INFORMATION RETRIEVAL IN CONTEXT

SIGIR 2004 IRiX WORKSHOP 29th JULY 2004, SHEFFIELD, UK

ACM SIGIR 2004 Workshop on "Information Retrieval in Context"

Organised by

Peter Ingwersen, Royal School of LIS, Denmark
Keith van Rijsbergen, University of Glasgow, Scotland
Nick Belkin, Rutgers University, USA

Programme editor

Birger Larsen, Royal School of LIS, Denmark

Web Site

Ryen White, University of Glasgow, Scotland

http://ir.dcs.gla.ac.uk/context

SIGIR 2004 Sheffield July 25-29

27th Annual International ACM SIGIR Conference

Preface

There is a growing realisation that relevant information will be increasingly accessible across media and genres, across languages and across modalities. The retrieval of such information will depend on time, place, history of interaction, task in hand, and a range of other factors that are not given explicitly but are implicit in the interaction and ambient environment, namely the context. IR research is now conducted in multi-media, multi-lingual, and multi-modal environments, but largely out of context. However, such contextual data can be used effectively to constrain retrieval of information thereby reducing the complexity of the retrieval process. To achieve this, context models for different modalities will need to be developed so that they can be deployed effectively to enhance retrieval performance. Thus truly context-aware and -dependent retrieval will become feasible.

Context implies interactive IR and there may exist a stratification of contexts in association to IR engines and systems. For example, knowing where a user is focusing his or her attention during image retrieval can enhance the operation of relevance feedback to the system. The user's current task situation also acts as context as does his or her current information seeking situation of which IIR forms part. The underlying hypothesis (and belief) is that by taking account of context the next generation of retrieval engines dependent on models of context can be created, designed and developed delivering performance exceeding that of out-of-context engines.

This purpose of this workshop is to explore a variety of theoretical frameworks, characteristics and research approaches to focus on an agenda of activities to be recommended for future interactive IR (IIR) research.

Organisers

Peter Ingwersen, Royal School of LIS, Copenhagen, Denmark
Keith van Rijsbergen, University of Glasgow, Scotland
Nick Belkin, Rutgers University, USA

Program committee

Nick Belkin, Rutgers University, USA
Pia Borlund, Royal School of LIS, Aalborg, Denmark
Yves Chiaramella, IMAG, Grenoble, France
Peter Ingwersen, Royal School of LIS, Copenhagen, Denmark
Birger Larsen, Royal School of LIS, Copenhagen, Denmark
Andrew MacFarlane, City University, London, UK
Ian Ruthven, University of Strathclyde, Scotland
Eero Sormunen, University of Tampere, Finland
Amanda Spink, University of Pittsburgh, USA
Pertti Vakkari, University of Tampere, Finland
Keith van Rijsbergen, University of Glasgow, Scotland

Programme editor

Birger Larsen, Royal School of LIS, Copenhagen, Denmark

Web site

Ryen White, University of Glasgow, Scotland
http://ir.dcs.gla.ac.uk/context

Programme

09.00-09.15: **Welcome by Peter Ingwersen**

09.15-10.15: **Session I: Opening Panel**
IR in Context - the Conceptions of Context.
Peter Ingwersen, Royal School of LIS, Denmark; Ian Ruthven, University of Strathclyde, Scotland; Kalervo Järvelin, University of Tampere, Finland

10.30-12.15: **Session II: Models and General Approaches to IRiX**
Session Chair: Peter Ingwersen

Information Retrieval as Multi-tasking: an Exploratory Framework.
A. Spink and M. Park, School of Information Sciences, University of Pittsburgh, USA.

The Role of Context in Information Retrieval.
G.J.F. Jones, School of Computing & Centre for Digital Video Processing, DCU, Ireland, and P.J. Brown, CS Dept, University of Exeter, UK.

Using User's Context for IR Personalization.
N.J. Belkin, G. Muresan, X.-M. Zhang, SCILS, Rutgers University, USA.

Identifying the Significant Contextual Factors of Search.
E.G. Toms, and Christine Dufour, Dalhousie University, Canada, J. Bartlett, L. Freund, and S. Szigeti, University of Toronto, Canada.

Discussion (25 min).

12.15-13.00: **Lunch**

13.00-15.00: **Session III: IRiX Applications**
Session Chair: Ian Ruthven

A Context-sensitive Information System for Mobile Users (Demo).
A. Göker, M. Yakici, R. Bierig, School of Computing, The Robert Gordon University, Scotland, and H.I. Myrhaug, Trondheim, Norway.

Context in Active Groups (Demo).
S. Watt, School of Computing, The Robert Gordon University, Scotland.

Building a Test Collection for Investigating Contextual Information Retrieval.
D. Kelly, University of North Carolina, USA.

End Users in the Context of XML documents: Setting Up an Interactive Track at INEX.
B. Larsen, Royal School of LIS, Denmark, A. Tombros, Dept CS, Queen Mary, University of London, UK, and S. Malik, University of Duisburg-Essen, Germany

Discussion (20 min).

15:00-15.15: **Coffee break**

15.15-16.30: **Session IV: Mixing Contextual Evidence and Representations**
Session Chair: Kalervo Järvelin

Intelligent RSS News Aggregation Based on Semantic Contexts.
W. Huang and D. Webster, Centre for Internet Computing, University of Hull, UK.

On the Need for Annotation-based Image Retrieval.
M. Inoue, National Institute of Informatics, Japan.

Testing the Principle of Polyrepresentation.
M. Skov, H. Pedersen, B. Larsen, and P. Ingwersen, Royal School of LIS, Denmark

Discussion (15 min).

16.30-17.00: **Wrapping up and where do we go??**
Nick Belkin, USA & Peter Ingwersen, Denmark.

Table of Contents

Information Retrieval in Contexts

Peter Ingwersen[1] and Kalervo Järvelin[2]
[1]Department of Information Studies, Royal School of Library and Information Science – DENMARK
[2]Department of Information Studies, Tampere University – FINLAND
Email: pi@db.dk – kalervo.jarvelin@uta.fi

Categories and subject descriptors: H.3.3. Information Search and Retrieval

General Terms: IR Theory.

Keywords: Information Retrieval, IR in Context; Context Types; Work Task-based IR

1. INTRODUCTION

The integration of perspectives and models of information seeking and information retrieval (IS&R) into a holistic conceptual framework for research is currently under development (Ingwersen & Järvelin, in preparation). Epistemologically it is founded on the cognitive viewpoint (Belkin, 1990) and based on elements of the cognitive theory for interactive IR (IIR) put forward by Ingwersen (1992; 1996; 2001). Intentionality in the form of perceived work and search tasks or non-job related interests is central as the rationale underlying IS&R (see, e.g., Järvelin (1986)). Search tasks are the instrumental activities, cognitive-emotional as well as physical, that in IS&R serve to advance the fulfillment of the work task in terms of information provision.

The framework reflects the understanding that IS&R is a process of cognition for the information seeking actor(s) or team *in context*. Algorithmic and IIR, as well as information seeking (IS), involve cognitive and emotional representations from a variety of participating actors. Such representations are seen as manifestations of human cognition, reflection, emotion or ideas forming part of IS&R components and kinds of interaction in context — as shown in Fig. 1.

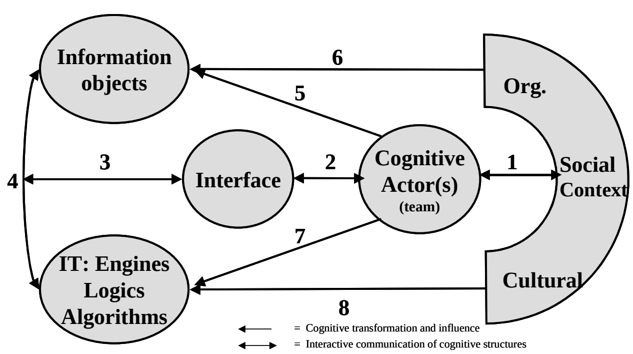


Fig. 1. General model of cognitive information seeking and retrieval, Ingwersen & Järvelin (in preparation). Arrow numbers in ( ) refer to kinds of interaction or influence.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

The framework operates with several kinds of contexts. First, algorithmic and IIR processes cannot stand alone, but are *nested* in IS behavior as special cases of information behavior (Wilson, 1999) – Fig. 2. Algorithmic IR, i.e., the study of the interaction between information objects and IT-based algorithms, arrow (4) Fig. 1, has no real meaning without human information interaction with IR systems, arrows (2-3). IIR itself functions in the context of IS – but reversely IS becomes increasingly *also* only meaningful when considering the involvement of formal (algorithmic) IR engines and information structures. This is because progressively more informal communication channels, like mail, become formalized due to the overpowering and integrative effects of modern IT.

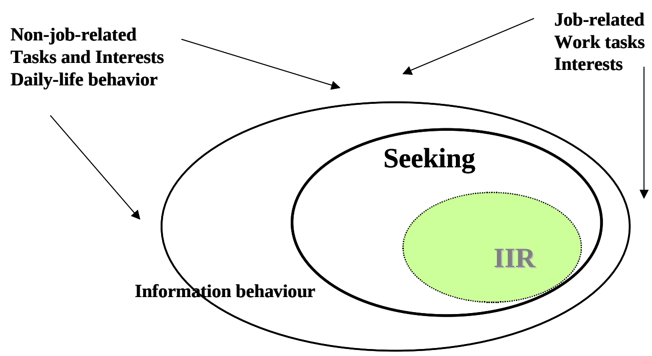


Fig. 2. Information retrieval embedded in IIR, information seeking and behavior and instigated by work tasks or interests from daily-life or (non)-job related behavior, Ingwersen and Järvelin (in preparation).

As information behavior is regarded, for instance, generation, acquisition, use and communication of information – as well as information seeking. Typical information seeking behavior is acquisition of information from knowledge sources, for instance, from a colleague, through (in)formal channels like social interaction in context (arrow 1), or via an IR system (arrows 2-4), Fig. 1; IIR involves information acquisition via formal channels like the Internet, or from other organized sources. In a cognitive sense, information behavior, information seeking and all forms of IR are activities driven by work tasks or interests – Fig. 2. We extend the notion of work task also to cover non-job-related emotional and cultural *interests*, e.g., entertainment, as depicted in Fig. 2.

Secondly, every information actor (or team of actors) operates in, and is influenced by, a dual contextual frame: that of the IT and information spaces surrounding the actor(s) — the systemic context on the left hand side, Fig. 1 — and the socio-cultural-organizational context to the right. By manifestations of practice and authorships the latter context influences over time the information space on the one hand (arrow 6) and the IT infrastructure (arrow 8) on the other.

Below, we will discuss in more detail the types of context that comes into play during information retrieval.

2. TYPES OF CONTEXT IN IR

As already emphasized by Wilson (1981), the current *situation in context* plays a central role in information seeking. Our cognitive framework extends this understanding by *also* stressing the role of the historic context – both contexts driving the shape of the current situation of the information seeker. In a narrow sense only the five central components of the framework, Fig. 1, and their interaction condition the shape of the situation at hand; but also the societal contexts in a wide sense – like *economic and techno-cultural infrastructure*, (5) Fig. 3 – influence the current situation as perceived by the information seeker or communicator[1]. Fundamentally, the situation is a personal cognitive construct in contexts.

The seeking actor's perception and interpretation of many levels and dimensions of contexts is central to how that situation develops. In this process the perceived work task or interest is an important interpretation outlining – but not totally defining – that situation at hand. Also included in the construct are the perceptions and interpretations of:

- Knowledge gap or ASK and relevance;
- Uncertainty and other emotional states;
- The potential sources for the solution (if any) of the work task or interest;
- The intentionality, i.e., goals, purposes, motivation, etc.;
- Information preferences, strategies, pressures (costs, time)
- Self, i.e., of own capabilities, health, experiences – and
- Systemic and interactive features and information objects.

According to the framework, the current context of a component is constituted by the other components immediately surrounding that component – Fig. 1. Hence, the study of one component (or an element), say interface functionality, should incorporate an awareness or direct involvement of the IT and information object components *as well as* of the searchers using the interface in some conceivable context. To the information seeker the real work task and preferences, influenced by the socio-organizational environment, forms a context to the situation at hand, as do the interface characteristics, information object presentation and the present mode of interaction. The searcher's experiences act as a *historic context* – (6) Fig. 3.

[1] A typical example of this (economic) influence - also directly on empirical research settings - was the high *cost* of public online searching in the 1970-80s. This lead to many investigations of the so-called "pre-search interview" – a phenomenon not applicable in free-of-charge in-house online systems or realistic to systems design.

Context is not only a searcher phenomenon. The system itself can be context-aware in use. Interacting with searchers, Fig. 1, means more to the system than capturing simple input data. Rather, temporal searcher interaction with a system forms a rich network of potential information regarding preferences, style, experience and knowledge as well as interests. This information helps to constitute a *session context* (arrow 2, Fig. 1, (3), Fig.3) that can be made available for the system to interpret current searcher actions (Ruthven et al., 2003). In principle the IIR session context should be seen in the broader perspective of information seeking behavior (thus including arrow 1) and nested in information behavior as a special case – as stated above and shown in Fig. 1.

From the system's point of view ergonomic behavior, like mouse or eye movements, patterns of relevance feedback or *evidence* of the immediate perceptions and interpretations by the searcher constitute this session context, with the seeking actors and *their* current situations in context (arrow 1) as more remote contextual phenomena – (4), Fig. 3. The latter may be manifestations of cultural conventions, organizational preferences, or domain-specific traditions. For IR systems design it is crucial to uncover patterns of objective, *tangible evidence* of actors' interpretations as well as of the socio-organizational or cultural context. Without it the system cannot react properly during session time.

The application of evidence algorithmically from the session context involves intensive knowledge of what such evidence implies. For instance, Spink et al. (1998) found that searchers often assess a large proportion of the retrieved and viewed documents as partially relevant when in a state of uncertainty or they are unfocussed on their work task. Later, during session time, a bigger proportion of documents is commonly judged highly relevant. This evidence may thus inform the system about the cognitive state of the current searcher and signal from which documents to derive potential novel search keys for query modification. Also the positioning of documents on the ranked output lists, assessed more or less relevant, may be applicable information to the weighting algorithms of the system – or give raise to research assumptions to be tested empirically.

The system in turn has technical characteristics relating to how information is presented during interaction (arrows 2-3, Fig. 1), that influence the session context and the searchers' perceptions of the system competence and the quality of information sources. Consequently, models of context are representations of a shared process of interpretation and adaptation on both sides of the interface.

Within each component of the cognitive framework, at quite elementary levels, the representations of different cognitive nature form *intra-component* structures in IS&R – (2), Fig. 3. They form a group of elements, like information objects in the information space component – (1), Fig. 3, that themselves are contextual to one another. Parts of documents, like references or outlinks to other information objects as well as citations or inlinks, are seen as giving and taking context to the content of other objects. Within objects, for instance, images are contextual to a surrounding text or other structures attached to them, and *vice versa*, and paragraphs serve as context for their own sentences and words: signs are seen in context of sign structures that are media-dependent.

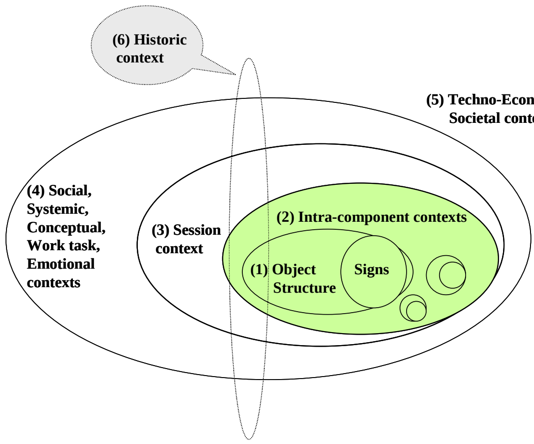


Figure 8.10. Nested model of context stratification for IS&R. From Ingwersen & Järvelin (in preparation).

Other components, like the IT platform or the interface, similarly contain inter-related elements and objects, such as pieces of software and algorithms. The seeking actor also constitute a component of interactive IR with his/her own inter-dependent (cognitive) structures, such as the central work task perception, problem state, uncertainty level and other emotional factors, information need, knowledge states, and search task comprehension. Such intra-component structures and manifestations actually constitute the evidence which, when tangible, is utilized in the principle of polyrepresentation (Ingwersen, 1992; 1994; 1996).

The historic context, (6) Fig. 3, functions across all other contexts at a given point in time and serves to produce expectations (models) concerned with the future steps in the IIR process and the surrounding components. However, present expectations relying on past experiences may indeed *not* always be satisfied by the conditions offered by the current context. For instance, the interface does not present documents in the expected form, the search algorithm seems incomprehensible, or the documents do not immediately satisfy the requirements as good as in previous IS&R situations. From the systems point of view similar disappointments may occur as to the searcher behavior and provision of information.

We summarize our *stratified* context definition for IS&R – Fig.3. In our framework five distinct but *nested levels* of context exist associated with each central component or actor: (1) The divergent representations of cognitive structures, often nested and always contextual to one another, like signs in context of sign structures constituting objects, embedded as (2) intra-component contexts in our framework; (3) the session context dealing with features of the interaction between two components or actors – with the situation at hand as a central cognitive-emotional element. Session context is embedded in broader seeking and information behavior. The situation at hand is constructed by the actor's perception of work and search tasks (interest), knowledge gap and potential sources, etc. in the (4) context of the conceptual, emotional, systemic and social properties immediately surrounding the actor or component. All actors, components and interactive sessions are influenced to a certain extent by (5) remote contextual constructs, such as general techno-economic infrastructures and socio-cultural factors in society. Across this stratification operates an additional dimension, that of the historic context of actors' experiences forming their expectations. All IIR processes and activities are under influence of this sixth form of context.

3. CONCLUDING REMARKS

Context, work and search tasks are interconnected concepts in IS&R. The various kinds of context are commonly nested, with the historic context in the form of experiences and learning of actors crossing the former. At the moment of IS&R the searcher is not only surrounded by a socio-organizational and cultural-epistemic context but *also* by session-based and systemic ones, which interrelate the IT and interface components as well as the information space. The situation at hand, dealing with work task or non-job related interest perception, is a cognitive-emotional construct thus heavily influenced by all those kinds of context. Similarly, the current search task execution is forced by the same contexts – but the historic context of the searcher, i.e., his/her earlier experiences with such contexts, plays the most central role in the situation. This can be seen as the main reason behind inter-searcher inconsistency, even though all contexts (except the historic one) including the work task or interest are the same across the searchers involved. Both search task implementation and relevance assessments are thus expected to be slightly different over several test persons or assessors (Vorhees, 1998).

IR in Context does not only deal with the contexts of searchers or searchers *as* context. IRiX also concerns the interaction between documents and IT platform in context of domains and different kinds of work tasks (and situations), i.e., an extension of the laboratory model, not necessarily involving test persons.

3. REFERENCES

[1] Belkin, N.J. (1990). The cognitive viewpoint in information science. *Journal of Information Science: Principles and Practice, 16*(1): 11-15.

[2] Ingwersen, P. (1992). *Information Retrieval Interaction.* London: Taylor Graham.

[3] Ingwersen, P. (1994): Polyrepresentation of information needs and semantic entities: elements of a cognitive theory for information retrieval interaction. In: *Proceedings of SIGIR 1994*, p. 101-110.

[4] Ingwersen, P. (1996): Cognitive perspectives of information-retrieval interaction - elements of a cognitive IR theory. *Journal of Documentation,* 52(1), p. 3-50.

[5] Ingwersen, P. (2001). Cognitive information retrieval. In: Williams, M. (ed.), *Annual Review of Information Science and Technology*, 34: 3-51.

[6] Ingwersen, P. & Järvelin, K. The Turn: Integration of Information Seeking and Retrieval in Context. Kluver. (In preparation).

[7] Järvelin, K. (1986). On information, information technology and the development of society: An information science perspective. In: Ingwersen, P., Kajberg, L. & Pejtersen, A.M. (eds.), *Information Technology and Information Use.* London, UK: Taylor Graham: 35-55.

[8] Ruthven, I., Lalmas, M. & van Rijsbergen, K. (2003). Incorporating user search behaviour into relevance feedback. *Journal of the American Society for Information Science and Technology,* 54(6): 529-549.

[9] Spink, A., Greisdorf, H. & Bateman, J. (1998). From highly relevant to not relevant: Examining different regions of relevance. *Information Processing & Management,* 34(5): 599-621.

[10] Voorhees, E. (1998). Variations in relevance judgments and the measurement of retrieval effectiveness. In: W.B. Croft, A. Moffat, C.J. van Rijsbergen, R. Wilkinson and J. Zobel (eds.), *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* New York: ACM Press: 315-323.

[11] Wilson, T.D. (1981). On user studies and information needs. *Journal of Documentation, 37*(1): 3-15.

“and this set of words represents the user’s context…”

Ian Ruthven
Department of Computer and Information Sciences
University of Strathclyde
Glasgow. G1 1XH
Ian.Ruthven@cis.strath.ac.uk

1. INTRODUCTION

A practical means of handling and exploiting contextual information has always been one of the long-term goals of Information Retrieval research. Indeed, our forebears in library science were always aware that the situations in which information needs arose were as important as the topic being searched. Qualitative work in the information seeking area has characteristically been rich in noting the importance of the contextual issues involved in searching such as the user’s task, the search environment and, of course, the searchers themselves. However the response to this agenda, in terms of context-intelligent search systems, has been limited.

