS

The work reported in this paper has been supported by the PROMISE network of excellence (contract n. 258191) project as a part of the 7th Framework Program of the European commission (FP7/2007-2013).

6. REFERENCES

- [1] Foster, J. (2006) Collaborative information seeking and retrieval. In: Annual Review of Information Science and Technology, Volume 40, 2006, 329–356.
- [2] O'Day, V. and Jeffries, R. (1993) Orienteering in an information landscape: how information seekers get from here to there. Proceedings of the conference on Human factors in computing systems. INTERCHI '93, (pp. 438-445), Amsterdam, The Netherlands, New York: ACM Press.
- [3] Karamuftuoglu, M. (1998) Collaborative Information Retrieval: Towards a Social Informatics View of IR Interaction. JASIS, 49(12), 1070-1080, 1998.
- [4] Fidel, R., Bruce, H., Pejtersen, A. M., Dumais, S., Grudin, J., and Poltrick, S. (2000). Collaborative Information Retrieval (CIR). The New Review of Information Behaviour Research, 2002, pp. 235-247.
- [5] Hansen, P. and Järvelin, K. (2005) Collaborative information retrieval in an information intensive domain. Information Processing and Management (IPM), 41 (5), pp. 1101-1119.
- [6] Reddy, M.C., and Spence, P.R. 2008] Collaborative information seeking: A field study of a multidisciplinary patient care team. Information Processing & Management 22(1), 242-255.
- [7] Morris, M.R. and Horvitz, E. 2007] SearchTogether: An Interface for Collaborative Web Search. In: Shen, C., Jacob, R. and Balakrishnan, R. (Eds.), Proceedings of the 20th annual ACM symposium on User interface software and technology, Newport, Rhode Island, USA, October 7 - 10, 2007, 3 – 12.