Most IR systems have smoothed context out of the retrieval programme. Initially, this was a necessary simplification of the IR task to facilitate the creation of working search systems. Partly it also arises from a general trend in IR research to provide one-size-fits-all systems: systems that will work for most people most of the time [1]. The implications of this can be seen, for example, in user evaluations which are typically discrete, snapshot investigations of interactive systems with limited consideration of the context in which such systems would be used. Contextual aspects, such as user motivation and user knowledge, are seen as experimental variables to be controlled not as important evidence, integral to our understanding of how IR systems support searching.

Similarly, few IR systems have any memory of the user. The traditional model of a search engine supports individual *sessions* of querying or browsing but not long-term repeated use by an individual. In its most limited form context has usually been seen as a way of disambiguating search topics. User profiles, for example, attempt to model long-term topical interests but the profiles themselves usually lack context regarding their own creation: we have sets of important concepts but not the knowledge of why they are important.

There are significant attempts to incorporate *learning* within search systems, mostly utilising user modelling, e.g. [2], however these often model search artefacts not the broader search context. This is, of course a tricky objective as, no matter how often a user searches, the system only has access to small, possibly tangential, indications of context. Sifting out what evidence of context is available, and more importantly what evidence is significant, is one area where modern computing technology can help take a lead.

The presence of large-scale information resources, fast (and cheap) machines, and users who *want* to search, have provided the impetus to many new and stimulating areas of context-based IR research. Much of this research is focused around gathering contextual information through behaviour monitoring, environment detection, gesture recognition and perceptual evidence, such as eye-tracking. This begs the question of what we do with all this context information now that we have it? Work by Kelly [3], amongst others, has shown that we cannot simply examine small fragments of contextual evidence in isolation: we need larger frameworks for describing and exploiting computational issues of context.

This has several implications. From a viewpoint it means that developing systems which provide some *reasoning* about context is as important as how contextual information is gathered. It may also mean rethinking some of our core beliefs about IR techniques. What do notions such as algorithmic similarity or relevance mean within context-sensitive situations? Are terms the basic contextual unit? It also means that we need a meaningful dialogue between research communities that deal with context. Context itself is not only an IR issue, any research area involving people necessarily involves context [4].

Whether we work with context or ignore it, context is always with us. Broad knowledge resources such as ontologies, currently underpinning the semantic web, are popular for context-sensitive systems because they are seen as objective and context-free. Hence, they can be used for many applications requiring some semantic component. However, any information resource is not context-free in its creation: context often comes built-in and understanding how and why resources are constructed is important to understanding how to use them effectively. Context is complex and needs complex, multi-layered solutions. IR systems form an ideal test-bed for investigating the computational nature of these solutions.

2. REFERENCES

[1] K. Spark-Jones. IR lessons for AI. Proceedings of Searching for Information: Artificial Intelligence and Information Retrieval Approaches. IEEE Special Events. Glasgow. 1999.

[2] W. B. Croft and R. Thompson. I3R: a new approach to the design of document retrieval systems. Journal of the American Society for Information Science 38. pp 289-404. 1987.

[3] D. Kelly and N. J. Belkin. Reading time, scrolling and interaction: Exploring implicit sources of user preference for relevance feedback. In Proceedings of the 24th Annual International ACM Conference on Research and Development in Information Retrieval. pp 408-409. New Orleans. 2001.

[4] N. A. Bradley and M. D. Dunlop. Towards a multidisciplinary user-centric design framework for context-aware applications. Proceedings of the 1st UK-UbiNet Workshop. London. 2003.

Extending Information Seeking and Retrieval Research Toward Context

Kalervo Järvelin[1] and Peter Ingwersen[2]
[1]Department of Information Studies, Tampere University – FINLAND
[2]Department of Information Studies, Royal School of Library and Information Science – DENMARK
Email: kalervo.jarvelin@uta.fi - pi@db.dk

Categories and subject descriptors: H.3.3. Information Search and Retrieval;

General Terms: IR Theory.

Keywords: Information Retrieval, IR in Context; Context Types; Work Task-based IR

1. INTRODUCTION

Analyzing information retrieval and information seeking (IS&R) from a task perspective puts new requirements on research in IS&R - requirements which have not been taken into account to a sufficient degree. We propose nine broad classes of variables that interact in IS&R processes, here called dimensions:

1. The *work task dimension*: the work task[1], (social) organization of work, collaboration and the system environment.
2. The *search task*, i.e., seeking and retrieval tasks, as understood in the organization.
3. The *actor dimension*: the actor's declarative knowledge and procedural skills.
4. The *perceived work task dimension*: the actor's perception of the work task
5. The *perceived search task*, the actor's perception of the search task including *information need types* regarding the task and the task performance process; emotions.
6. The *document dimension*: document genres and collections in various languages and media, which may contain information relevant to the task as perceived by the actor.
7. The *algorithmic search engine dimension*: the representations of documents / information and information needs; tools and support for query formulation; matching methods.
8. The *algorithmic interface dimension*: tools for visualization and presentation.
9. The *access and interaction dimension*: strategies of information access, interaction between the actor and the interface (both in social and in system contexts).

[1] The notion 'work task' implies also non-job-related daily-life tasks and/or interests.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

Each of the dimensions is complex and contains multiple variables. It is obvious that IS&R is performed in very diverse work and leisure situations characterized by diverse values on the variables of the broad dimensions. Thus also IS&R becomes quite different. In many, if not in the most, situations actors performing their work tasks are ignorant about IS&R – professionally mediated information retrieval being a notable but no more so frequent exception to the contrary. Mostly the actors view IS&R instrumentally, not as a goal in itself, and want to get over with it fast. They want just to cope with the tools and practices supplying information usable for augmenting their deficient knowledge. Therefore, they may consider IS&R just a pain in the neck and use various tools for information access in uninformed and ineffective ways – from the tool designer's viewpoint.

With this perspective in mind we do not really know how well current IR systems serve their users in various situations. At least the systems have been evaluated in IR research only for some limited use scenarios, mostly excluding searchers in context with their work tasks. Neither provides current information seeking research much help in this regard. While the information seeking practices of various actor populations have been investigated, much remains still unexplored. Moreover, the majority of information seeking studies does not look at IR systems at all or not at the level of system features, interaction and support for query formulation and searching. This situation is illustrated in Figure 1.

The real issue in IR systems design and evaluation is not whether a proposed method or tool is able to improve recall / precision by an interesting percentage with statistical significance. The real issue is whether it helps the searcher better solving the seeking and retrieval tasks (faster, with less resources, with better result quality). This has to do with learning about the search task, formulation of the request, a variety of tactics. Quite different needs (types and formulations), with accordingly found information, may serve the work task. One source may indeed not provide all the information required. Recall and precision only become relevant *after* the need formulation. Systems for information access have a job to do *before* the actor commits on a formulation.

Section 2 discusses the nine broad dimensions presented above. Section 3 analyzes current IR research and Section 4 design and evaluation frameworks for IS&R. Section 5 views upon IR from a knowledge work augmentation viewpoint and Section 6 gives the conclusions. This article is based on the forthcoming book by the authors (Ingwersen & Järvelin, 2004).

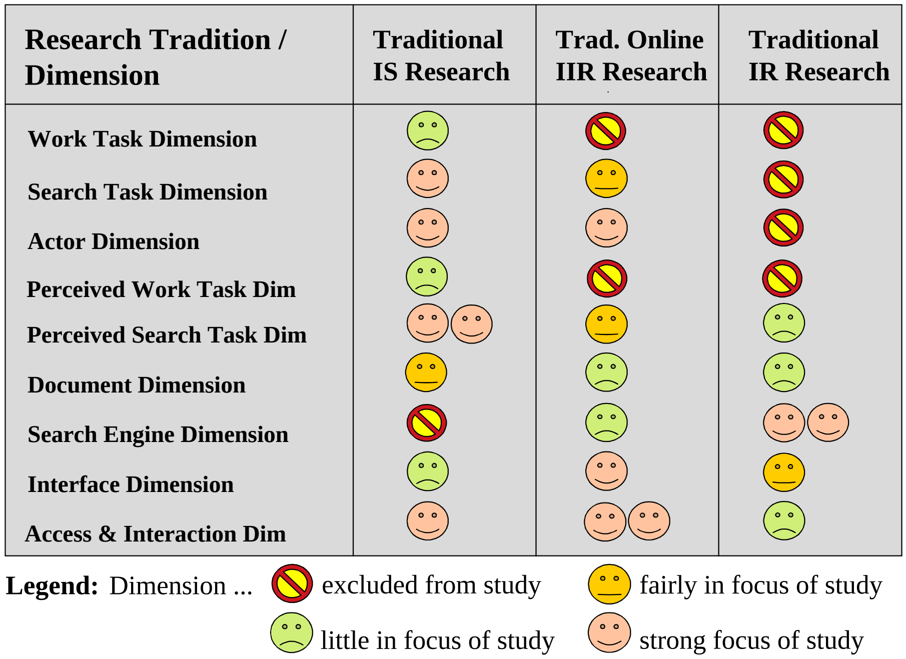


Figure 1. Foci of traditional IS&R research

2. DIMENSIONS AFFECTING IS&R

The Organizational Task Dimensions. This category contains two dimensions – the work task and search task dimensions. The latter covers both the seeking task and the retrieval task and the corresponding task processes. Likewise the work task subsumes the search task and process. The embedded ones serve the goals of the subsuming ones. Each work task may induce several search tasks and each search task several seeking and retrieval tasks, and the former direct the latter. They may run in parallel. The complexity of each task may vary and its process (or stages) may be more or less defined in its social / organizational environment. The social-organizational environments provide various systems and tools, as well as more or less articulated expectations regarding how each task should be carried out, often in collaboration with other actors.

The Actor Dimensions. The actor's perception and interpretation of the work task at each stage, with varying level of cooperation with other actors – the perceived work task dimension - greatly affects her search task and information needs – the perceived search task dimension – as do her prior knowledge, skills and experience, the third dimension. The actor's perception of the organizational and systemic environment, and her experience regarding them, together with the information needs, are the main factors in the formation of seeking tasks, the choice and use of systems and tools. The actor's perception and interpretation of various tasks are not independent – they have a history in the actor's entire career and the present organization. Also the pressures (e.g., hurry) and emotions affect her situation, perception and interpretation.

The Document Dimension. Various types of documents may be relevant for a given work task. The documents form different genres in different contexts of generation and use, e.g., orders, invoices, applications, plans and designs, guidelines and instructions, research reports, novels and poems, photos, films, musical records – to name just a few. From a task (interest) viewpoint, documents in such genres may (not) have been carefully selected and organized in collections with provided access tools, but may also lie unorganized in the actor's vicinity with her personal memory as the only access tool. Documents (genres) may come in many languages and representations – some of which being digital – and all can be exploited for IS&R.

The Algorithmic Dimensions. The two algorithmic dimensions deal 1) with the representations of documents / information and information needs, methods for matching these representations, tools and support for query formulation, and 2) tools for presentation via an interface. In addition to content, document representations may (not) cover explicitly their structure and layout. Likewise, information need representations may (not) cover explicitly their structure, content and motivation. A range of best match and exact match matching methods are available. The tools and support for query formulation may cover ontologies, thesauri, relevance feedback, and query modification. Access to documents / information may be through any combination of their metadata, full content, structure and layout. Document / information presentation may be based on visual abstracts, best matching snippets, extracted facts or structural. The alternatives are many – what makes sense depends in a complex way on contexts, i.e., works tasks, search tasks, other actors, and other available information objects, systems and tools.

The Access and Interaction Dimension. Topical well-defined requests on content (only) is just one approach to document retrieval albeit the most popular in IR research. Requests may be vaguely defined, non-topical (e.g., by journal or genre) and/or non-content-based (e.g., on given substructures). This will probably influence the nature of relevance and relevance assessment. The strategies of information access cover interaction modes like browsing and navigation in addition to retrieval. These may alternate and evolve from instance to instance of short-term interaction over session time and longitudinally due to the searcher's perception, line of progress, and learning. The alternatives are many – what makes sense depends in a complex way on works tasks, search tasks, other actors, and other available systems and tools.

3. IR RESEARCH IN ISOLATION

With a view on the eight broad dimensions presented above, traditional IR research is quite limited. While it has progressed considerably over the years, the context of use of IR systems has not developed sufficiently in IR research. Typically, the core of traditional IR is the Algorithmic Dimension in close interaction with the Document Dimension. That is the reason for trying out the same retrieval algorithms on many different types of media. But much more could be done exploring that association alone.

IR research typically considers only *retrieval tasks.* Moreover, these tasks are most often (a) purely topical, (b) content-only, (c) well-defined, (d) static, and (e) exhaustive retrieval tasks – one should find as many documents as possible matching the well-defined static topical need irrespective of document quality (binary topical relevance) and document overlaps. When designing and evaluating IR systems to serve such tasks one should identify the real-life seeking tasks that give rise to such retrieval tasks and their frequency. One should also identify alternative types of retrieval tasks, e.g., non-topical, non-content or structural, weakly defined, dynamic, and non-exhaustive – and various combinations. These have received much less attention in IR research.

Focus on the standard type of retrieval task is justified if (a) it clearly is the most frequent type in real life, and (b) by solving such tasks well all other types of retrieval tasks become easy to solve. Both points are at least questionable – perhaps incorrect while nobody knows the answers yet. Therefore IR should look into the non-standard retrieval tasks.

Still, one may claim the standard focus justified if the study of the alternatives would not make any difference in the design on IR systems. Several of the objections to the the laboratory model in IR culminate at this point. What are IR systems? – Algorithms for the representation and matching of documents and requests? Or tools for solving human information seeking tasks, contributing to work task performance? More fundamentally, what *is* IR as a discipline about? – About the algorithms for the standard retrieval task? Or about solving human information seeking problems through computers, with a focus on information represented in documents, as opposed to knowledge personally possessed by humans, and to data or collections of facts. If IR is about the algorithms only, the laboratory model may be justified. We believe however, based on our cognitive viewpoint, that IR should have a much broader focus than the focus on representation and matching of documents and requests.

Information seeking research was over the years often criticized for uselessness. Those working in the area have not been very critical anymore in the nineties but – we believe – the sentiment has been, and still is, shared by many working in information retrieval. One should therefore consider the motivations of the study of information seeking. In principle, the motivations, and benefits, may lie in (a) theoretically understanding information seeking, (b) empirically describing information seeking in various contexts, and (c) providing support to the design of information systems and information management.

4. AN IS&R DESIGN AND EVALUATION FRAMEWORK

Basically, we approach IS&R design and evaluation as embedded contexts of retrieval, seeking and work tasks/interests – Figure 2. IR serves the goals of seeking, and information seeking the goals of the work task (interest). The same person symbol in all the three contexts denotes the same or another actor(s) performing the work task, the seeking task and the retrieval task – interpreting the tasks, performing the process and interpreting the outcome – possibly resulting in task reformulation in each context. The person symbol in IR context signifies the possibility of applying human relevance feedback during a traditional two-run IR experiment as well as real interactive IR over several short-term interactions. Possible evaluation criteria in each context are given: A – D. The eight dimensions of variables outlined above are rewrapped in Figure 2.

As de-contextualized, IR may be designed and evaluated in its own context – the laboratory IR approach. In this confined context the evaluation measures are the traditional ones, recall and precision, or some novel measures. In addition, one may assess the system's efficiency along various dimensions during IR interaction, the quality of information (documents) retrieved, and the quality of the search process like searcher's effort (time), satisfaction, and various types of moves/tactics employed.

However, IR belongs to the searcher's information seeking context where it is but one means of gaining access to required information. This context provides a variety of information sources/systems and communication tools, all with different characteristics that may be used based on the seeker's discretion and in a concerted way. The design and evaluation of these sources/systems and tools needs to take their joint usability and quality of information and process into account. One may ask what is the contribution of an IR system in the end result of a seeking process – over time, over seeking tasks, and over seekers. Since the knowledge sources, systems and tools are *not* used in isolation they should not be designed nor evaluated in isolation. They affect each other's utility in context.

By looking at work task situations one may learn about the typical handles actors have available for accessing relevant information/documents.

Modern work is increasingly knowledge work where access to recorded information or human sources is essential. Task requirements must affect the design of information access. Nowadays, the means of access and sources increasingly become electronically networked and formalized in systems. This integration of e-generation, e-access, and e-use makes IR engineering complex – but not unmanageable. The question for IR engineering is: *which* additional *variables* from the immediate contexts one wishes to include in a controlled relationship with one another. The use of only one variable, as commonly attempted in laboratory IR, is insufficient and pursues only a limited case of IR.

Further, it is not just retrieval that matters, information systems also need to support reading (watching) as well as document processing and information use.

5. IS&R AND TASK PERFORMANCE AUGMENTATION

There are many *work task types* relevant for IS&R since they cause different kinds of information requirements and thus seeking and retrieval tasks *by actors*, and because they affect information use. The goal of IS&R is to augment work task performance and fulfillment. Figure 3 illustrates means and ends in task performance augmentation. Its upper part is inspired by D.C. Engelbart's (1963) framework for knowledge work augmentation, where a human is augmented by language, artifacts and methods in which (s)he has been trained. [2]

In Figure 3, information seeking is somewhat remote from the work task – with document retrieval even more remote and behind many decisions. In line with Figure 2 this underlines our view that IS&R belongs to a context in real life. The distance however does not make IR independent of work tasks – it needs to contribute to the work task, which sets a number of requirements on IR.

The work task type space hardly has been explored in Information Seeking and IR.

6. CONCLUSION

Based on this analysis we may conclude that:

- the *focus areas* in IS&R have been on one hand IR engines in strictly confined contexts and on the other hand information seeking behavior mostly without a work task context or with a narrow type of work task context;
- the *neglected areas* deserving more attention are work tasks and organizational contexts in general, and the interaction of several important dimensions in explanatory study designs; also any IT components other than algorithms for indexing, query formulation and matching deserve more attention.

Two action lines are therefore needed.

On the one hand, IR research needs to be extended to capture more context but without totally sacrificing the laboratory experimentation approach – the controlled experiments. Only by this line of action one may approach real *IR engineering*. IR engineering allows one to specify necessary IR system features by looking at the description of IR systems use in terms of tasks, users, documents and access requirements. Such features are, for instance, document and request representation, their matching, and various support tools. IR systems are thus seen in context of the other central components of the framework.

On the other hand, current information seeking research needs to be extended both toward the task context *and* the technology. We appreciate the efforts in information seeking so far exploring information seeking in diverse task/actor contexts but also think that the diversity of contexts is far from exhausted. Therefore lots of research is needed exploring IS&R in further task/actor contexts. Moreover, the systems context in information seeking research so far has been limited and often nonexistent. This research should reach toward system and interaction features so that communication with system design is facilitated.

7. REFERENCES

[1] Engelbart, D. (1962). *Augmenting Human Intellect: A conceptual Framework.* Stanford Research Institute, Menlo Park, CA.

[2] Ingwersen, P. (2001). Cognitive information retrieval. In: Williams, M. (ed.), *Annual Review of Information Science and Technology*, 34: 3-51.

[3] Ingwersen, P. & Järvelin, K. (2004). The Turn: Integration of Information Seeking and Retrieval in Context. Kluver. (In preparation).

[4] Järvelin, K. (1986). On information, information technology and the development of society: An information science perspective. In: Ingwersen, P., Kajberg, L. & Pejtersen, A.M. (eds.), *Information Technology and Information Use.* London, UK: Taylor Graham: 35-55.

[5] Kekäläinen, J. & Järvelin, K. (2002b). Evaluating Information Retrieval Systems under the Challenges of Interaction and Multi-Dimensional Dynamic Relevance. In: Bruce, H, Fidel, R., Ingwersen, P. & Vakkari, P. (2002) (eds.), Emerging Frameworks and Methods. *Proceedings of the 4th International Conference on Conceptions of Library and Information Science (CoLIS 4)*, July 21-25, 2002, Seattle, USA. Greenwood Village, CO: Libraries Unlimited: 253 – 270.

---

[2] Engelbart (1963) proposes a framework for *augmenting human intellect.* This is the ultimate goal of *instrumental* IS&R no matter whether it takes place in professional or leisure contexts. This is a strong legitimization to our *cognitive* viewpoint – IS&R should augment human intellect – in context.

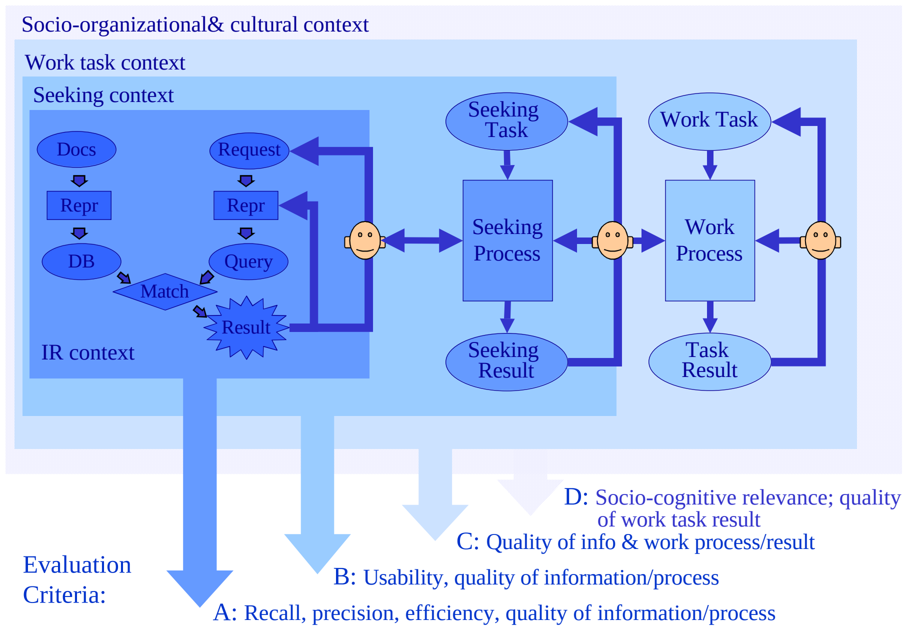


Figure 2. Nested contexts and evaluation criteria for task-based IS&R (based on Kekäläinen & Järvelin, 2002)

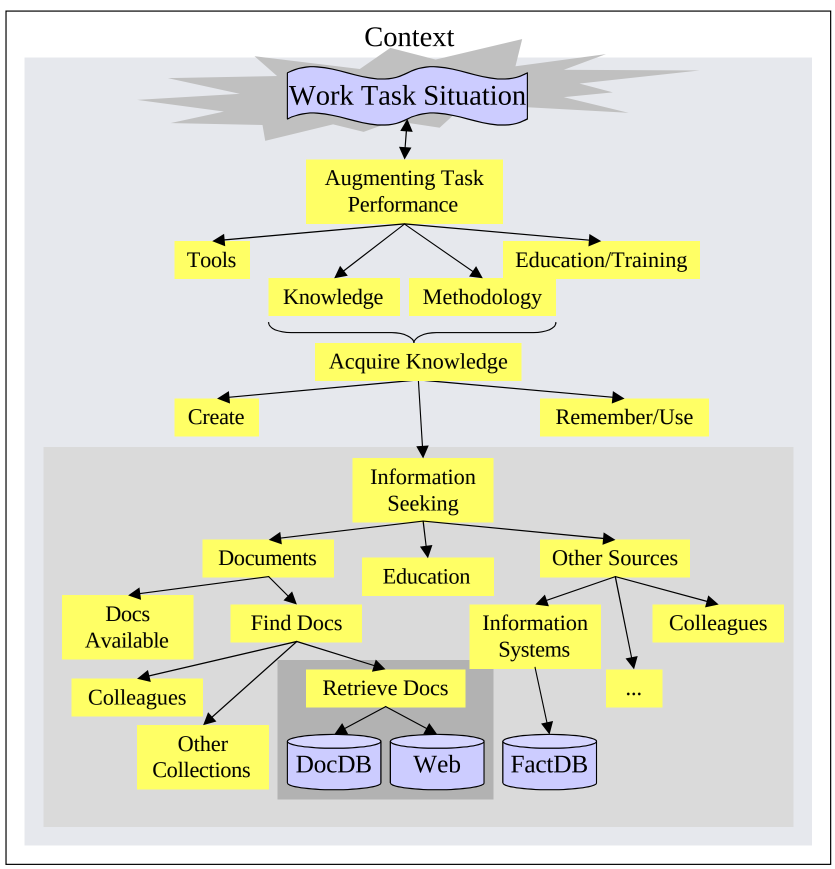


Figure 3. Augmenting work task performance – perhaps by IS&R (based on Järvelin, 1986)

Information Retrieval as Multitasking: An Exploratory Framework

Dr. Amanda Spink & Minsoo Park
School of Information Sciences
University of Pittsburgh
610 IS Building, 135 N. Bellefield Avenue
Pittsburgh PA 15260
Tel: (412) 624-9454 Fax: (412) 648-7001
E-mail: aspink@sis.pitt.edu, mpark@sis.pitt.edu

1. INTRODUCTION

This paper proposes an approach to interactive information retrieval (IR) contextually within a multitasking framework. During multitasking, humans cognitively and physically coordinate multiple tasks through task switching. The paper proposes that interactive IR is contextually a multitasking behavior on two levels. First, on an i*nteractive search task* level, people construct an interactive IR session as a series of tasks, including an embedded interplay of information problem, interactive search and other tasks.

For example, embedded between telephoning and computing tasks, a search engine user coordinates many tasks when looking for medical information, such as translating their information problem into a set of search terms and strategy, search engine and search term selection, relevance judgments, etc.

On a second level, people engage in *multitasking information behaviors* or are seeking information on more than one topic concurrently. For example, a search engine user switches between seeking fashion information and medical information. Interactive IR can be conceptualized as interplay between different types of tasks and often different information problems.

Conceptualizing interactive IR as a multitasking process embeds interactive IR within the broader framework of multitasking research in the cognitive/behavioral sciences.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

Recent cognitive/behavioral science studies suggest that dual tasking (driving and talking concurrently) and workplace multitasking behaviors on unrelated task are counterproductive [1, 2]. However, one can argue that effective interactive IR is coordinating switching between related tasks.

To elaborate on this notion further, this paper first provides a brief overview of multitasking behavior studies, and then conceptualizes interactive IR within two levels of multitasking, and finally provides a model of interactive IR as multitasking.

2. MULTITASKING RESEARCH

2.1 Multitasking Behavior

In the broad context of human behavior studies, due to the increasing complexity of the global information environment, people are increasing engaged in critical multitasking and task switching behaviors [3, 4]. Multitasking is the ability of humans to simultaneously handle the demands of multiple tasks via task switching [1, 2].

Wickens [5] suggests that *time sharing* allows the simultaneous performance of multiple tasks and time swapping allows the sequential performance of tasks. Alternatively, dual tasking is the human ability to simultaneously handle the demands of multiple tasks, such as talking on the telephone whilst driving a car [1]. Multitasking is also an important research area for technology designers [6].

However, in the IR context, critical multitasking behaviors, particularly multitasking information behaviors, are still largely under-researched. In addition, current search technologies are designed to largely support only limited types of searching by specifying queries using terms to select documents/Web sites to fulfill a single information task. But, interactive IR is in fact accomplished by people in much more complex ways than just this method of query specification and selection.

The next section of the paper conceptualizes outlines interactive IR as two levels of multitasking.

3. INTERACTIVE INFORMATION RETRIEVAL: MULTITASKING LEVELS

3.1 Interactive Information Retrieval Tasks: Coordinating and Multitasking

On one level, people construct an interactive IR session as a series of tasks and interplay of interactive search tasks. For example, embedded between telephoning and computing tasks, a search engine user may coordinate their interactive search tasks when looking for medical information.

People must coordinate the translation of their information problem(s) by performing search term selection task, tactic and strategy tasks, search engine interaction tasks, relevance judgments, etc. Research shows that human's have different levels of cognitive coordination [7, 8]. Interactive IR occurs as series of coordinated task actions. To achieve interactive IR, humans' *coordinate* a number of tasks, including their cognitive state, level of domain knowledge, and their understanding of their information problem, into a coherent series of activities that may include seeking, searching, interactive browsing and retrieving and constructing information.

Humans cognitively coordinate their information seeking level behaviors with their interactive searching level (human-system interaction) level behaviors; including the recognition and making sense of and cognitively articulating an information problem or a gap in their knowledge. Human's then coordinate these processes to construct an interactive IR process embedded within their broader information and non-information behaviors [9].

Establishing and sustaining an effective interactive IR process require humans' to coherently coordinate and multitask their information problem and interactive search tasks. In other words, an information seeker must coordinate a number of tasks, including their cognitive state, level of knowledge, their understanding of their information problem, into a coherent series of sustained activities that may include seeking, searching, retrieving and using information. We know that hand-eye coordination is a physiological process that humans develop from childhood. But, how do humans learn the process of cognitively coordinating their information problems into coherent processes of human information behavior and interactive IR?

3.2 Multitasking Information Behaviors

On a second level, people often engage in multitasking information problems concurrently. Studies have highlighted the nature of task in information behavior [10], but have focused on single tasks and generally consider information task in isolation form other tasks. Recent empirical studies show that information retrieval system users often engage in *multitasking information behaviors* [11 12, 13, 14, 15, 16, 17].

Information behaviors are not limited to single discrete information problem tasks, but often range over multiple topics or browsing behavior on specific multiple topics. Therefore, many humans engage in information related multitasking behaviors during interactive IR. Spink, Ozmutlu and Ozmutlu [13] first identified information multitasking processes in four studies conducted within different information environments, including library use. They interviewed library users and found that people often seek information in a library for information on more than one information task during a single or multiple library use episodes.

Spink [16] found from library use episodes recorded in a diary that library use often includes many task switches between different topics. Library users' also optimized their interaction time by engaging in multitasking searches rather than separate searches on discrete topics over time. Multitasking information behaviors are often driven by the complex nature and frequency of users' information problems.

Research studying Web search transaction logs shows the increasing complexity in users' behaviors related to searching that are reflected in successive searching and multitasking search [12]. Multitasking searches include more keywords, queries and query reformulation than single topic searches. Users reported problems in coordinating, tracking and managing their multitasking searches. Spink, Jansen and Park studied Alta Vista Web search transaction logs of two-query and three or more query search sessions [14, 15, 17). Findings included: (1) 81% of two-query sessions included multiple topics, (2) 91.3% of three or more query sessions included multiple topics, (3) there are a broad variety of topics in multitasking search sessions, and (4) multitasking sessions contained frequent topic changes.

In the interactive IR context, an IR system user may multitask (either begin their search with multiple topics or develop further topics during the search process, and information task switch (switch back and forth between different topics during a search session). Users often search on more than one information task (topic) during a single search or multiple search interactions. Users may engage in many related multitasking search episodes over time. For example, a person switches between seeking health information and fashion information as they are thinking and working on multiple information problems concurrently. However, search technologies require them to search sequentially.

The next section of the paper provides an initial model of interactive IR as a multitasking behavior.

4. MODEL OF INTERACTIVE IR AS MULTITASKING

Figure 1 provides a model of interactive IR as multitasking.

Figure 1. Model of interactive IR as multitasking.

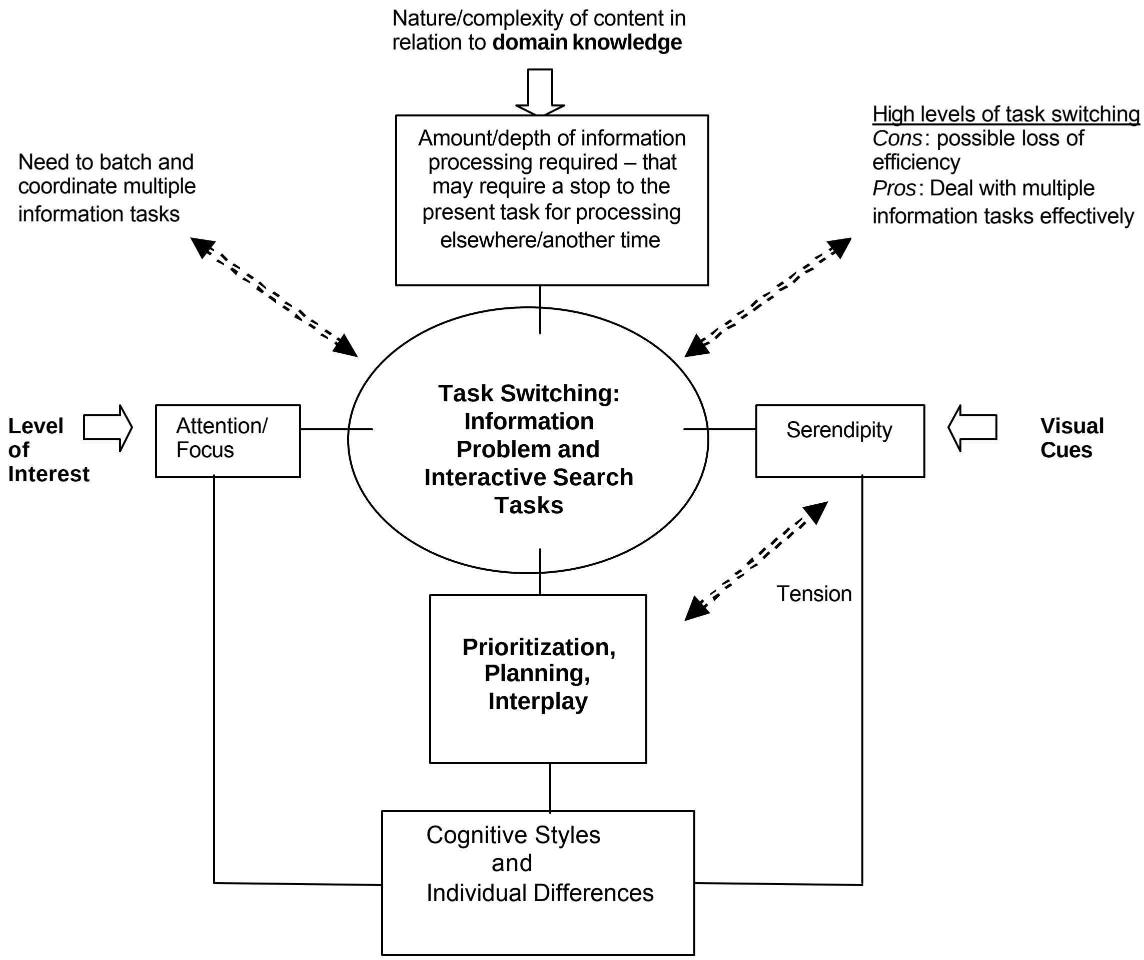


Figure 1 includes task switching between information problem and interactive search tasks. This process may occur over different timeframes and levels of rapidity. This task interplay could also affected by a number of factors, including:

1. Nature and complexity of content in relation to the information seeker's *domain knowledge.*
2. Amount and depth of *information processing* required for different information tasks.
3. Information seeker's *level of interest*, including their attention and focus, in the information task.
4. Level of *planning and priorities* by the information seeker in relation to their information tasks.
5. Pros and cons or the *effects on effectiveness, efficiency and productivity* of information task switching.
6. S*erendipity* by the information seeker that is prompted by visual information cues and the tension with the planning and priority goals, and task interplays.
7. *Cognitive styles and individual differences* associated with attention/focus, task prioritization, task planning, and task interplay.

tasks.

Multitasking information behavior research is a significant area of study. Despite the focus on tasks [10], current models of interactive IR do not consider multitasking behaviors. Human information behavior is more complex that the consideration of information tasks in isolation from people's other tasks.

Understanding and modeling multitasking information behaviors, requires an understanding of the coordination and interplay between information problem, interactive search and other tasks.

4. CONCLUSION

This paper has proposed that theoretically and practically, interactive IR can be conceptualized as a multitasking and coordinating processes on various levels as interplay of information problem and interactive search tasks. Exploring multitasking and coordination behaviors are relatively new and heuristic direction for interactive IR research. The authors are currently conducting further studies to extend our understanding of the nature, patterns and impacts of interactive IR as multitasking.

5. REFERENCES

[1] Just, M. A., Carpenter, P. A., Keller, T. A., Emery, L., Zajac, H., and Thulborn, K. Interdependence of non-overlapping cortical systems in dual cognitive task. *Neuroimage, 14*, (2001), 417-426.

[2] Rubenstein, J., Meyer, D., and Evans, J. Executive control of cognitive processes in task switching. *Journal of Experimental Psychology, 27,* 4 (2001), 763-797.

[3] Burgess, P. W. Real-world multitasking from a cognitive neuroscience perspective. In: Stephen Monsell and Jon Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII*. Cambridge, MA: The MIT Press, pp. 465-472, 2000.

[4] Carlson, R. A. and Sohn, M.-Y. Cognitive control of multistep routines: Information processing and conscious intentions. In: Stephen Monsell and Jon Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII*. Cambridge, MA: The MIT Press (pp. 443-464), 2000.

[5] Wickens, C. D. *Engineering Psychology and Human Performance*. New York: HarperCollins, 1992.

[6] MacIntyre, B., Mynatt, E. D., Voida, S., Hansen, K. M., Tullio, J., and Corso, G. M. Support for multitasking and background awareness using interactive peripheral displays. *UIST'01: ACM User Interface Software and Technology 2001, Nov. 11-14, Orlando, FL*.

[7] Ivry, R. B., and Richardson, T. C. Temporal control and coordination: The multiple timer model. *Brain and Cognition, 48,* 1 (2002), 117-132.

[8] Olsen, G. M., Malone, T., and Smith, J. B. (Eds.), *Coordination Theory and Collaboration Technology*. Mahwah, NJ: Lawrence Erlbaum Associates, 2001.

[9] Spink, A., and Park. M. (Forthcoming). *Multitasking Interplay of Information and Non-Information Tasks: A Model.*

[10]Vakkari, P. Task-based information searching. *Annual Review of Information Science and Technology, 37* (2003), 413-464.

[11]Ozmutlu, S., Ozmutlu, H. C., and Spink, A. Multitasking Web search: Implications for design. *ASIST'03: Proceedings of the American Society for Information Science and Technology, Oct 18-22, 2003. Long Beach, CA* 2003a.

[12] Ozmutlu, S., Ozmutlu, H. C., and Spink, A. A study of multitasking Web searching. *IEEE ITCC'03: Conference on Information Technology: Coding and Computing, April 3-5, 2003, Las Vegas, NV* 2003b.

[13] Spink, A., Ozmutlu, S., and Ozmutlu, H. C. Multitasking information seeking and searching processes. *Journal of the American Society for Information Science and Technology, 53,* 8 (2002), 639-652.

[14]Spink, A., Jansen, B. J., and Park, M. Multitasking during Web search. *ASIST'04: American Society for Information Sciences and Technology Conference, Providence, RI, November 2004*, 2004a.

[15]Spink, A., Jansen, B. J., and Park, M. Multitasking on Alta Vista. *IEEE ITCC'04: International Conference on Coding and Computing, April 2-5, Las Vegas, NV* 2004b.

[16]Spink, A. Multitasking information behavior and information task switching: An exploratory study. *Journal of Documentation,* 60, 4 (2004).

[17]Spink, A., Park, M., and Jansen, B. J. *Multitasking During Web Search Sessions*, Forthcoming.

The Role of Context in Information Retrieval

Gareth J. F. Jones
School of Computing &
Centre for Digital Video Processing
Dublin City University
Glasnevin, Dublin 9, Ireland
Gareth.Jones@computing.dcu.ie

Peter J. Brown
Department of Computer Science
University of Exeter
Exeter EX4 4QF, United Kingdom
P.J.Brown@exeter.ac.uk

ABSTRACT

The scope for information retrieval applications continues to growth to encompass larger and more diverse archives and new computing environments. The consequent increased demands on information retrieval systems continue to motivate research into new and more efficient search technologies. Users of information retrieval systems work in personal and physical contexts, while the documents containing the information they seek often relate to specific contexts. It is argued that taking account of context can improve information retrieval effectiveness. While potential use of context has certainly been under explored, it is already present in established techniques such as relevance feedback. The emergence of new information retrieval environments, such as those associated with mobile computing, raises new challenges for information retrieval for which greater use of context may form a important component. Among the many questions raised by attempting to perform information retrieval in context are issues of how algorithms might be extended or developed to facilitate use of context, does the user need to be actively involved to make use of context, where should the context information come from, and how might the effectiveness of such methods be tested in extended evaluation frameworks?

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Algorithms, Experimentation

Keywords

context, information retrieval

SIGIR'04 July 25-29, 2004, Sheffield, UK.

.

1. INTRODUCTION

Information retrieval (IR) is often defined in terms of the location and delivery of documents to a user to satisfy their information need. This apparently simple task turns out to be highly complex. In order to facilitate research and move towards real systems various pragmatic assumptions and simplifications have traditionally been made in IR research. A key factor that has generally been factored out is explicit modelling of the *context* in which the search is being carried out, although as we demonstrate later it is already used implicity in at least one IR technique.

The notion of context has been one of the mostly widely interpreted, or possibly abused [1], terms in IR, encompassing topics such as interaction, domain specific search, personalization, [2] and environmental sensors [3]. None of these or other interpretations are necessarily wrong, but neither we feel are they exclusive.

IR has conventionally been concerned with users working inside an "information world" using desk-based systems. However, developments in mobile and wireless computing mean that users of IR systems may now be embedded in the "real/physical world" surrounded by environmental sensors providing rich information of physical context [4].

An important question then is, how should "information retrieval in context (IRiX)" be interpreted? Should it refer only to taking account of the information world, and within this refer only to what has often classified as "interactive IR"; should the scope be broader encompassing personalization and implicit feedback [2], or should it be broader still to covering mobile IR as well [4]. Further than this, from our analysis, once one begins to look for issues of context within IR it starts to emerge everywhere. Should IRiX then focus on all forms of context information wherever it can be found in the IR process? One interesting question with regard to IR and context that we ask ourselves is, does IRiX really imply interaction in IR? Are there opportunities for context to be exploited without any user involvement or modelling of the individual user, we argue that the answer appears to be yes.

However the scope of IRiX is defined, the overall aim in taking account of context within IR is to improve the efficiency with which the user's information need is satisfied. The appropriate measures of effectiveness here will themselves often depend on the informational or physical context in which the search is being conducted. This may perhaps imply that the user has to do less work to satisfy their information need if context is incorporated into the search process, although depending on how the context data is entered and used, it may actually be harder work for the user. By some measures the complexity of the retrieval process might be reduced by incorporating context, but by others it may well be increased.

The remainder of this paper analyses the use of context within current IR, discusses the challenges and opportunities of IR for mobile and ubiquitous computing environments, and outlines some possibilities for exploring the extension of the role of context within IR.

2. CONTEXT IN CURRENT IR SYSTEMS

While it is to a large extent true that current IR systems take no account of context explicitly, any information given to the IR system by the user and used in computation of the output can be regarded as some form of search context.

Current IR models generally assume that document attributes, typically words, are independent [5]. This is a practical approximation enabling the development of computationally tractable algorithms with parameters that can be reasonably estimated from the data available. However, even within this model we can think of attributes appearing in the context of other ones. For example, relevance feedback (RF), already implicity makes use of attribute context. Within RF documents are placed in the context of their relevance to information need rather than seen as independent items within an overall document collection [6]. Feedback from the user provides context dependent information to the IR system on the relevance of a specific document. Using this feedback information, RF seeks to model the importance of terms to the information need and expands the request to include those terms deemed to make it a better expression of the information need. Interestingly pseudo or blind RF, which operates without relevance input from the user effectively, automatically develops context between documents retrieved using the initial search. Search context is being inferred here without any user involvement. Some RF methods take account of word context within documents, for example local context analysis [7] and document summaries [8].

IRiX for searching in the information world relates to taking account of the user and their cognitive attributes. This may result in the search being personalized or the user placed within a user class [2]. While it is sometimes difficult even to elicit user feedback of document relevance, personalization of search often seems to be based on elaborate interfaces requiring considerable engagement from the user [9] or marking relevant material for feedback [1]. While these interfaces may be effective when used properly, even by novice users, it must surely be open to question whether the typical user with a real information need will make the effort to use them properly. Perhaps we should focus our attention for IRiX more within the system. As the next section illustrates this is an even more significant question for mobile IR.

3. CONTEXT IN MOBILE AND UBIQUITOUS INFORMATION ACCESS

Mobile computing devices range from "semi" portable, e.g. a laptop, to truly mobile, e.g. mobile phones and PDAs, indeed sometimes the device may effectively be embedded in the user's environment, e.g. a car computer. This change in the working environment has many implications for the user needs in respect of IR which involve information and context.

Various studies have considered the implications of this new environment for IR [10][4][11][3]. Some important conclusions of these and other studies are that:

- the user is often engaged primarily in another activity other than information searching and is likely to want rapid access to relevant information to assist them.
- the user is embedded in the "real world", and will often require information accurate up-to-date information relating to their current physical circumstances.
- particularly for small handheld devices anything more than cursory searching and browsing is impractical.
- the potential of physical context data to improve IR precision appears attractive and should be explored.

Context in these studies has tended to focus on the physical world, rather than the information world more typically associated with interactive IR. The issues of information world IRiX still apply here, although they may be modified with respect to the types of user activities or searching. Here we consider only the additional context features associated with the physical world.

In mobile IR applications much physical context data is potentially available via personal and environmental sensors; for example, the user's location, who they are with, ambient temperature, traffic conditions, nearby shops and offices, and their attributes. Some of this information might be used directly, while other data must aggregated to infer higher level context or user activity before it is useful in IR. As well as its use in IR, this context data might also be used to determine the best mode of information delivery, e.g. via audio output while the user is driving [3].

In previous work we have explored issues of IRiX relating to the physical world for mobile computing environments. In this work we developed the concept of the *context-of-interest* [11]. The basis of this proposal is that mobile users are often likely to be most interested in information referring to their future context rather than their current one, which may be out-of-date when the information is actually delivered. This led us to consider the topic of *prediction* of physical context. This may be based on user independent physical context tracking or be personalised through incorporation of information from a user's *context diary* of recent, current and expected activities.

It has also been noted that users of small mobile devices are less able to interact with them [10][12]. Thus, IR systems on these devices must behave more autonomously than desk-based systems, and attempt to provide users with the relevant information they need without significant browsing.

Also as noted above users may often be engaged in other activities rendering them unable to search for information of use to them in their current physical context, or they may not be aware of the existence of information which they might find useful. In this situation it would be useful if the IR system were to behave *pro-actively* to look for potentially interesting information by making use of context data to form form search queries automatically. This is more akin to information filtering or routing than IR, but is nevertheless a further example of non-interactive IRiX.

4. CONTEXT IN IR ALGORITHMS

While the physical operating environment and means of gathering context information are different for traditional desk-based IR and mobile IR, the issues of how to use context within IR algorithms raise fundamentally the same questions, and may have the same or related solutions.

A key concern is how to model context data as attributes within an IR system, and how to incorporate these within IR algorithms. This raises questions of establishing context classes and defining context attributes within these classes. Linguistic context features can be sometimes be handled in a straightforward way using relevance feedback type algorithms for learning and personalization, although handling even these features will often be more complex than this. More problematic in general is other context data for which the attributes are not obvious and for which suitable similarity coefficients must be sought.

One simple way to use context in IR would be to use them to introduce boolean type constraints on existing search mechanisms, e.g. to limit the time range searched for documents. This is a very limited vision of the use of context. In some situations there may be a balance between the level of context match and the level of content match. A document which is highly matched based on context while poorly matched based on content, may be judged by the user to be more or less useful than one which matches poorly on context and highly on content [13]. Matching algorithms for IR in context should be able to take account of this.

Consider the popular context features of time and location. These have continuous values that can be described in precise or vague terms at different levels of granularity. For example, an event may take place at `9:57`, at `about 10 o'clock`, `in the morning`, and at a given numerical grid reference, or in a named street, district, or city. In this case it is not clear what the appropriate representation is and how the significance of each feature should be defined in terms of selectivity. Determining how to compute a scored match between context fields of varied specificities, and how to integrate these scores with traditional IR content-based query-document matching in a well-founded way is a key issue here. A novel approach to matching between temporal and location fields for topic detection and tracking is introduced in [14]. These methods assign a score to degree of temporal overlap between an event in two documents and compute a matching score between locations using a geographical ontology tree, e.g. `Paris` is in `France` which is in `Europe`. While interesting, these methods do not currently improve retrieval effectiveness beyond standard content matching. A further important issue is the extent to which the context associated with documents and queries has to be stated explicitly or can be extracted automatically from their contents [2][3]. Finding potential context information within a document is only part of the problem, it may then be necessary to try to find associations between context data using shallow natural language processing methods, perhaps in a manner similar to event construction in information extraction methods as applied in question-answering systems.

Once some similarity measure between context of the same type has been established this could be incorporated in an extended form relevance feedback. Related context information in retrieved documents could be identified and used to extend the contextual annotation of the individual documents, to determine the likely importance of this context information across the document collection, and then to extend the search request to include significant context labels. The notion of using cross collection statistics to establish reliable features associated with an event has been already been in explored in [15], and this might be extended to determine significant and reliable context labels. The success of this idea requires the development and extension of a number of techniques, but could offer a way to integrate context of various forms into an IR search.

5. REFERENCES

[1] L. Finkelstein, et al. Placing Search in Context: The Concept Revisited. *ACM Transactions on Information Systems*, 20(1):116–131, 2002.

[2] S. Lawrence. Context in Web Search. *IEEE Data Engineering Bulletin*, 23(3):25–32, 2000.

[3] G. J. F. Jones and P. J. Brown. *Context-Aware Retrieval for Ubiquitous Computing Environments*. In *Proceedings of the Mobile HCI 2003 International Workshop on Mobile and Ubiquitous Information Access*, pages 227–243, Udine, 2003.

[4] P. Coppola, V. Della Mea, L. Di Gaspero and S. Mizzaro. *The Concept of Relevance in Mobile and Ubiquitous Information Access*. In *Proceedings of the Mobile HCI 2003 International Workshop on Mobile and Ubiquitous Information Access*, pages 1–10, Udine, 2003.

[5] C. J. van Rijsbergen. *Information Retrieval* (2nd edition), Butterworths, 1979.

[6] G. Salton and C. Buckley. Improving Retrieval Performance by Relevance Feedback. *J. American Society of Information Science*, 41(4):288–297, 1990.

[7] J. Xu and W. B. Croft. Improving the Effectiveness of Information Retrieval with Local Context Analysis. *ACM Transactions on Information Systems*, 18(1):79–112, 2000.

[8] A. M. Lam-Adesina and G. J. F. Jones. Applying Summarization Techniques for Term Selection in Relevance Feedback. In *Proceedings of ACM SIGIR 2001*, pages 1–9, New Orleans, 2001. ACM.

[9] J. Pitkow, et al. Personalized Search. *Communications of the ACM*, 45(9):50–55, 2002.

[10] B. J. Rhodes and P. Maes. Just-In-Time Information Retrieval Agents. *IBM Systems Journal*, 39(3&4):685–704, 2000.

[11] P. J. Brown and G. J. F. Jones. *Exploiting Contextual Change in Context-Aware Retrieval*. In *Proceedings of the 17th ACM Symposium on Applied Computing (SAC 2002)*, pages 650–656, Madrid, 2002.

[12] C. Gurrin, et al. *Mobile Access to the Fischlár-News Archive*. In *Proceedings of the Mobile HCI 2003 International Workshop on Mobile and Ubiquitous Information Access*, pages 124–142, Udine, 2003.

[13] G. J. F. Jones and P. J. Brown. *Information Access for Context-Aware Applications*. In *Proceedings of the ACM SIGIR 2000*, pages 382–384, Athens, 2000.

[14] J. Makkonen, H. Ahonen-Myka and M. Salmenkivi. Simple Semantics in Topic Detection and Tracking. *Information Retrieval*, 7(3&4):347–368, 2004.

[15] S. Dumais et al. Web Question Answering: Is More Always Better? In *Proceedings ACM SIGIR 2002*, pages 291–298, Tampere, 2002. ACM.

Using User's Context for IR Personalization

N.J. Belkin, G. Muresan, X.-M. Zhang
School of Communication, Information and Library Studies
Rutgers University, New Brunswick, NJ 08901
+1-732-932-7500
[belkin, muresan, xzhang]@scils.rutgers.edu

1 INTRODUCTION

We believe that knowledge of the user's context is vital to personalizing information retrieval (IR) interaction, and furthermore, that such knowledge is best obtained through implicit sources of evidence, e.g. inferences made on the basis of the user's past or current behaviors. We further believe that it is now well past time to *test* whether such knowledge really does affect the interactive IR experience. Some contextual factors that have been suggested as being important to consider in improving IR performance, by us and others, include:

- searcher's familiarity with or knowledge of the topic;
- searcher's experience of searching for information;
- documents which the searcher has previously found (un)useful;
- genre of desired documents;
- purpose of the search (use to which retrieved documents would be put);
- task which led the searcher to information seeking;
- what else the user is doing at the time of information seeking.

In order to test our ideas about what aspects of a user's context might be important in this sense, and how knowledge of these characteristics could be utilized to affect various IR techniques, our group at Rutgers University participates in the TREC HARD track. Here, we present an overview of the general HARD approach, our position on how this issue should be addressed, and how we have attempted, and are attempting to implement such knowledge.

2 THE HARD TRACK

The HARD Track investigates the effect of knowledge of user's context on IR system performance in the following way. Search topics are constructed by assessors, with respect to issues of interest to them. These topics follow the general TREC format, with the addition of categories of metadata whose values describe various aspects of the assessor's context. In TREC 2003, the metadata were:

- familiarity with the topic
- desired genre of retrieved documents
- purpose of the search
- specification of geographic focus of documents

In TREC 2004, the categories of metadata were reduced to:

- knowledge of the topic
- desired genre of retrieved documents
- documents should be about USA, or not USA

In addition to these categories of metadata, assessors also specify one or two documents related to the topic, and the granularity of response they require.

Participating sites are initially given the document corpus and a set of training topics, each with the metadata and 100 documents which have been judged either *not* relevant, *soft* relevant (i.e. on topic), or *hard* relevant (i.e. on topic, *and* satisfying the metadata conditions). Then, the set of test topics is distributed, *without* the user metadata. Each site constructs a query for each topic, searches the corpus and returns a ranked list of documents for each topic. This constitutes the *baseline* run. Then, the metadata and other information for each topic are distributed to all sites. The sites then use information derived from the metadata to modify the retrieval techniques (e.g. modify the query, re-rank the baseline list). In addition, they may submit a clarification form to the assessor, asking one simple, limited question concerning some aspect of the initial retrieval performance (e.g. which of these clusters of retrieved documents do you find most interesting). The sites then submit one or more new runs, based on the information received. These are the *test* runs.

The results of baseline and test runs are pooled, and evaluated by the original assessors according to the three categories of relevance. The test of the utility of the modifications that have been made is the difference in performance between the baseline and test runs, *judged according to hard relevance*. For a detailed explanation of the TREC 2003 HARD track, see Allan (2004).

3 RUTGERS' APPROACH TO HARD

3.1 Introduction

The goal of our work in the HARD track is to test techniques for using knowledge about various aspects of the information seeker's context to improve IR system performance. We are particularly concerned with such knowledge which could be

SIGIR'04 July 25-29, 2004, Sheffield, UK.

gained through implicit sources of evidence, rather than explicit questioning of the information seeker. We therefore do not submit any clarification form. Of the categories of metadata and related information which are available in the HARD track, we have chosen to investigate:

- searcher familiarity with (or knowledge of) the topic, since there is some evidence that it is important (Kelly and Cool, 2002) and evidence of this could be gained through observation of previous behavior;
- desired genre, for the same reasons (e.g. Rauber & Müller-Kögler, 2001); and
- related texts, since they could be inferred on the basis of both past and current behaviors.

3.2 HARD 2003

In TREC 2003 we attempted to test the following hypotheses:

H1: People familiar with a topic will want to see documents which are detailed and terminologically specific; people unfamiliar with a topic will want to see general and relatively simple documents. This we operationalized by promoting the value of documents which scored toward the unreadable end of readability scales for people highly familiar with the topic, and by promoting the value of documents which scored toward the easily readable end of the scales for people unfamiliar with the topic.

H2: Different document genres can be identified by their vocabularies. This we operationalized by constructing language models for all the retrieved and for just the hard relevant documents for each training topic. By comparing language models, we then identified words which occurred with greater than expected probability in the relevant documents, for all topics which had the same genre. These words were considered indicators of the genre and were used for query expansion for topics which requested that genre.

H3: Certain document sources will be relevant, or not, to different desired genres. This we operationalized by promoting or demoting the score of documents, or by removing documents from the ranked list, according to their source and the requested genre.

H4: If there are texts which the information searcher has identified as related to the topic, using them as the basis for automatic query expansion will improve retrieval performance. This was operationalized by choosing terms for query expansion from the related texts.

We understood that there are, in general, two ways in which to take account of the metadata. One is to modify the initial query from the (presumed) searcher, before submitting it for search; the other is to search with the initial query, and then modify (i.e. re-rank) the results before showing them to the searcher. We used both of these techniques in taking account of the different types of metadata.

Our results in the 2003 HARD Track indicated a few interesting trends, but were generally poor. Detailed analysis suggests that we did gain some advantage from using the metadata to modify the baseline queries, in some respects, and query expansion via related documents did help. But the ways in which we used the metadata to modify rankings and queries were quite ad hoc, and without real theoretical justification, which could go some way toward explaining negative results. A more detailed report on our HARD 2003 experience can be found in Belkin et al. (2004).

There were some significant problems with HARD 2003. The training data were insufficient, the familiarity scale did not actually judge the assessor's real familiarity with the topic, and there was insufficient representation of different values of the different metadata for training and testing purposes. Thus, for HARD 2004, both the number of metadata factors, and the number of values which they could take, were reduced.

3.3 HARD 2004

In HARD 2004, we are attempting to make our hypotheses more formal, and to move from ad hoc implementations of our hypotheses to more principled ones. This is still in progress, but we outline them here. In addition to query expansion from related documents, Rutgers is investigating the following two issues:

- how can we take account of a searcher' knowledge of a topic to improve retrieval performance; and
- how can we take account of knowledge of desired genre to improve retrieval performance.

In HARD 2004, there are only two values of knowledge of a topic: *little*; and *a great deal*, and there are only three values of genre: *news-report*; *opinion*; and *other* (the corpus consists of news sources). With respect to these issues, we consider the following hypotheses:

H1: People with a great deal of knowledge of a topic will want to see documents which are detailed and terminologically specific; people with little knowledge of a topic will want to see general and relatively simple documents. This is the same hypothesis that we had with respect to familiarity in TREC 2003. However, we are investigating the use of different readability measures, which are more directly concerned with terminology than those we used in TREC 2003.

In addition, we are investigating two new hypotheses with respect to knowledge of the topic. These have to do with findings that people with little knowledge of a topic cannot interpret and understand *abstract* concepts in the topic domain as well as those who have good knowledge, and that words indicating *concrete* concepts are in general more easily understood than abstract ones. This leads us to the following:

H2: People with little knowledge of the topic will prefer documents with a low ratio of abstract words to total words, and of abstract words to concrete words. People with good knowledge of a topic will prefer documents which have a high ratio of abstract words to total words, and of abstract words to concrete words. This hypothesis leads to a re-ranking strategy.

H3: Adding concrete terms to the initial query from the topic domain (as determined by initial retrieval results) will lead to more effective results for people with little knowledge of the topic; adding abstract terms from the topic domain will lead to more effective results for people with a great deal of knowledge of the topic. This is a query modification strategy.

With respect to genre, we have several hypotheses, one of which is directly related to those of TREC 2003. They are all based on the idea that no matter what the topic, documents of specific genres will share some common characteristics which can be identified through different sorts of analyses.

H4: The differences between the genres of news-report and opinion can be identified according to the degree of subjectivity or objectivity of a document, as determined by various linguistic features of the documents (cf. Rittman, 2004). This leads to a classification and re-ranking strategy.

H5: Different document genres will have different characteristic vocabularies, regardless of topic. This is essentially the same hypothesis as for TREC 2003, and we investigate it by again developing language models for the topic in general (i.e. soft relevant), and those for the different genres within each topic. Words which occur with greater than expected frequency with respect to the topic models for a particular genre, across all topics, will be indicative of the genre's vocabulary. This technique can be used both to identify words which can be added to a query (query modification strategy), and to classify documents which belong to a specific genre (re-ranking strategy).

H6: Different document genres will have different discourse-level features characteristic of each genre, regardless of topic. We will determine these features with the training collection, and use them to classify initially retrieved documents. This leads to a re-ranking strategy.

4 CONCLUSION

We have outlined a general experimental approach to evaluating the effectiveness of taking account of different aspects of a searcher's context in personalizing retrieval to that person. Although initial results are still very sketchy, we believe that following this route is necessary in order to determine whether context really is important, what aspects of context are really important, and how knowledge of such aspects can best be taken account of.

5 REFERENCES

Allan, J (2004) Overview of the TREC 2003 HARD track. In Proceedings of the 12th Text REtrieval Conferenc (TREC 2003). Retrieved from http://trec.nist.gov/ pubs/trec12/t12_track.index.html on 28 May 2004.

N.J. Belkin, D. Kelly, H.-J. Lee, Y.-L. Li, G. Muresan, M.-C. Tang, X.-J. Yuan, X.-M. Zhang (2004) Rutgers' HARD and Web Interactive Track Experiences at TREC 2003. In Proceedings of the 12th Text REtrieval Conferenc (TREC 2003). Retrieved from http://trec.nist.gov/ pubs/trec12/t12_track.index.html on 28 May 2004.

Kelly, D. & Cool, C. (2002) Effects of topic familiarity on information search behavior. In *Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries – JCDL 2002* (pp. 74-75). New York: ACM.

Rauber, A. & Müller-Kögler, A. (2001) Integrating automatic genre analysis into digital libraries. In *Proceedings of the First ACM/IEEE-CS Joint Conference on Digital Libraries - JCDL 2001* (pp. 1-10). New York: ACM.

Rittman, R. (2004) Adjectives as Indicators of Subjectivity in Documents. In Proceedings of the 2004 Annual Meeting of the American Society for Information Science and Technology. in press.

Identifying the Significant Contextual Factors of Search

Elaine G. Toms[1], Joan Bartlett[2], Luanne Freund[2], Christine Dufour[1], and Steve Szigeti[2]

[1]Dalhousie University
Halifax, Nova Scotia
Canada
(902) 494-8374

[2]University of Toronto
Toronto, Ontario
Canada

etoms@dal.ca; christine.dufour@umontreal.ca; bartlett, freund, szigeti {@fis.utoronto.ca}

ABSTRACT

In this paper, we explore the contextual boundaries of search and briefly describe some of the research that we have been conducting to isolate and examine selected contextual elements.

Categories and Subject Descriptors

H.3.3 [**Information Search & Retrieval**]: Search Process

General Terms

Design, Human Factors

Keywords

Context, Task, Domain, Searching.

1. INTRODUCTION

Context is a commonly used term in everyday parlance, assumed to be so fully understood, but is rarely defined, and used so broadly as to almost be meaningless. Context "can be the whole world in relation to an utterance act" (Pinkal, 1985), or used more narrowly from the point of view of a single word, e.g., a "sequence of semantically related terms" (Stairmand, 1997).

In research, context is generally acknowledged to be an elusive construct and particularly so in studies of information retrieval (IR) (see discussion in Case (2002)). In IR, context represents a constraint on the interpretation of the situation: a searcher may implicitly (or explicitly) choose which elements of context to use for any single problem, in any single search situation, for any single search task, and indeed for any single event within that search task. But systems rarely enable this type of discrimination. The problem is in detecting which of the many contextual elements (that may be applied to the process) actually make a difference to search outcomes.

This paper identifies some of those contextual elements that need exploration and research, and describes some of the work that we have been doing over the past few years to identify key contextual elements that impact the search process.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

2. Contextual elements – a Taxonomy

The search process contains three key ingredients: individual or group who seek(s) the information for some purpose, the information resource(s) – the content – in which the information is being sought, and a system to facilitate the exchange. These elements belong to an information ecology, "a system of people, practices, values, and technologies in a particular local environment"(Nardi & O'Day, 1999) as illustrated in Figure 1.

An individual brings a host of attributes to the search process, some of which have been addressed by research (e.g., Allen, 1997; Sonnenwald, 1998). An individual has a set of cognitive abilities, from inductive reasoning to perceptual speed and accuracy. Twenty of these primary ability factors have been described by Kline (2000), although few (if any) have been assessed in studies of the search process, although selected ones have been assessedin other task environments, e.g. e-mail (Gwizdka & Chignell (2004). In addition, the individual has experiences that are socially and culturally induced, a certain knowledge base, a language, and a learning style. All of these are contextual factors that affect to some degree a user's interaction with the system and a user's interpretation of what is presented.

The content resource also has a set of characteristics that add additional contextual elements. A content source is represented in a particular medium such as video or text, is of a certain age, is presented in a certain language, is of a genre (as defined by its format, content and function), and is from a particular domain. The resource is created for a particular purpose or use by an individual or organization which also contributes to its authoritativeness and authenticity.

The transactions between individual and content take place using some form of system – a medium – from paper and papyrus to digital. All three interact under a set of conditions: within a particular environment, and situation as well as a work task that normally but not necessarily brought the individual into the interaction.

Complicating the assessment of these interactions are desired outcomes, as not all search results can be characterized in the same way (e.g., Hersh (1996)). Relevance means different things to different people in same or different situations. While novelty may be the goal in one situation, interestingness may be the goal in others, and the goal may change over multiple interactions within the same task. One situation may be 'mission critical' while another may be perceived as somewhat frivolous. Often there may be no goal at all – in this case it is the experience that counts from the individual's perspective.

From a research perspective, all of these factors and their elements potentially impact search, and this list is by no means complete. Our challenge is to identify the ones that have the most significant impact.

3. Case Studies of Contextual Factors

In an effort to examine the role of some of these contextual elements in the search process, we have been conducting a series of studies. Each is briefly described below.

3.1 Role of Domain

To understand the role of domain in the search process, we studied 48 user interactions with a modified Google interface using four information domains: consumer health, general research, shopping, and travel. In this between subject/within subject design, participants searched both researcher defined and user defined topics within each domain. Multiple types of data were collected to understand the process. From surveys, interviews, transaction logs and screen capture data, we identified distinctive differences in search processes and needs among the four domains. From these results we prototyped four search interfaces for domain-specific search systems. The conclusion we reached was that the one-size-fits-all is not the most efficient approach to searching. In addition, each domain also appeared to have other unique contextual factors that affected the search process, e.g., authenticity was perceived differently within the domains. (See Toms et al for more details of this study).

3.2 Identifying Task Processes

In a study of information search and use by bioinformatics analysts in their work places, we analyzed the process used by 20 analysts performing a specific but generic task: the functional analysis of a gene. To collect data, the 20 analysts were interviewed, and the results from an analysis of those interviews were internally and externally validated. From the data analysis, 16 sub-tasks were identified to complete the work task; each sub-task had its own data input and output and required a distinctive set of tools and/or data resource(s). Notable also about this research was the need for human decision making at each step in the process. From this research, elements of the task focused the search process. Like our earlier domain work, this research also demonstrated the value of domain specificity, the critical role of task elements, and the need for unique resources. (See Bartlett & Toms (2004) for more details; this is the PhD work of Bartlett)

3.3 Relationship between Task and Genre

In another information use environment – IBM software engineers, we are examining several contextual elements concerning situation, task and resource. In this case the tasks are typified as aspects of the job situation, and a task topology is being developed. The content resources are also being assessed with a view to creating a document topology from a genre perspective. The goal is to develop a system that integrates both topologies to augment query input and to provide more useful output. A preliminary pilot study has been completed, but this is still very much a work in progress. (See Freund & Toms (2004) for details; this is the PhD work of Freund).

3.4 Finding Information in Immersive Spaces

In a different approach to context, we examined how adolescents find information in immersive worlds. In this case context is embedded implicitly in the way the information pathways are presented. To date, a small exploratory pilot study has been conducted. Twelve adolescents were given a set of assigned search tasks with specific answers to find in an immersive world while each participant occupied the role of a personalized avatar. This is a novel information-finding scenario, one that is typically not examined in IR research. In this case the context is contained within elements of the physical space metaphor. (See Szigeti & Toms (2004) for details; this is based on data collected for the PhD work of Szigeti).

3.5 Interaction with digital videos

In a more recent study, we are exploring user interaction with digital videos. We examined how characteristics of this particular format and of the system developed to handle the videos influence the user search process. Sixteen participants performed three types of typical IR tasks using ePresence, a webcasting system developed for both live and stored videos. Using qualitative and quantitative data collected through on-line questionnaires, screen capture and interviews we found that the video constrained the usual information seeking process. For example, participants expressed difficulties in easily skimming a video to find specific pieces of information. The analysis of the data is demonstrating how different interface tools are useful for different types of search tasks. (see Dufour et al (2004) for more details).

4. SUMMARY COMMENTS

Isolating contextual factors that affect the search process and that may be gathered relatively unobtrusively is one of the leading challenges in interactive IR research. The five instances described here represent our current efforts to explore context selectively.

5. ACKNOWLEDGMENTS

This work has been funded by NSERC, and SSHRC grants, and a Heritage Canada sub-grant to Toms, an IBM Centre for Advanced Studies Student Fellowship to Freund, and a Teletoon grant-in-kind to Szigeti. Three of these studies are the PhD research of three co-authors.

6. REFERENCES

[1] Allen, B. (1997). Information needs: a person-in-situation approach. In. Vakkari, P., Savolainen, R. and Dervin, B. (eds.) Information Seeking in Context: Proceedings of an International Conference on Research in Information Needs, Seeking and Use in Different Contexts, pp. 111-122. Taylor-Graham.

[2] Bartlett, J.C. & Toms, E.G. (2004). Capturing and modelling the research process of bioinformatics experts: an integrated information behaviour and task analysis approach. *Journal of the American Society for Information Science and Technology*, accepted.

[3] Callan, J, Smeaton, A. et al. (2003). *Personalization and recommender systems in digital libraries*. Joint NSF-EC DELOS Working Group Report.

[4] Case, D.O. (2002). *Looking for Information: A survey of Research on Information Seeking, Needs, and Behavior*. Academic Press.

[5] Dufour, C., Toms. E.G., Bartlett, J., Ferenbok, J., and Baecker, R.M. (2004). Exploring User Interaction with Digital Videos. Poster presentation at Graphics Interface 2004, London (Ontario).

[6] Freund, L & Toms, E.G. (2004). Modeling the information practices of software consultants in a work and task framework . *Submitted*

[7] Gwizdka, J, Chignell. M.H. (2004). Individual Differences and Task-based User Interface Evaluation: A Case Study of Pending Tasks in Email. Interacting with Computers. Elsevier Science. 16 (4) forthcoming.

[8] Hersh, W., & Pentecost, J. (1996). A task-oriented approach to information retrieval evaluation. Journal of the American Society for Information Science, 47(1), 50-56.

[9] Kline, P. (2000). *The New Psychometrics*. London: Routledge.

[10] Pinkal, M. (1985). Kontextabhangigkeit, vagheit, mehrdeutigkeit. Translated and cited in: Quasthoff, U.M. (1994). Context In *Encyclopedia of Language and Linguistics*, pp. 730-737. Pergamon Press.

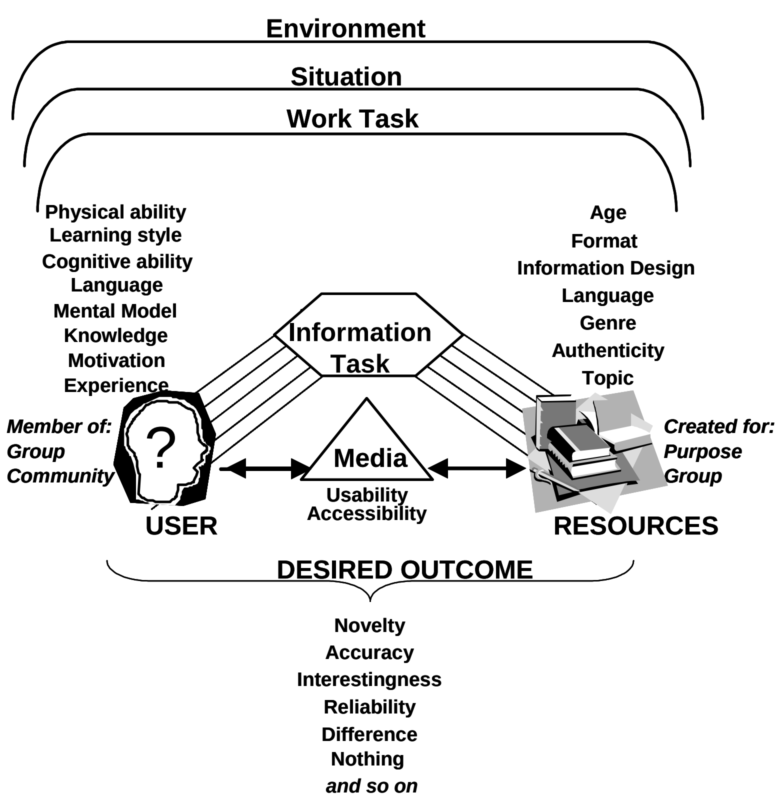
[11] Sonnenwald, D. H. (1998). Evolving perspective of human information behaviour: contexts, situations, social networks and information horizons. In T. D. Wilson & D. K. Allen (Eds.), Exploring the Contexts of Information Behaviour: Proceedings of the Second International Conference on Research in Information Needs, Seeking and Use in Different Contexts (pp. 176-190). London: Taylor Graham.

[12] Stairmand, M.A. (1997). Textual context analysis for information retrieval. In SIGIR '97, pp. 140-147.

[13] Szigeti, S. & Toms, E.G. (2004). Finding information in immersive worlds (*in production*)

[14] Toms, E.G., Kopak, R. Freund, L. & Bartlett, J. (2003). The effect of task domain on search. In *Proceedings of CASCON 2003, Markham, Ontario, Canada, 6-9 October 2003*, pp. 1-9.

Figure 1. Contextual Factors that affect Search. (Figure is adapted from Callan et al; original was Toms' contribution to the working group report)



A Context-sensitive Information System for Mobile Users

Ayse Göker[1], Hans I Myrhaug[2],
Murat Yakici[1], Ralf Bierig[1]
[1]School of Computing, The Robert Gordon University
St. Andrew Street, Aberdeen, Scotland, UK
[2]SINTEF ICT, N-7465 Trondheim, Norway
{asg, my, rb}@comp.rgu.ac.uk; hansim@sintef.no

1. INTRODUCTION

The motivation for this work is to support the information needs of people with handheld wireless devices. The information needs change more rapidly as the people encounter new situations and environments, and hence it is likely to change with the changes of the user's contexts. From the user's perspective, the thrust is about having access to and receiving relevant information in the situation. Seen from the information retrieval perspective, it is about populating people's pockets and the surroundings with search facilities engines that operate on ambient and distributed content repositories.

At present, due to slow and expensive network connections, digital content for handheld devices is not hugely abundant. This is, however, already changing as we can observe the trend amongst telecom operators trying to channel content more than just providing network access. Digital content is, therefore, fully on its way into handheld and wireless networked devices, as was the case for personal computers and the Web in the early 1990s.

In short, an increasing amount of information, services and applications are becoming available to a mobile user, but it is hard to get what you need when you need it, and users have to wade through all the information, irrespective of their location or information needs.

2. DESCRIPTION OF DEMONSTRATION

As part of our vision about users, their interaction with the surroundings and their information needs, this demonstration shows the hardware and software we have developed for this purpose.

We have developed and implemented a system comprising general context-aware technology that is proposed as a solution with a unifying framework for exploiting user contexts in ambient computing. The overall architecture includes three cornerstones: a specifically developed tag (context tag), a content service provider, and the mobile user (with a mobile device), see Figure 1. The system integrates the developed tag technology with information from content service providers in order to deliver personalised, context-sensitive information wirelessly to the handheld device. The tag is a miniaturised computer, more than just an RFID tag. It can hold 128 MB of content and is Bluetooth enabled. It can also be WLAN enabled, and can be updated remotely via Ethernet. The demonstration will focus on the interaction between the tag, the mobile device, and the content as provided by a content service provider and how this relates to meeting mobile users' information needs.

We will bring several of our tags and demonstrate the working system which participants will be able to access using a mobile device. If they own a device such as a Sony Ericsson P900, they will be able to access using their own hardware, but otherwise a few mobile handheld devices will also be made available for those without their own.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

Figure 1: Tag and mobile device

The demonstration involves a range of information:

- Travel guide content (Lonely Planet Publications) which can be browsed or searched for,
- Conference related information and business cards via the tag,
- Short news headline, event, or advert pop-up.

Figure 2: Screenshots of system in operation on handheld device

The content on the tags can be pushed to or pulled from the mobile device and presented on the screen - it can be delivered remotely from a content service provider through the wireless network infrastructure, or via the tag.

Users will also be able to see the types of content when in a relevant information zone. In an outdoor scenario, for example, as they approach the vicinity of the small and wireless tags a traveler can access content as he/she walks around a city. In an indoor case, this would be when close to objects at a museum, or in meeting rooms, for instance. The result is that relevant information can be provided to mobile users through more automatically captured information about the situation and by means of context-aware ambient technology. The retrieved content is shaped by the users' contexts. A context is, in general, constructed through the user profiles (typically local to a device) and the wireless tags that can be embedded in the surroundings.

Search engines are deployed on both the handheld device and tags. The screenshots in Figure 2 show how once a user logs in to the system (picture 1), they can key in a search (picture 2) and see an ordered list of content items (picture 3). Clicking on a specific item would then display further details for that item (picture 4). It is also possible to browse content via categories. The last picture (picture 5) shows a popup item with a news headline at the top and an image advert. The installation of search engine and content can be done from the surroundings.

3. SYSTEM DESCRIPTION

The system integrates our own tag technology with information from content service providers covering both general travel guides and local information, as described above. The overall system architecture consists of three main cornerstones: the content service provider; the mobile user; and the context tag (Figure 3).

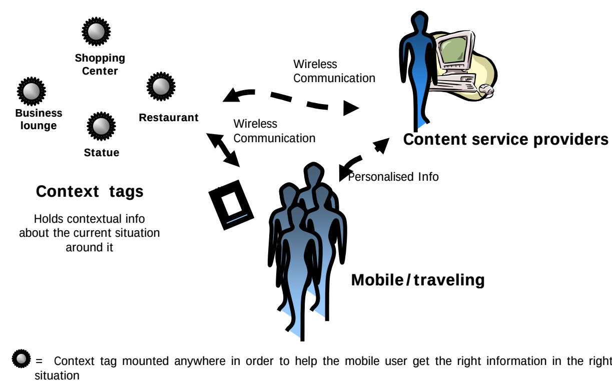


Figure 3: Overview of system architecture

The diagram illustrates the flexible ways in which users can access information. Content service providers may provide online information directly to a user (usually at a significant cost to the mobile user) or also via tags deployed in various strategic places thus creating an information zone. Information can be uploaded and maintained remotely (by content service provider or building owners, for example) but be accessible locally to the user who is in that environment and situation. For example, in the context of a conference event, the program, schedule and announcements can be communicated in this way. Discussion papers could be uploaded on tags at relevant points, and be accessible to users at relevant places.

The main components of the system consist of:

Content bases: Ambient and distributed content access is achieved via the content bases – a technology created within the project. The content can either be stored locally, or referenced by a URL. The content bases store content items. This type of information object holds meta-information about the actual physical content itself. The physical content can also be stored in the content base when the application stores a content item. This enables the application in the handheld device to be aware of and to access much more content than it has the capacity to store. In the demonstration there is a set of content items from a travel guide and news. It is relevant for both tourists and business travelers within a city. These items are structured with XML and are accessible at certain distributed locations in the surroundings. The content is indexed both on the handheld device and on the context tags.

User context middleware: This enables the system to deliver context-sensitive and personalized information. Information about the user and the user's environment is structured in a user context model. It contains information needed for a particular scenario and is attached to a specific location in the mobile environment. It also contains more individual information such user interest profiles (predominantly acquired automatically). When a user enters the vicinity of a tag, the user context is automatically enriched by aspects of the current surroundings. Moreover, contexts stored in the context middleware can also provide links to content items and content, thereby providing additional flexibility for retrieving/highlighting content.

Information search: Context-aware search (based on a probabilistic model) is performed to help retrieve relevant information objects. The information retrieval component is proactive and suggests results to the user. User context, at present, is used for a form of query expansion to improve precision measures.

Information zones: In order for the content service providers to maintain information on the tags more easily, it is possible to define information zones. An information zone object groups several context tags together and ensures that the same content appears on all tags at the same time.

Mobile clients on the handhelds: These can be downloaded when a user is in the vicinity of a tag. It provides information search and distribution mechanisms.

Context tags in the surroundings: The tags mounted at various places in the surroundings provide an information channel for the user. A tag consists of hardware and software. It can detect the proximity ('physical closeness') of handheld devices. These proximities can be configured for each tag to suit the situation or application needs. It is possible to channel the automatic distribution of content based upon the proximity.

4. RELATED WORK AND DISCUSSION

Contextual information provides an important basis for identifying and understanding users' information needs. Cool and Spink in a special issue on Context in Information Retrival [1] provide an overview of the different levels in which context for information retrieval interest exists. They refer to information environment level (e.g. Taylor [2]), information seeking level (e.g. Belkin [3]), information retrieval interaction level (this refers to user-system interactions but from a cognitive perspective can be said to relate to Ingwersen's cognitive communication model [4]), and the query level (e.g. as also discussed within TREC lately).

These categories are related and overlap. To this extent, the work described here has aspects in each of the four categories, but the first three in particular. The query level parts are not so much based on a linguistic analysis of the query but a case of augmenting or expanding it with contextual information.

Others have viewed context-aware retrieval more as a way of filtering results from normal retrieval techniques [5]. Related work can also be found in the fields of ubiquitous and context-aware computing. Dey et al, [6] in a special issue on Situated Interaction and Context-aware computing provide an overview. The focus from this perspective, however, has tended to be on location-based approaches and device contexts. Examples of these can also be found in few applications for tourists. Wider perspective of context has been discussed in some forums e.g. [7].

More specifically, however, related previous work involved the development a context learner for a probabilistic information retrieval system [3] in a traditional environment with a bibliographic search system. This was based on observations (within that environment) that users will tend to repeat searches or conduct a series of closely related searches over a period. Whilst each search must be regarded as representing a different information need they could be said to occur within a particular context. At present, the more general user context model [9] has a wider range of aspects which have been enriched as a result of the earlier work.

In summary, we have presented an overall architecture for enabling context-sensitive information systems which can take on board changes in both location and user preferences to achieve an efficient and extensible platform for information provision. Our system provides a method of achieving access to far greater amounts of information than normally manageable on a handheld mobile device while using context-sensitivity to make prevent problems of user overload.

5. REFERENCES

[1] Cool, C., and Spink:, A. Special issue on context in information retrieval. Information Processing and Management 5, 38 (2002).

[2] Taylor, R.S., Information use environments. In 'Progress in communication sciences'. Eds Dervin B., Voight, M. 217-255. Ablex: Norwood, NJ., 1991.

[3] Belkin, N.J., Anomalous States of Knowledge as a basis for information retrieval. Canadian Journal of Information Science 5, 133-143, 1980.

[4] Ingwersen, P. Information Retrieval Interaction. London: Taylor Graham, 1992 (0-947568-54-9).

[5] Brown, P.J., and Jones, G. Context-Aware Retrieval: Exploring a new environment for information retrieval and information filtering. Personal Ubiquitous Computing 5(4): 253-263, 2001.

[6] Dey, A., Kortuem, G., Morse, D., and Schmidt, A., Special Issue on Situated Interaction and Context-Aware Computing, Journal of Personal and Ubiquitous Computing, 5, 1, (2001).

[7] Gross,T., and Specht, M.Workshop on User Modeling for Context-aware Applications. User Modeling Conference UM2001, 2001.

[8] Goker, A. Context learning in Okapi. Journal of Documentation (53), 80-83, 1997.

[9] Myrhaug, H. I., and Goker, A. AmbieSense - interactive information channels in the surroundings of the mobile user. 10th International Conference on Human-Computer Interaction (HCI International 2003)}, Crete, Greece, (2003).

Context in Active Groups

Stuart Watt
School of Computing, The Robert Gordon University
St. Andrew Street
Aberdeen, Scotland, UK
+44 1224 262723
S.N.K.Watt@rgu.ac.uk

1. INTRODUCTION

We have been working on email information retrieval systems for several years now, and Active Groups represents the latest in a range of systems we have developed to support collaborative information retrieval in communities. In this paper, we'll discuss how the needs of this particular problem and domain have influenced the design of a system that uses context extensively.

First, it is worth discussing the domain and problems that Active Groups was designed to address. It was developed to support a multinational oil and gas company whose engineers routinely used email discussion lists to share problems, ideas, and experiences around the world. These discussion lists were implemented by Microsoft's Exchange, which provided a very weak — and very slow — search facility. The archives of these discussion lists provided an extensive resource of valuable knowledge and experience, although these were almost entirely inaccessible due to the inadequacy of the search tools.

Our initial solution to these problems — drawn from our earlier work on the Virtual Participant [1] and Sentinel [2] — was a purely email-based interface. The system participated in the discussion lists in exactly the same way as human engineers, and if anyone raised an issue that had been discussed earlier, or in another discussion forum, the system would post a summary message of the earlier discussion. This message could then contain follow-up questions, which people could use to reply to the system if the initial message did not answer their question.

Proactive retrieval is very different from the interaction of conventional search engines. Not every action on the part of the user needs or deserves a response, but it is important that when a response is given, it really is relevant to the needs of the user at that point in time. To achieve this, we exploit the genres of email messages — genre reflects the social context of discussion, and has a layout component as well as a content component that can often help determine the social intent behind a particular message.

2. DESIGN OF ACTIVE GROUPS

In our original Virtual Participant system, there was a base of messages which was hand-crafted, based on a corpus of discussion messages from previous years. This was laborious to set up, and not sustainable for larger discussion forums. In our later systems, with tens of thousands of messages rather than a few hundred, we set out to index the discussion automatically as far as possible, recognising that the same quality of indexing would not be possible.

The architecture of Active Groups is shown in Figure 1. This clearly shows the two modes of interaction with the system, through email and through the web interface.

At the core of Active Groups is a fairly simple information retrieval engine, based on BM25 [3], and now published as an open source Perl module (Search::FreeText, available through CPAN — www.cpan.org). This is complemented by a range of filters that clean the messages and provide additional indexes through a semantic network. The system is completed by email

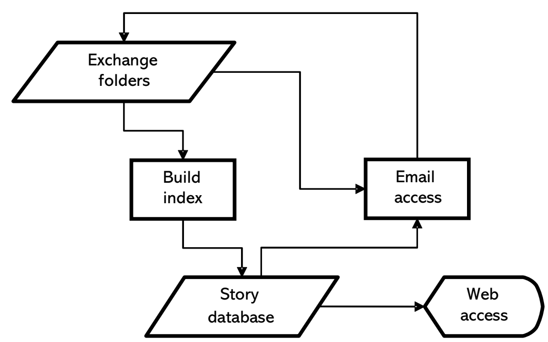


Figure 1. Architecture for Active Groups

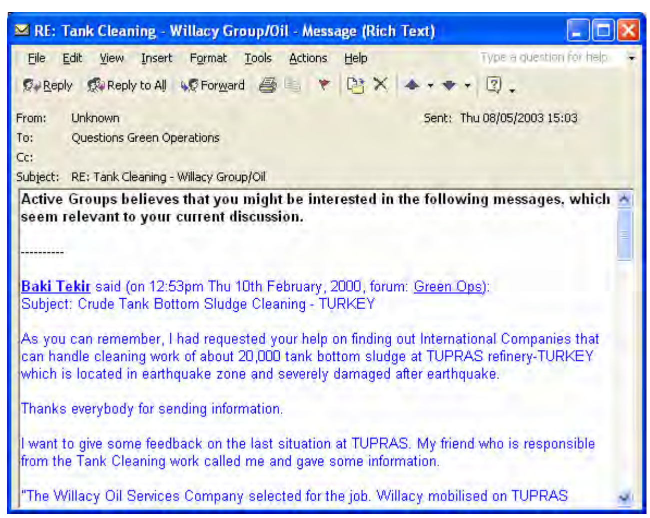


Figure 2. An example email sent by Active Groups

SIGIR'04 July 25-29, 2004, Sheffield, UK.

and web interfaces that use these indexes to provide relevant information to users as and when they need it.

The primary and initial interface to Active Groups is through an email client such as Microsoft's Outlook. Any client will do, as the program interacts with the server rather than the client. A typical view is shown in Figure 2 below. There is no search box in this view — Active Groups monitors the discussion through the email server, and when a posting is available which is relevant to the current topic of discussion, it is posted publicly to the forum. Individual users can then pursue any further interests in the matter by clicking a link, which will jump them into the right place in the web interface.

Although not part of our original design, we have added a web interface— a screen snapshot is shown in Figure 3. We originally included it as a debugging tool — at the time even follow-up enquiries were handled through email. At the client's request, it has subsequently become a core part of the system.

3. CONTEXT IN ACTIVE GROUPS

Active Groups is a proactive information retrieval system, and needs to make postings to the discussion groups only when they are useful. Postings are not required after each and very message — indeed, this would be socially unacceptable. Active Groups needs to determine when — and if — a response is appropriate. This is strongly dependent on the context of discussion — making the same response a few days later would not be useful.

A simple solution is to use a threshold function on the results of the search. This works reasonably well as a starting point — Active Groups uses the 'long query' form of the BM25 heuristic as this starting point. There is a sound reason for using a probabilistic model for information retrieval in Active Groups. It provides a sound basis for a threshold-based assessment of whether or not to post a response — in effect, if the probability is sufficiently high, it is worth posting.

But Active Groups does not use BM25 alone. First, it indexes questions designed to elicit opinions and suggestions as well as whole messages, and uses these questions when they are available. Second, Active Groups builds a semantic network, linking the messages and associating them with the message

Figure 3. Web interface to Active Groups

sender's, the topics they discuss, and so on. The method used in Active Groups is a simplified version of the 'point linking' in [4], which is intended to produce a properly indexed base of texts for navigation in a learning environment. This structure was the inspiration for much of the design of the system. Unlike point linking, which has never been implemented in practice, the 'object linking' in Active Groups is not yet of a standard where it can identify the point in a message — this is the goal of future research work. Even so, the semantic network representation does make it easy to identify experts in particular areas, and to find people who form connections between the different discussion forums. The combination of traditional information retrieval with a semantic network representation has proven a promising approach to integrating the different information sources needed for context reasoning.

Another important reason is that it is easier to combine several different sources of evidence if they are all providing probabilities of relevance. In the case of Active Groups, the relevance probabilities of the query results are combined with a number of other aspects, all of which assess the context of each message. These other aspects include the following:

Discussion thread context. This measures the role of the message in a discussion. The first message in a thread typically asks for opinions and suggestions, through questions of the form "Does anybody have experience of X?". In practice, the form is usually more complicated than this — it is a distinct message genre which contains three parts, a scene-setting paragraph, a bulleted technical specification, and a set of specific questions. Active Groups uses the questions as well as the whole message to index the message properly. Follow-up messages from other people often either help orient or focus the discussion, or provide opinions or suggestions. This interplay of messages is essential to deciding whether or not to respond to a particular message.

Technical context. Many questions involve a particular technical configuration, which forms the background environment for a problem to be solved. In the future, we expect Active Groups to use a model of the configuration to provide recognition of context; currently, we use an augmented stemmer that can manage the range of measures used. This is often enough; for example, "1 3/4 in" is the same measure as "1.75"", but is entirely different to "1.75 degrees F" — this can help to match technical specifications more accurately.

Discussion space context. The discussion space itself is also part of the context, and shape whether or not a response will be posted by Active Groups. A recent message from the same forum will not be posted as a response, as it will have been seen by the participants anyway. A message from a different forum may well be posted.

Temporal context. Time is another important part of context in Active Groups. It is important to respond to a message at the right time. When posting a response, Active Groups will wait an appropriate time before the message is actually sent. This time is usually a few hours after the previous message, to give other participants a chance to respond.

Although Active Groups does begin to address some aspects of context [5], there are other areas where its support is less well developed. For example, the social role of the sender also should play an important part in the processing of a message. Some people are moderators, and tend to forward messages rather than answer them, or to pass on the contact details of known experts in particular fields. Active Groups does not yet manage any evidence about the roles people adopt in its processing of messages.

Email contains other context cues which can be a real problem. Many people embed quoted chunks of an original message, and respond to them a section at a time. Only a small fraction of a message may be new. Also, people often embed greetings, and especially signatures, which say little about this particular message, but a lot — indirectly — about the sender. Active Groups has a large set of heuristic rules which it uses to 'clean up' the messages, so that information retrieval can be focused on the distinctive content of each particular message. Although this additional information may be useful when, for example, the phone number of an expert in a given field is needed, its use in information retrieval is limited, and there are often existing databases and other sources that provide the same information.

Active Groups has yet to be evaluated formally, and there are good reasons for this — evaluating large-scale group systems which use context — especially social context — introduces many problems not found in more traditional information retrieval systems. Many users are vicarious, and only use the system by observing other people's behaviour. Relevance is strongly linked to the time of response, so independent relevance judgements can be hard to obtain and assess. Given these issues, more evidence about the strengths and weaknesses of systems like Active Groups is needed to guide further work in this area, and the lack of a formal user study is a limitation of our work to date. A user study has been planned, and is expected to be conducted in the summer of 2004.

4. DISCUSSION

Proactive email-based information retrieval has proven an intriguing area of research. Although technologies from traditional information retrieval have helped make Active Groups into an effective solution to the knowledge management needs of the client, these technologies needed to be combined with other solutions in an imaginative manner to achieve this success.

To deliver relevant information effectively in the proactive mode, Active Groups uses a range of different technologies. First of all, it uses a standard but effective information retrieval system, complemented by a semantic network and a wide range of heuristics to support use of different kinds of context.

Finally, there is a real need for significant evaluation work in this field. Although we have evaluated our earlier systems, evaluation of Active Groups is still ongoing, although it was well-received and is actively used by our client. Our earlier evaluation work on previous systems indicated that there may be a tendency for people to over-estimate relevance in the proactive mode. However, the essential dependence on context makes controlled studies of the effectiveness of systems like this exceedingly hard to design. We feel that although Active Groups has made considerable progress towards managing and interpreting context, a concerted and collaborative effort towards an effective evaluation programme would significantly strengthen work in this area. Email discussion groups seem to make an ideal test case.

5. REFERENCES

[1] S. J. Masterton and S. N. K. Watt, "Oracles, bards, and village gossips, or, social roles and meta knowledge management," *Journal of Information Systems Frontiers*, vol. 2, 2000.

[2] E. Moreale and S. N. K. Watt, "An agent-based approach to mailing list knowledge management," presented at AAAI Spring Symposium on Agent-Mediated Knowledge Management (AMKM-03), Stanford, CA., 2003.

[3] K. Sparck Jones, S. Robertson, and S. Walker, "A probabilistic model of information retrieval: development and comparative experiments," *Information Processing and Management*, vol. 36, pp. 809-840, 2000.

[4] C. Cleary and R. Bareiss, "Practical methods for automatically generating typed links," presented at Seventh ACM Conference on Hypertext (Hypertext '96), Bethesda, MD, 1996.

[5] C. Cool and A. Spink, "Special issue on context in information retrieval," *Information Processing & Management Journal*, vol. 5, 2002.

Building a Test Collection for Investigating Contextual Information Retrieval

Diane Kelly
University of North Carolina
100 Manning Hall
Chapel Hill, NC 27599-3360 USA
+00 919.962.8071
kelly@ils.unc.edu

The importance of building a test collection which can be used to explore contextual information retrieval (CIR) is great. Test collections such as those generated for TREC have proven invaluable to both TREC participants and others in the IR community. The creation of sharable test collections facilitates discovery and allows for more rapid progress since building a good test collection is such a difficult, laborious, and time-consuming task. Standard test collections also allow for multiple modes of inquiry including those that involve the comparison of various IR techniques, examination of alternative hypotheses and replication of previous findings.

The design and construction of a test collection for CIR introduces numerous challenges that are not present during the construction of more traditional IR test collections. For instance, a test collection for CIR should contain more than just documents, topics and relevance judgments. In addition to this information, such a collection should also contain information about users, their information needs and goals, their information-seeking context and their behaviors within this context. Collecting this type of information necessarily implies the construction and evaluation of new tools for data collection since these types of information are neither obvious nor explicit in the interaction. Furthermore, these tools should produce valid, reliable and usable data. Constructing a collection with tools that do not meet these fundamental criteria does little to further research in CIR and instead, is likely to yield erroneous results. For such a collection to have the greatest ecological validity, and thus generalizability, it should be constructed within a natural use environment with real tasks and topics, rather than within a laboratory setting, with artificial tasks and topics. Moreover, such a collection should be representative of activities that occur over an extended period of time to allow for the investigation and modeling of more complex types of interaction such as successive searching and long- and short-term needs. Finally, because the notion of context is so complex, building a test collection for CIR necessarily implies that some discussion about what context is, what elements of context matter in IR, and how these elements can be measured explicitly has to occur.

In this presentation, I will discuss the results of a naturalistic, longitudinal study that was designed to collect information about users' information-seeking activities, contexts and behaviors in a natural setting over an extended period of time. One of the outcomes of this study was the development of a method for collecting data about users' information-seeking behaviors in natural search environments, with user-defined tasks and topics. Another outcome of this study was the development and evaluation of techniques for measuring aspects of users' information-seeking context. In this study, the entirety of users' on- and off-line interactions with their computers were unobtrusively monitored and recorded using both client- and proxy-side logging software. This included applications used, URLs visited, start, finished and elapsed times for all interaction, operating system commands, and all keystrokes such as queries and word processing text. In addition, a copy of every document requested by each user was saved on a local server. Throughout the study, users identified the various tasks and topics about which they were seeking information, and classified the documents that they viewed according to these tasks and topics. At weekly intervals, users updated each context measure, which included things such as their familiarity with a topic, and judged the usefulness of the documents that they viewed during the previous week. At the close of the study, users provided qualitative feedback about the study method including the various instruments and procedures used to measure context.

Although this study was not necessarily concerned with building a test collection for CIR, the resulting data makes for a potentially useful collection for exploring CIR. Furthermore, the study method can be viewed as a pilot for a larger data collection effort. In this presentation, I will discuss the method used to collect this data, including my attempt at measuring context, the results of the data collection effort, and the lessons learned from this effort. My purposes in sharing these results are to provide the audience with an overview of (1) how much effort is involved in planning and assembling such a collection; (2) how much data can be potentially collected; and (3) what are some potentially fruitful measures of information-seeking context. I will conclude by suggesting future directions for the construction of a new test collection that can be used to investigate CIR, and issues that need to be addressed before such an effort can commence.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

End users in the context of XML documents – setting up an interactive track at INEX

Birger Larsen
Department of Information Studies
Royal School of LIS,
Copenhagen, Denmark
blar@db.dk

Anastasios Tombros
Department of Computer Science
Queen Mary, University of London
United Kingdom
tassos@dcs.qmul.ac.uk

Saadia Malik
Fak. 5/IIS, Information Systems
University of Duisburg-Essen
Duisburg, Germany
malik@is.informatik.uni-duisburg.de

ABSTRACT
The widespread use of XML has stimulated research in developing appropriate methods for searching and browsing XML documents. However, relatively little research has been carried out that studies user interaction with IR systems that take advantage of the additional features offered by XML documents. In this paper, we describe the efforts to establish an interactive track at INEX and discuss its main motivation and aims.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval]**: Information Search and Retrieval – *search process*; H.3.7 [**Information Storage and Retrieval]**: Digital Libraries – *user issues*.

General Terms
Experimentation, Human Factors.

Keywords
XML documents; User Contexts; INEX; Interactive Track.

1. INTRODUCTION
The widespread use of XML and other high level mark-up languages has stimulated research in developing appropriate methods for searching and browsing XML documents. From the point of Information Retrieval (IR), highly structured XML documents are attractive because the mark up makes it possible to identify separate parts of the documents easily rather than to view them as a uniform bag of words.

The Initiative for the Evaluation of XML Retrieval (INEX) was initiated in 2002 as a large scale international effort to improve the efficiency of research in content-oriented XML retrieval, and to promote evaluation procedures that can assess the effectiveness of proposed methods [2; 3]. Evaluating the effectiveness of XML retrieval systems requires a test collection where the relevance assessments are provided according to a relevance criterion, which takes into account the imposed structural aspects. Such a test collection has been built as a result of INEX2002 and INEX 2003. The test collection consists of:

a) A corpus of 12,107 full text scientific documents from the IEEE Computer Society's 18 journals formatted in XML.

b) A set of topics each expressing an information need. The topics are developed by INEX participants.

c) Relevance assessments provided by the topic authors. Relevance assessments are non-binary and are provided on two dimensions (*coverage* and *topicality*).

In INEX 2002 and INEX 2003 two types of topics were considered: Content and Structure (CAS), which contain explicit references to the XML structure, and Content Only (CO), which disregard the XML structure in the queries.

Research under the INEX initiative has started to shed light in aspects of effectiveness of XML retrieval. However, relatively little research has been carried out to study user interaction with IR systems that take advantage of the additional features offered by XML documents and so little is known about how users behave in the context of such IR systems. One exception is the work done by Reid and associates, who studied end user interaction with a small test collection of Shakespeare's plays formatted in XML [4].

In order to learn more about end user interaction with XML-based IR systems an interactive track is being set up as part of INEX 2004. In this paper we describe this effort and discuss its main motivation and aims.

2. INTERACTIVE TRACK MOTIVATION
Issues relating to interactive IR have been extensively investigated in the last decade. A major advance in research has been made by co-ordinated efforts in the interactive track at TREC. These efforts have been in the context of unstructured documents (e.g. news articles) or in the context of the loosely-defined structure encountered in web pages. XML documents, on the other hand, define a different context, by offering the possibility of navigating within the structure of a single document, or of following links to another document.

This context is different to the one encountered in the conventional case of unstructured documents, and has provided the main motivation for the establishment of an interactive track

SIGIR'04 July 25-29, 2004, Sheffield, UK.

at INEX. The main aims for the interactive track are twofold. First, to investigate the behaviour of users when interacting with components of XML documents, and secondly to investigate and develop approaches for XML retrieval which are effective in user-based environments.

In the first year, we plan to address the first issue: to investigate the behaviour of searchers when presented with components of XML documents that have a high probability of being relevant (as estimated by an XML-based IR system). Presently, metrics that are used for the evaluation of system effectiveness in the INEX ad hoc track are based on certain assumptions of user behaviour [6]. These metrics attempt to quantify the effectiveness of IR systems at pointing searchers to specific relevant portions (or elements) of documents. Some of the assumptions behind the metrics include that users would browse through retrieved elements in a linear order, that they would “jump” with a given probability *p* from one element to another within the same document’s structure, that they would not make use of links to another document, etc. These assumptions have not been formally investigated in the context of XML retrieval; their investigation forms the primary aim for the first year of the interactive track. Through appropriate logging software (see section 3) we aim to collect data which will allow us to examine user behaviour in this context.

In addition, the track will aim to investigate the effect that task type has on search behaviour in the context of XML documents. This forms a second type of context that we aim to take into account. The effect that the context determined by task type has on the behaviour of online searchers has been demonstrated in a number of studies [e.g. 7]. One way to categorise tasks is according to the “type” of information need they correspond to. In [7] the categorisation included background (find as much general information on a topic as possible), decision (make a decision based on the information found) and many-items tasks (compile a list of items related to the information need) types. It was shown that different task types promote the use of different criteria when assessing the relevance of web pages. It is likely that a similar effect, in terms of user behaviour within structured documents, may exist in the context of XML documents. Searchers may exhibit different browsing patterns and different navigational strategies for different task types.

In this way, the format of the track for the first year differs to that followed by e.g. the interactive track at TREC. The main difference is that a comparison between different interactive approaches is not the main focus for the first year. Instead, a more collective effort is planned, where the outcome of the studies will benefit the INEX initiative. Participating sites will have the option to develop and evaluate their own interactive approaches but this will not be required.

3. PLANNED EXPERIMENTAL SETUP

In this section we briefly outline the experimental set up for the first interactive track at INEX.

A number of the 2004 CO topics will be used in the study, along with the standard INEX collection. In order to make the topics comprehensible by other than the topic author, it is required that the INEX 2004 topics not only detail *what* is being sought for, but also *why* this is wanted, and in what *context* the information need has arisen. Thereby the INEX topics are in effect simulated work task situations as developed by Borlund [5]. Compared to the regular topics, more context on the motives and background of the topic is provided in the simulated work tasks. In this way, the test persons can better place themselves in a situation where they would be motivated to search for information related to the work tasks. The aim is to enable the test persons to formulate and reformulate their own queries as realistically as possible in the interaction with the IR system. Also, in order to examine the effect of task type, CO topics will be selected that represent different types of information needs.

The test persons, employed locally in each site, will need to identify documents which are useful for completing the requirements specified in the simulated work task. They can either identify these documents explicitly (e.g. by marking down a relevance score for each document) or implicitly (e.g. by saving or bookmarking useful documents). A time limit will be set for each simulated work task.

For the first year of the track, all participating sites will use the same system which will be made available. The system will provide a basic functionality which will be agreed upon. The present system under consideration is HyRex[1] with a web-based interface with a range of visualisation features. Additional systems may be employed locally if a participating site wishes to develop and compare their own interactive approach to the official “baseline”.

Analysis of the collected data will be required in order to extract conclusions from the studies. The collected data will comprise (as minimum) of questionnaires completed by the test persons and the logging of searcher interaction with the system. The logged data will consist of the queries issued, the components returned by the system, the components actually viewed and the order in which they are viewed, relevance assessments of these, any browsing behaviour, as well as time stamps for each act of interaction between the test person and the system. Participating sites will also be given the opportunity to employ site-specific data collection methodologies (e.g. think-alouds, desktop monitoring software, etc.).

The results of the track will be reported in the next INEX workshop, to be held in December 2004 in Schloss Dagstuhl, Germany.

4. CONCLUSION AND FUTURE DIRECTIONS

The interactive track at INEX is an effort to systematically investigate issues relating to user behavior and effectiveness of interactive approaches in the context of XML retrieval. The track has so far received interest from 23 out of the 51 participating sites in INEX 2004. For the first year we expect that few of the participants will be evaluating their own interactive approaches. Instead, the anticipated benefits in this first effort are twofold: first, the instigation of an interest in the research community for interactive aspects of XML retrieval and second, the collection of sufficient data for the investigation of user behavior – this data will also be used for the validation of assumptions related to evaluation metrics. This is expected to be a significant contribution of the track to the wider INEX initiative.

[1] See http://www.is.informatik.uni-duisburg.de/projects/hyrex/

In the coming years, the aims of the track will shift towards the evaluation of the utility of various interactive approaches for XML retrieval. The context of the evaluation will also shift towards that of digital libraries, taking into account different users, document collections, systems and uses. This shift will also be facilitated by the expansion of the INEX initiative, with the inclusion of more tracks in INEX 2004.

5. ACKNOWLEDGMENTS

This work is done as part of the DELOS Network of Excellence on Digital Libraries. The authors wish to thank the participants of the 2nd INEX workshop in Schloss Dagstuhl, December 2003, for useful discussions on the format of the track.

6. REFERENCES

[1] Chiaramella, Y. (2001): Information retrieval and structured documents. In: *Lectures on information retrieval : third European summer-school, ESSIR 2000, Varenna, Italy: Revised Lectures*. Berlin: Springer, pp 286-309.

[2] Fuhr, N., Gövert, N., Kazai, G. and Lalmas, M. (2002): INEX : initiative for the evaluation of XML retrieval. In: *ACM SIGIR'2002 workshop on XML and information retrieval*, pp 62-70.

[3] Gövert, N. and Kazai, G. (2003): Overview of the initiative for the evaluation of XML retrieval (INEX) 2002. In: *Proceedings of the 1st workshop of the initiative for the evaluation of XML retrieval (INEX)*, pp 1-17.

[4] Finesilver, K. and Reid, J. (2003): User Behaviour in the Context of Structured Documents. In: *Advances in Information Retrieval: 25th European Conference on IR Research, ECIR 2003*, pp 104-119.

[5] Borlund, P. (2000): *Evaluation of interactive information retrieval systems*. Åbo: Åbo Akademi University Press. vi, 276 p. (PhD dissertation).

[6] Kazai, G. (2003): Report of the INEX 2003 metrics working group. In: *Proceedings of the 2nd workshop of the initiative for the evaluation of XML retrieval (INEX)*, pp 184-190.

[7] Tombros, A., Ruthven, I. and Jose, J (2004): Searchers criteria for assessing web pages. In: *Journal of the American Society for Information Science and Technology*. To appear.

Intelligent RSS News Aggregation Based on Semantic Contexts

Weihong Huang
Centre for Internet Computing
The University of Hull, Scarborough Campus
Scarborough, YO11 3AZ, United Kingdom
W.Huang@hull.ac.uk

David Webster
Centre for Internet Computing
The University of Hull, Scarborough Campus
Scarborough, YO11 3AZ, United Kingdom
icr03dew@cic.hull.ac.uk

ABSTRACT
Information aggregation has been considered as one of the most important methods in personal or organisational information and knowledge management. To improve the intelligence of information aggregation and consequently reduce the workload of manual filtering, this paper presents a context-aware information servicing approach based on semantic contexts. An RSS (RDF Site Summary / Really Simple Syndication) news aggregation prototype using intelligent agents is developed. The implementation not only shows the practicability of the approach in service, but also provides a good foundation for further extension to other applications and future developments towards fully automatic agent-assisted knowledge services.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: Information filtering; I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents; H.3.5 [**Online Information Services**]: Web-based services

General Terms
Management, Design, Human Factors

Keywords
RSS, aggregation, semantic context, agents

1. INTRODUCTION
As one of the most outstanding inventions in last decade, the WWW has become a huge information pool to global users. Traditional information retrieval services over the WWW through search engines have been quite successful in collecting information from the Web, despite of the fact that users still have to filter the information manually. With the development and deployment of XML (eXtensible Markup Language) technologies, more and more semi-structured datat [1] and structured data [2] are available online and open to certain user groups. In the future, more information will be represented in semantic description formats such as XTM (XML TopicMaps) [15], and RDF (Resource Description Framework) [14] towards the Semantic Web [3]. Accordingly, information aggregation services will involve more software assistance such as intelligent agents [5] in the future, since agents are expected to reduce the workload of users in services and service aggregation.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

This paper presents a context-aware approach to promote the information aggregation service based on semantics. The proposed approach features a context-aware service architecture with a context description model. Agents as the service operators are to understand the resource content in semantic contexts referring to common knowledge bases (e.g. the Open Directory DMOZ [12] taxonomy).

The chosen example information packaging format in this paper is RSS (RDF Site Summary / Really Simple Syndication) [13], which is one of the most popular RDF-based applications in real use on the current Web. Another reason of choosing RSS is the diversity of its formats: RSS/0.9x and RSS/2.0 in XML, and RSS/1.0 in RDF, which is a good study case for proving the proposed context-aware approach works in different contexts (e.g. different structure, different encoding format, different levels of enclosed semantic information).

Our recent related work includes semantic annotation of multimedia resource using annotation graph [7], contextual knowledge management for e-Learning [6, 9] and context-awareness in e-Enterprise knowledge management [8].

2. WHAT IS SEMANTIC CONTEXT
The concept of *context* has been widely used in many computing areas. For example, pervasive computing [11], contextual logical reasoning [10, 1]. In this paper, we define the

[1]Typical semi-structured data are schemaless self-describing data such as personal profile including email, address, telephone numbers.
[2]Typical structured data are un-normalized relational data such as product catalogs, medical records, and stock quotes.

concept of *semantic context* as follows:

Definition: Context of an entity (i.e. an object, an event, or a process) is a collection of semantic situational information that characterizes the entity's internal features or operations and external relations under a specific situation.

Typical contextual information includes:

- general metadata of entities, such as title, author, key words, publish date, version, etc.;
- literal statements, such as free annotations of multimedia resources such as images, audio, video and presentations, etc.;
- conceptual models, such as system models, learning processes, mind maps, etc.;
- hybrids of statements and conceptual models to represents contextual knowledge;
- interlinks with other knowledge descriptions, links between description elements resided in different context descriptions that show the relations between different contexts and their reusability of the common elements and links.

To give you a direct impression of what "Semantic Context" is, we can subtract some information directly from the IRiX workshop web page on SIGIR2004, and transform it into an RSS item with three basic elements: *title*:"Information Retrieval in Context", *link*:"http://ir.dcs.gla.ac.uk/context/", and *description*:"There is a growing realisation that relevant information will be increasingly accessible ...". It looks very straight-forward to humans, but when agents come to this RSS item, they will have some questions in machine understanding. Generic questions could be like these: is this a news link? What is the subject of this item? What is the RSS version of this item? Go deep into the content, when agents come to the term "context", as we mentioned above, there are more than one areas related to context in computing. So here agents will need more category descriptions of the term, for instance, some ACM category descriptors such as *Information Search and Retrieval*, *Online Information Services*, which could help agents locating the semantic context of this term. Furthermore, when another conjunct term "information retrieval" in the same statement comes to the process, the previously located semantic context could be verified. As the RSS format allows flexible extension in its "description" field, all related information could be included if necessary. In this way, all situational descriptions in relation to the IRiX04 workshop in RSS form its semantic context, which will be understandable to agents referring to common sense knowledge bases (such as ACM category classifications, generic online dictionaries, e.g. DMOZ taxonomy).

3. CONTEXT-BASED RSS NEWS AGGREGATION

3.1 Context-Aware Service Architecture

To enable semantic-based information services (not just for RSS news aggregation), we present a context-aware service architecture (as shown in Figure 1). This architecture is not only designed for agents, but also open to other non-agent interfaces such as web services.

The most important contribution of this architecture is introducing a context description model with distinguished service descriptions from content descriptions. This means the same content could be accessible to various semantic-aware services, for example, the same book review comment could be used on eBay for audition and on Amazon for selling recommendation as well.

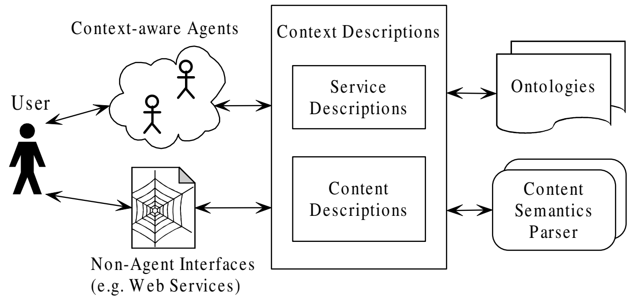


Figure 1: Context description architecture

3.2 Context Description Model

Figure 2 illustrates the context description model in details with an example of real RSS description model in development.

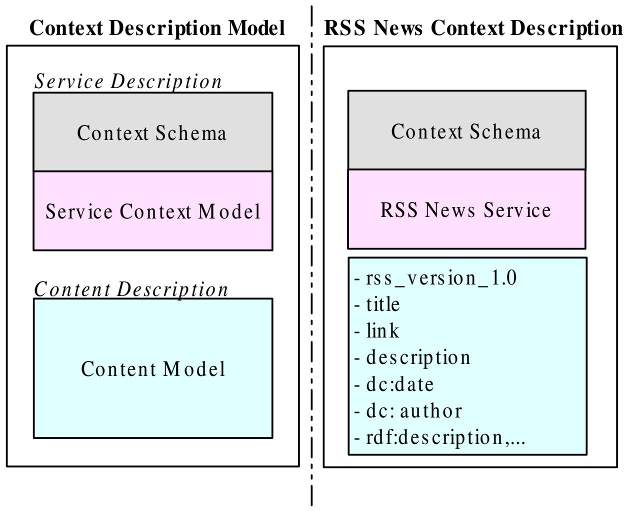


Figure 2: RSS context description model

The context description model consists of two parts: service description and content description. The context schema in the service description part contains generic information about the context model. In another word, it explains what this context model is about, where the schema is located, what the basic vocabulary is, and so on. The service context part describes the basics about the service. For instance, in RSS news aggregation service, it shows the service name (i.e. RSS news aggregation), service type (i.e. agent-based), service version (i.e. V1.0), and so on. If we say the context description part is generic and abstract, then the content description part is concrete, because it describes the real content of the resource object (e.g. RSS payload).

3.3 Aggregation in Practice

When one user invokes his/her personal agents or a news aggregation program interface, a group of agents who can provide the RSS news aggregation service will be available on a list. After he/she inputs the keywords, the personal agents will choose the capable aggregation service agents to carry out the service. This process is invisible to end users, but the agents have already retrieved and understood the guidelines of RSS news aggregation service, and located the generic reference knowledge base. With the input keywords, agents will start retrieving the content (payload) of the RSS items, and verifying the relevance between the input and the content of the RSS item through the parsing process, where XML-based RSS/0.9x/2.0 items are thrown to Informa [3] and RDF-based RSS/1.0 items are thrown to Jena [4]. The semantically matched items are then to be returned back to the personal agents in a list.

Figure 3 shows an example of the results of a completed RSS item search to a number of service agents. In the case of this example search a simple keyword of *Windows* was used.

The prototype is developed in Java. The agent platform in use is called JADE, a FIPA-compliant multi-agent software framework [2]. Considering the portability issue, the system is designed to be agent platform independent. Future migration to other agent platforms are supposed to be carried out easily.

By using Java agents to perform services, this system has a better extensibility than existing stand-alone programs such as FeedReader[5] or web browser plug-ins such as NewsMonster [6]. In terms of functionality, the service we provide in the prototype is somehow similar but different from the semantic search service provided by TAP at Stanford [4]. Further extension to new services other than RSS aggregation using the same service agent could be carried out easily, for example understanding document outlines in OPML (Outline Processor Markup Language) [7]. Furthermore, interoperation between different types of agents within the generic service architecture are also possible in future development.

4. CONCLUSION AND FUTURE WORK

Aiming at improving the intelligence in information aggregation and consequently reducing the workload of manual filtering, this paper presents a context-aware approach using agents to assist RSS news aggregation by understanding the context description of services and content of resources. The proposed approach features an application-independent service architecture with a context description model. System prototype implementation shows the feasibility of this approach.

[3]RSS Library for Java, http://informa.sourceforge.net/
[4]Jena SW Framework, http://jena.sourceforge.net/
[5]FeedReader, http://www.feedreader.com/
[6]NewsMonster, http://newsmonster.org/
[7]OPML, http://www.opml.org/

Future work might involve extension of the context-aware approach to trust-aware service based on semantic context, which aims at crediting the resources, services and agents themselves in a cognitive manner.

5. REFERENCES

[1] V. Akman. *Context in Artificial Intelligence: A Fleeting Overview.* McGraw-Hill, Milano, 2002.

[2] F. Bellifemine, G. Caire, A. Poggi, and G. Rimassa. Jade - a white paper. *exp*, 3(3):6–19, September 2003.

[3] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, May 2001.

[4] R. Guha, R. McCool, and E. Miller. Semantic search. In *Proceedings of the twelfth international conference on World Wide Web(WWW03)*, pages 700–709, Budapest, Hungary, May 2003.

[5] J. Hendler. Agents and the semantic web. *IEEE Intelligent Systems*, 16(2):30–37, 2001.

[6] W. Huang and M. S. Hacid. Contextual knowledge representation, retrieval and interpretation in multimedia e-learning. In *Proceedings of the Fourth IEEE International Conference on Information Reuse and Integration*, Las Vegas, USA, October 2003.

[7] W. Huang, Y. Prie, P. A. Champin, and A. Mille. Semantic context representation of resources using annotation graph. In *Proceedings of the Eighth International Workshop on Multimedia Information Systems*, Tempe, USA, September 2002.

[8] W. Huang and T. Tao. Adding context-awareness to knowledge management in modern enterprises. In *Proceedings of the First IEEE International Conference on Intelligent Systems*, Varna, Bulgaria, June 2004.

[9] W. Huang, T. Tao, M. S. Hacid, and A. Mille. Facilitate knowledge communications in multimedia e-learning environments. In *Proceedings of the First ACM International Workshop on Multimedia Databases*, New Orleans, USA, November 2003.

[10] J. McCarthy and S. Buvac. Formalizing context(expanded notes). Technical report, Stanford University, 1998.

[11] T. P. Moran and P. Dourish. Introduction to this special issue on context-aware computing. *Human-Computer Interaction*, 16, 2001.

[12] Open Directory Project. `http://dmoz.org/`.

[13] M. Pilgrim. What is rss? Technical report, XML.com, December 2002.

[14] Resource Description Framework (RDF). `http://www.w3.org/RDF/`.

[15] XML TopicMaps(XTM). `http://www.topicmaps.org/xtm/`.

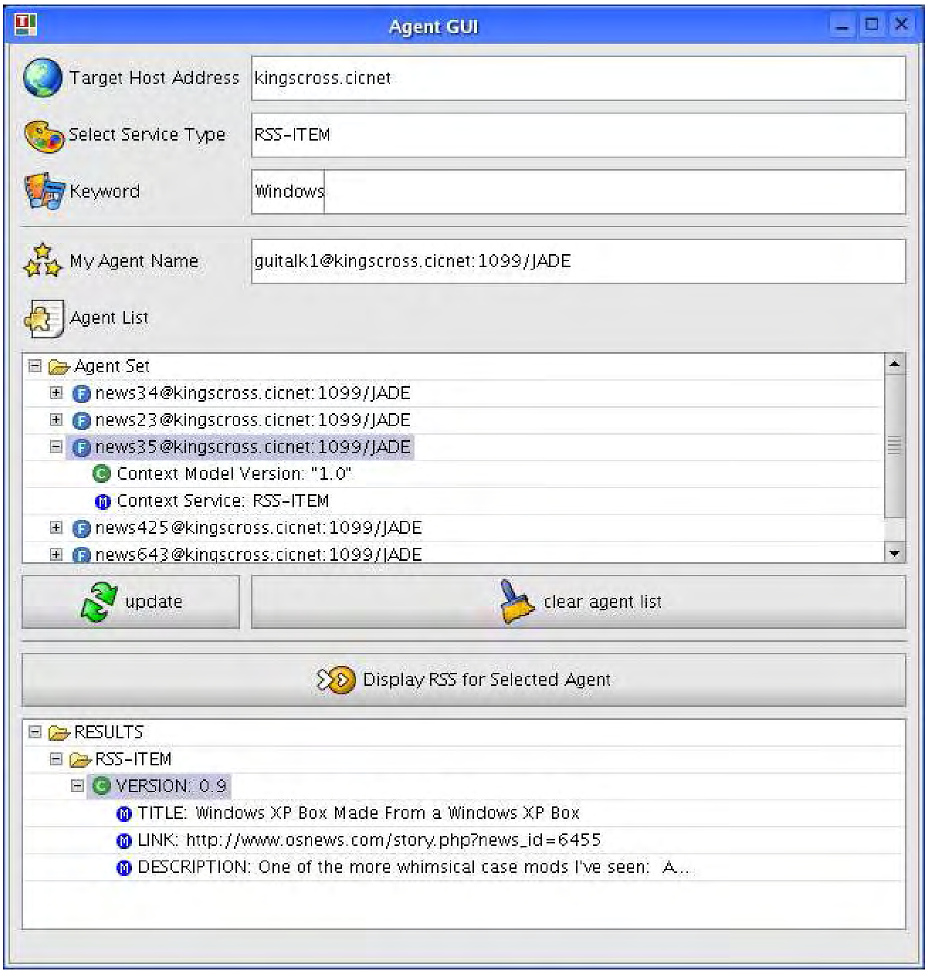


Figure 3: Agent-based RSS aggregation GUI

On the need for annotation-based image retrieval

Masashi Inoue
National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku
Tokyo, Japan
m-inoue@nii.ac.jp

ABSTRACT

Compared with content-based image retrieval, annotation-based image retrieval is more practical in some application domains. Users' information needs and the semantic contents of images can be represented by textual information more easily. We describe two problems which are unique to annotation-based image retrieval and would be worthy of further research. Contextual information embedded in data may be used to address these problems.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information Search and retrieval—*Retrieval models*

General Terms

Algorithms

Keywords

Image retrieval, cross-media retrieval, annotation-based retrieval, word sparseness, noisy annotation

1. INTRODUCTION

1.1 Two Types of Image Retrieval

Image retrieval procedures can be divided into two approaches: query-by-text (QbT) and query-by-example (QbE). In QbT, queries are texts and targets are images. That is, QbT is a cross-medium retrieval. In QbE, queries are images and targets are images. Thus, QbT is a mono-medium retrieval. For practicality, images in QbT retrieval are often annotated by words. When images are sought using these annotations, such retrieval is known as annotation-based image retrieval (ABIR). In contrast, annotations in a QbE setting are not necessary, although they can be used. The retrieval is carried out according to the image contents. Such retrieval is known as content-based image retrieval (CBIR). This paper explains the practicality and research issues of ABIR.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

A basic difference between ABIR and CBIR is related to the values of textual and visual information in image retrieval. Two user studies suggest the importance of textual information in image retrieval. Hughes et al. revealed that users of video retrieval systems tend to use textual information more often than visual information to validate their search results [6]. Another study found a similar result for photo images [3]. As can be seen in many of todays image retrieval systems, such as WWW search engines and clip-art searching software, ABIR is considered practical in many general settings. Consequently, textual information should play a central role in visual information retrieval. However, CBIR has been researched far more than ABIR. Next, we will briefly review some reasons why CBIR research is supported.

1.2 Advantages and Disadvantages of CBIR

Researchers working on CBIR claim that ABIR has limitations. For example, Brahmi et al. mentioned the following two drawbacks in text-based image retrieval [1]: First, manual image annotation is time-consuming and therefore costly. Second, human annotation is subjective. In addition, Sclaroff et al. indicated that some images could not be annotated because it is difficult to describe their content with words [12]. This may be one of the main causes of above two problems. We agree that the above two problems of annotation seem valid; however, we do not think that we should support CBIR instead of ABIR. There are two reasons to support ABIR. First, CBIR has its own problems, which are probably more crucial. Second, the negative effects due to the above problems in ABIR may be mitigated.

First, let us start with the analysis of the problems in CBIR. It is obvious that there are many applications where the use of CBIR is advantageous. As examples, CBIR is suitable for medical diagnoses based on the comparison of X-ray pictures with past cases, and for finding the faces of criminals from video shots of a crowd. These examples can be categorized as "find-similar" tasks; the images to be searched may not differ significantly in their appearances, and so the superficial similarities of the images are more important than the semantic contents. Other applications that involve more semantic relationships cannot be dealt with by CBIR, even if extensive image processing procedures are applied. For instance, in the gathering of the photos regarding the 'Iraq war', it is not clear what kind of images should be used for the querying. This is simply because visual features cannot fully represent concepts. Only texts or words can do that.

Also, it should be noted that in a QbE setting, which is usually the premise of CBIR, users must have an example image at hand. In contrast, in a QbT setting, users simply need to have their search requests in mind because they can compose queries freely using their natural language. Now that we have clarified the advantages and disadvantages of CBIR, we can now consider the problems of ABIR.

1.3 ABIR and Context

Among two major problems in ABIR, let us first consider the problem of subjectivity in annotations from the perspective of context. We can divide the contextual information in IR into two types, depending on when the information is obtained. The first type occurs at the querying step: for example, the interactions between users and systems, users' preferences and skills, or cultural and linguistic differences. The context in IR studies usually refers to this type of user information. The other type of context is found during the process of data creation. For example, the background knowledge of the annotator, the work environment (e.g., amount of time and money), or the potential users (e.g., family members or general public) may influence the characteristics of the resulting data collection. An extreme case might be the context for the same person at two different times. For example, suppose that the context information assigned to a photo image as an annotation immediately after the photo is taken is "lovely silver car." The feeling and the clarity of memories may change over time. When the same person tries to retrieve the image at a later date, the context at that time might be "boring black car." The presence of such embedded type of context is not obvious in IR. In ordinary textual documents, such contextual information is mixed with the thematic subject, since both of the embedded context and the thematic subject are represented by words. In image retrieval, on the other hand, images are just signals and the context of the data creation is implicitly found in the images themselves. If images are annotated, their context can be accessed through the subjective words.

In actuality, subjectivity in annotations may not always cause difficulties in ABIR. Subjectivity implies that the annotations contain some contextual information derived from the annotators' view on the images. Although this subjectivity might cause some mismatches between the users' intentions and retrieval system behavior, as suggested by the studies mentioned in 1.1, such contextual information embedded in annotations is sometimes useful for interpreting images. For the most part, subjective context (e.g., the school of an art work that appears in a photo image) is accessible only by the annotation words assigned to images. Therefore, in ABIR, we can think of the context information as a guide to relevant documents. In CBIR, sometimes the contextual aspect embedded in data, such as the change of visual appearances because of the different illumination conditions, must be eliminated for the objective matching of images.

As we have seen, the subjectivity of annotation, which can be regarded as an obstacle to ABIR, can be useful if we can model the annotations appropriately. In contrast, the first problem—the problem of annotation cost—is obviously an issue that must be solved. Thus, the remainder of this paper focuses on the problem of cost. The difficulties resulting from this problem are twofold: 1) the lack of abundant textual information, 2) the lack of reliable textual information. In the next section, we will explain these two issues in detail.

2. RESEARCH TOPICS IN ABIR

2.1 Word Sparseness

The word information used in IR, such as word co-occurrence frequencies, is often sparse. In annotated images, the occurrences of words are especially limited because they must be assigned only for indexing purposes and the need for such extra effort is not appreciated. The worst annotation may be only one word, which is the file name of the image. Handling such severe word sparseness is one important research topic in ABIR.

The problem of word sparseness may be mitigated by incorporating external knowledge such as thesauri that explicitly identify the relationships between words. This approach is frequently studied in textual IRs and may be applicable to ABIR as well. In addition to explicit knowledge, implicit information can be utilized in ABIR. Zhou et al. suggested that CBIR is limited because it relies solely on low-level visual features. They proposed the use of textual information within the CBIR framework [14]. They also mentioned the problem of word sparseness. They used relevance-feedback (RF) for estimating word associations in annotated images. RF can be considered contextual information at the user-system interaction level.

To overcome the word sparseness problem in ABIR, there has been an attempt to utilize image similarity information for estimating word associations [7]. An interpretation of this approach is that the images are viewed as the context of annotations rather than the annotations are viewed as the representation of image contexts. We think that using such contextual information extracted from target data to overcome word sparseness is a relatively untried research direction.

2.2 Noisy Annotation

When we retrieve images based on annotations, the quality of the annotations should be taken into account. We assume that manually assigned annotations are usually more reliable than automatically assigned ones. Because of the cost, however, annotations are sometimes assigned automatically. Two types of methods are frequently used to assign textual information to images. One method is based on information extraction techniques. For example, some textual information corresponding to images on the WWW can be extracted from their surrounding texts or anchor texts linked to the images. If the extraction rules are carefully designed, the acquired annotations may be relevant to the images. However, because there are usually exceptions in the data that are not consistent with the assumptions, the extracted annotations may contain noise.

The other method is based on classification techniques. The development of procedures for assigning keywords to a given image is an active research topic (e.g., [9]). Such automatic annotation can be regarded as a type of multi-class image classification. Although classification itself has been relatively well studied, automatic annotation cannot be performed easily; the number of classes is large, as large as the vocabulary size, and the amount of data for each class is small, possibly just one example per class. Furthermore— and this could be the most crucial problem in automatic annotation— there may not be simple correlations between the visual features and linguistic concepts.

With the above issues in mind, we should expect that automatically assigned annotations are inevitably noisy. In the field of classification, cleansing of class label noises— for example, by using an ensemble of classifiers [2]—has been studied. In image retrieval, such an ensemble could be constructed from multiple textual information sources. For example, sometimes different texts may concern the same image in different domains. Multiple models of the same collection could refine each other by eliminating the noise. In addition to cleansing, the creation of techniques that utilize the noise as a contextual information source for retrieval may be worth pursuing.

3. CONCLUSIONS

In this paper, we compared two types of image retrieval, namely CBIR and ABIR. We showed the importance of studying ABIR. ABIR is powerful because it can utilize the power of natural language to represent users' search needs and the semantic contents of images.

Despite its practicality, there have been less research conducted on ABIR than on CBIR. (Although there are some exceptional and important research projects that explicitly deal with ABIR: [13],[5],[4],[10],[11].) One possible reason for this lack of popularity may be that researchers believe there are no more interesting research topics in ABIR. Since ABIR is basically mono-medium retrieval, most of the retrieval process can be done by applying conventional textual IR methods. Another reason for the lack of ABIR research may be that users (or perhaps researchers) are skeptical about the possibilities of image retrieval that goes beyond the simple term-matching scheme. Although the above opinions seem reasonable, this paper has nonetheless attempted to show that there are still some interesting research topics —such as solving the problems of word sparseness and noisy annotations— that may lead to improvements of ABIR.

In our view, the use of contexts, especially those which are embedded in target data rather than the commonly utilized contexts found in user-system interactions, may provide some interesting solutions to the problems of ABIR. Recently, in the field of classification, the use of unlabeled data as an implicit information source has been studied (e.g., [8]). Unlabeled data are automatically collected but not manually labeled for use. In the field of image retrieval, such readily available but conventionally unnoticed information, which can be categorized as context, may also be utilized.

4. ACKNOWLEDGMENTS

This study is partly supported by a Grant-in-Aid for Scientific Research on Priority Area "Informatics" (Area #006).

5. REFERENCES

[1] D. Brahmi and D. Ziou. Improving CBIR systems by integrating semantic features. In *Proc. RIAO*, pages 291–305, Vaucluse, France, April 2004.

[2] C. Brodley and M. Friedl. Identifying mislabeled training data. *Journal of Artificial Intelligent Research*, 11:131–167, 1999.

[3] Y. Choi and E. M. Rasmussen. Users' relevance criteria in image retrieval in American history. *Information Processing & Management*, 38(5):695–726, September 2002.

[4] D. Elworthy. Retrieval from captioned image databases using natural language processing. In *Proc. 9th Int'l Conf. on Information and Knowledge Management*, pages 430–437, McLean, VA, 2000.

[5] S. Flank. A layered approach to NLP-based information retrieval. In *Proc. 36th Annual Meeting of the Association for Computational Linguistics and 17th Int'l Conf. on Computational Linguistics*, pages 397–403, San Francisco, 1998.

[6] A. Hughes, T. Wilkens, B. Wildemuth, and G. Marchionini. Text or pictures? an eyetracking study of how people view digital video surrogates. In *Proceedings of the International Conference on Image and Video Retrieval*, pages 271–280, July 2003.

[7] M. Inoue and N. Ueda. Retrieving slightly annotated images. Submitted.

[8] M. Inoue and N. Ueda. Exploitation of unlabeled sequences in hidden Markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(12):1570–1581, Dec. 2003.

[9] J. Jeon, V. Lavrenko, and R. Manmatha. Automatic image annotation and retrieval using cross-media relevance models. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 119–126, July 28-August 1, 2003.

[10] L. Khan and D. McLeod. Effective retrieval of audio information from annotated text using ontologies. In *Proc. Int'l Workshop on Multimedia Data Mining*, pages 37–45, Boston, MA, August 2000.

[11] H. Lieberman, E. Rozenweig, and P. Singh. Aria: An agent for annotating and retrieving images. *Computer*, 34(7):57–62, 2001.

[12] S. Sclaroff, M. L. Cascia, S. Sethi, and L. Taycher. Unifying textual and visual cues for content-based image retrieval on the world wide web. *Computer Vision and Image Understanding*, 75(1-2):86–98, 1999.

[13] A. F. Smeaton and I. Quigley. Experiments on using semantic distances between words in image caption retrieval. In *Proc. 19th ACM Int'l Conf. on Research and Development in Information Retrieval*, pages 174–180, Zurich, 1996.

[14] T. S. H. Xiang Sean Zhou. Unifying keywords and visual contents in image retrieval. *IEEE Multimedia*, 9(2):23–33, April-June 2002.

Testing the principle of polyrepresentation

Mette Skov, Henriette Pedersen, Birger Larsen and Peter Ingwersen
Department of Information Studies, Royal School of Library and Information Science
Birketinget 6, DK-2300 Copenhagen S, Denmark
Email: {ms,blar,pi}@db.dk

Categories and subject descriptors
H.3.3. Information Search and Retrieval; H.3.7. Digital Libraries

General Terms:
Performance; Experimentation.

Keywords
Information Retrieval, Polyrepresentation; Cognitive Overlaps.

1. INTRODUCTION

The cognitive theory of Information Retrieval (IR) and the principle of polyrepresentation derived from it [1;2] provide a theoretical background for how to exploit different document features in order to improve performance in IR. In summary, the theory hypothesises that overlaps between different cognitive representations of both users' information situation as well as documents can be exploited for reducing the uncertainties inherent in IR, thereby improving the performance of IR systems. Good results are expected when cognitively unlike representations are used, e.g., the document title (made by the author) vs. intellectually assigned descriptors from indexers and vs. citations made by other authors over time.

Essentially the principle of polyrepresentation signifies to make use of a variety of *contexts* – in particular associated with information objects – but also in principle related to the searcher and other central components of interactive IR. Context in IR can take several forms [3]. Intra-document structures and representations (signs) are contextual to one another, as are the documents themselves (inter-document relationships). The current session constitutes a third kind of context dealing with features of the interaction between searcher(s) and documents. Examples are eye or mouse movements and pointing and searcher request features, like depth of knowledge of the work task. Further, the interaction processes are in context of the conceptual, emotional, systemic and socio-organizational properties immediately surrounding the searcher and documents. All actors, information systems, documents and interactive sessions are influenced to a certain extent by remote contextual constructs, such as general techno-economic and socio-cultural infrastructures in society. Across this stratification operates an additional dimension, that of the historic context of actors' experiences forming their expectations. In the present paper elements of the intra and inter-document contexts are involved, that is, functionally different content representations like titles and abstracts and cognitively different representations like descriptors assigned by indexers as well as inter-document relationships in the form of bibliographic references made by the authors.

SIGIR'04 July 25-29, 2004, Sheffield, UK.

When representations with different cognitive (and functional) origin point at the same documents via so-called cognitive overlaps, it is regarded as evidence of high probability of relevance. In this paper we present the results from an experiment testing the principle of polyrepresentation in a test collection. The main purpose of the experiment is to show whether the use of cognitively different document representations as suggested in the cognitive theory of IR can enhance performance. Although the cognitive theory of IR and the principle of polyrepresentation by nature are holistic, polyrepresentation is, however, inherently Boolean in much of its reasoning. This is apparent in the pronounced focus on cognitive retrieval overlaps, i.e., sets of retrieved documents based on different cognitive representations. The second purpose of this experiment is to show the consequences of implementing a highly structured search strategy into a best match IR system.

2. THE EXPERIMENT

The test environment was the Cystic Fibrosis test collection [4] indexed in the InQuery retrieval system (version 3.1). The test collection consists of 1,239 document representations from Medline, 100 requests and tripartite relevance assessments. Although small this test collection is ideally suited for testing polyrepresentation as it contains two cognitively different representations and a number of functionally different ones. More recent test collections, such as TREC, are significantly larger in volume, but do not contain the cognitively and functionally different representations needed in this study.

Hence, we made use of the functionally different titles (TI), abstracts (AB) and references (RF) all made by the author. Major (MJ) and minor MeSH (MN) descriptors used in the experiment are functionally different, both representing the indexer as a cognitive agent. In order to test the principle of polyrepresentation we made use of 29 requests searched as both natural language and highly structured queries. The former queries were weakly structured [5] in the sense that Boolean operators were used only to combine the representations for identifying overlaps in InQuery. The highly structured queries were structured by the use of InQuery's Boolean operators in combining query facets and also in combining representations for identifying overlaps. Furthermore proximity operators were used to search phrases and MeSH synonyms were added to the search keys. By using both natural language and highly structured queries it was possible to analyse which search configuration was optimal when implementing polyrepresentation in a best match system.

The experiment revealed that on average 88% of documents found searching AB were also found searching TI. Therefore, the two representations (TI/AB) were indexed in the same field. Combining the four representations (TI/AB, MJ, MN and RF) resulted in 15 overlaps[1] (see table 1). The overlaps were defined such that a document could appear in one, and only one, of the 15 overlaps. Inspired by an earlier study of polyrepresentation [6;7] references were included in the search without an *a priory* intellectual selection of seed documents. Instead a subject search was performed for each request in SciSearch. The cited references in the retrieved documents were ranked using the RANK command in Dialog. For each request the cited references ranked top three on the list were used as input in a (RF) search in the test collection. Those search results for the reference representation included documents containing one or more of the top three cited references in their reference lists. The 29 requests from the test collection were used without modification as direct bag-of-words input for the natural language queries searched in TI/AB, MJ and MN, respectively. The highly structured queries consisted of the same 29 requests, modified in a number of ways. First, a noun-phrase finder parsed them. Secondly, stop words were removed, and finally the remaining search keys were expanded intellectually using the MeSH thesaurus.

3. RESULTS AND DISCUSSION

The tripartite relevance assessments provided in the test collection made it possible to investigate the retrieval performance for both all relevant documents (relevant and highly relevant documents) and highly relevant documents. The results are presented in table 1. Columns A and D show the number of documents found in each of the 15 overlaps in natural and highly structured language. Not surprisingly the natural language queries result in many more documents than do highly structured queries.

Table 1 also shows that, in general, overlaps generated from three or four representations (overlap 1-5) have higher precision than overlaps generated from two or one representations (overlap 6-11 over overlap 12-15).

These findings support the principle of polyrepresentation suggesting that a high number of representations pointing towards a document are likely to be an indicator of high precision. Table 1 shows that for all 15 overlaps highly structure queries result in higher precision (column E) than queries in natural language (column B). From a polyrepresentative point of view this can be explained by looking at the query structure. The highly structured queries ensure that documents identified in an overlap have identical or synonym search terms present from all the representations searched. The weak structure in the natural language queries does not ensure that the search terms (or synonyms) are present in each of the document sets generating the overlaps. Therefore, proper polyrepresentation in the true sense of the concept cannot be achieved with weakly structured queries in natural language.

As described above, table 1 reveals, that overlaps generated from three or four representations have higher precision than overlaps generated from two representations etc. However, looking at the overlaps generated from 3 representations (overlap 2-5) in the highly structured queries, overlap 2 has a considerably lower precision than overlaps 3-5. This suggests that the RF as representation is important to obtain high precision. This high-precision trend when RF is included also pertains to overlaps generated from 2 different representations (overlaps 8, 10, 11). These findings stress the importance of including representations that are both cognitively dissimilar (TI/AB; MJ/MN) and functionally different (RF) as suggested in the cognitive theory of IR.

Table 1: Recall and precision for the 15 overlaps. (Ol = overlap, P = precision, R = recall, # doc. = number of retrieved).

	Natural language			Highly structured language				
	# doc.	P all relevant	R all relevant	# doc.	P all relevant	P highly relevant	R all relevant	R highly relevant
Overlap	A	B	C	D	E	F	G	H
Ol 1 (ti/ab, mj, mn, rf)	126	41%	5%	58	69%	53%	4%	6%
Ol 2 (ti/ab, mj, mn)	668	13%	8%	100	42%	20%	4%	4%
Ol 3 (ti/ab, mj, rf)	101	48%	4%	66	79%	45%	5%	6%
Ol 4 (ti/ab, mn, rf)	240	29%	6%	68	62%	47%	4%	7%
Ol 5 (mj, mn, rf)	3	0	0	11	64%	45%	1%	1%
Ol 6 (ti/ab, mj)	702	12%	7%	131	45%	22%	5%	6%
Ol 7 (ti/ab, mn)	1761	9%	14%	210	27%	13%	5%	6%
Ol 8 (ti/ab, rf)	1528	9%	12%	162	27%	19%	4%	6%
Ol 9 (mj, mn)	141	6%	1%	42	26%	14%	1%	1%
Ol 10 (mj, rf)	6	33%	0	16	38%	19%	1%	1%
Ol 11 (mn, rf)	42	21%	1%	68	34%	16%	2%	2%
Ol 12 (ti/ab)	16201	2%	25%	770	12%	5%	8%	8%
Ol 13 (mj)	106	10%	1%	109	27%	12%	3%	3%
Ol 14 (mn)	603	4%	2%	336	17%	7%	5%	5%
Ol 15 (rf)	872	5%	0	2458	6%	2%	12%	10%

[1] We use the term overlap even though the documents in overlaps 12-15 were retrieved form one representation only.

4. CONCLUSION

The experiment supports the principle of polyrepresentation that states that when a number of cognitively (and functionally) different representations point to a document, this fact is likely to be an indicator of high relevance.

Because of the structured (Boolean) nature of polyrepresentation, a strongly structured query language is necessary when implementing the principle of polyrepresentation in a best match IR system. Finally, the results indicate that scientific references (outlinks) serve as central contextual elements during IR and constitute an important type of representation in order to obtain high precision.

5. ACKNOWLEDGMENTS

The authors wish to thank the Center for Intelligent Information Retrieval, University of Massachusetts Computer Science Department, Amherst, MA for providing the InQuery software.

6. REFERENCES

[1] Ingwersen, P. (1994): Polyrepresentation of information needs and semantic entities: elements of a cognitive theory for information retrieval interaction. In: *Proceedings of SIGIR 1994*, p. 101-110.

[2] Ingwersen, P. (1996): Cognitive perspectives of information-retrieval interaction - elements of a cognitive IR theory. *Journal of Documentation*, 52(1), p. 3-50.

[3] Ingwersen, P. and Järvelin, K. (Forthcoming). *Integration of Information Seeking and Retrieval in Context.* Klüver.

[4] Shaw, W. M.; Wood, J. B., Wood, R. E., and Tibbo, T. R. (1991). The cystic fibrosis database: content and research opportunities. *Library and Information Science Research,* 13, p. 347-366. (The test collection may be downloaded from http://www.dcc.ufmg.br/irbook/cfc.html).

[5] Kekäläinen, J. and Järvelin, K. (1998): The impact of query structure and query expansion on retrieval performance. In: *Proceedings of SIGIR 1998,* p. 130-137.

[6] Larsen, B. and Ingwersen, P. (2002): The boomerang effect: retrieving scientific documents via the network of references and citations. In: *Proceedings of SIGIR 2002,* p. 397-398. (poster paper)

[7] Larsen, B. (2004): *References and citations in automatic indexing and retrieval systems : experiments with the boomerang effect*. Copenhagen: Royal School of Library and Information Science. XIII, 297 p. (PhD dissertation) [http://www.db.dk/blar/dissertation, visited 11-2-2004]