

Information Retrieval in Context – IRiX

Peter Ingwersen¹ & Kalervo Järvelin²

¹ Department of Information Studies, Royal School of Library and Information Science, Copenhagen, Denmark pi@db.dk

² Department of Information Studies, Tampere University, Finland
kalervo.jarvelin@uta.fi

<http://irix.umiacs.umd.edu/>

1 Introduction

IR research is now conducted in multi-media, multi-lingual, and multi-modal environments but largely in a context-free manner (Ingwersen & Järvelin, 2005). However, the retrieval of such information depends 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. Such contextual data can be used effectively to constrain retrieval of information thereby reducing the complexity of the retrieval process.

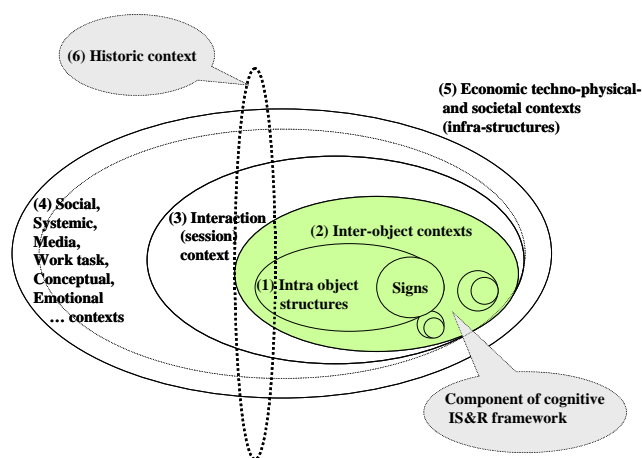


Fig. 1. Nested model of context stratification for IR. From (Ingwersen & Järvelin, 2005, p. 281).

Today real-life IR in context is nearly always interactive IR. One may observe a stratification of contexts in association to IR engines and systems see Fig. 1. Such strata range from the traditional content features of and between information objects, like words nested in paragraphs and hyperlinks, over eye and mouse movements during session-time into the searcher's current

work or daily-life task 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.

Some efforts of capturing context features have been made in previous decades of interactive IR research (Cool, 2001; Cool & Spink, 2002). However, there is a lot of room for investigating more intensively and distinctively the nature of context associated with IR.

The SIGIR 2005 Workshop on IR in context (IRiX) attempted more closely to investigate core issues of context. The 2005 IRiX continued last year's IRiX Workshop. In 2004 in Sheffield the participants observed and discussed the wide landscape of possible context features in IR and made available many paper presentations and some demos. This year's workshop changed this organization by focussing the efforts in a highly structured manner on *three major lines of action* or questions that explore the central features or evidence of context:

1. What are the elements of context, which are potentially significant to information retrieval?
 - Work or daily-life task or interest features
 - Searcher features
 - Interaction features
 - System features
 - Document features
 - Environmental/physical features
 - Temporal features
2. Which of these categories/elements are, or could be useful in improving information retrieval?
 - This is dependent upon the situation
 - This is also dependent upon the goal of the searcher
3. How can features of context be used to improve information retrieval?
 - We need to develop methods of investigation, test beds, evaluation measures and techniques
 - We need to explore potential operationalization of aspects of context

The goal was to concentrate on a workable agenda for future research on IRiX by substantiating the characteristics and recommendable methodological approaches.

2 Organization of IRiX

The Organizing Committee consisted of P. Ingwersen, Denmark, Nick Belkin, USA, and K. Järvelin, Finland. It was decided that in 2005 only a handful of short position papers (max. 2000 words) should be presented in plenum while other relevant contributions should be applied during the workshop as 'background papers'. Based on the papers and written proposals three

Discussion Groups were decided upon. All contributions were peer reviewed by the Organizing/Review Committee consisting of the following persons:

Nick Belkin, Rutgers University, USA
Pia Borlund, Royal School of LIS, Aalborg, Denmark
Peter Ingwersen, Royal School of LIS, Copenhagen, Denmark
Kalervo Järvelin, Tampere University, Finland
Jose Joemon, Glasgow University, Scotland
Birger Larsen, Royal School of LIS, Copenhagen, Denmark
Andrew MacFarlane, City University, London, UK
Ian Ruthven, Strathclyde University, Scotland
Amanda Spink, University of Pittsburgh, USA
Pertti Vakkari, University of Tampere, Finland
Ryen White, University of Maryland, USA

Birger Larsen, Denmark, and Ryen White, Scotland, organized the program and publication layout of presentations and the workshop website: <http://irix.umiacs.umd.edu/>

The final program became structured as follows: prior to the coffee break four theoretical papers were presented that led up to the three selected Discussion Groups. These groups addressed some of the action line issues shown above:

Group A) Context, Situation, and Task: Implications for IR

Chair: Nicholas J. Belkin

Group B) Adding Context to the Cranfield Laboratory Model

Chair: David Hawking / Ross Wilkinson

Group C) Central Research Variables in Contextual IR

Chair: Kal Järvelin

Ingwersen introduced the workshop in a presentation on *Selected variables for IR interaction in context*, which could also serve as a starter for group C. Kalervo Järvelin's contribution *Frameworks, models and theories in laboratory IR* discussed how existing models may limit laboratory IR research – the Cranfield Model – and proposed some lines of action. This contribution served as a starting point for the Discussion Group B, dealing with extensions of the laboratory IR model, as well as group C. In particular, the paper *Context in enterprise search and delivery*, by David Hawking, Cécile Paris, Ross Wilkinson and Mingfang Wu concerned the proceedings of Discussion Group B and was closely related to Järvelin's perspectives. Finally, Nick Belkin & Colleen Cool discussed the nature and features of *Context, situation and task*, leading to discussions in Group A. After lunch and in connection the Discussion Group work, Luanne Freund & Elaine Toms presented *Using contextual factors to match intent* providing some proposals of variables and evidence to be applied and/or captured during interaction.

A number of background contributions looked into the laboratory IR extension towards interaction from systems or information objects, e.g., Georgina Ramirez & Arjen P. de Vries in

their paper on *XML and context: Structural features relevant to search tasks*. Xiangji Huang looked into features to add in order to extend the probabilistic model (*Incorporating contextual retrieval into Okapi*) whilst Ryen White (*Contextual simulations for information retrieval evaluation*) and Andreas Komninou & Mark Dunlop (*Calendar based contextual information as an Internet content pre-caching tool*) discussed simulations of searcher behaviour as a gemmation of the laboratory tradition into confined searcher scenarios.

The latter contribution bridges over to the remaining papers that all are dealing with aspects (features) of IR interaction, also in non-textual media. They may thus serve as background partly for group A on the nature of situation, context and tasks, partly for group C on the research variables to take into account.

Personalized searching (Xuehua Shen, Bin Tan & ChengXiang Zhai), the roles of actors accessing their e-mail archives (Anton Leuski) and information need representations (Jimmy Lin & Dina Demner-Fushman) touched upon the information seeking actor side of interaction. Web IR (Jin Ha Lee, Xiao Hu & J. Stephen Downie) and knowledge generation on the web (Haakon Lund et al.) incorporated potentially other media. Finally, Elaine Toms et al. moved the focus to the working (organizational) context of IR and its influence on searching – in this case the impact and constraints of interruptions in the process. All the papers are available in full on the IRiX website.

The 24 participants divided themselves into the three Discussion Groups after the morning coffee break – with group A as the most popular one. Discussions took place over lunch and ended with their presentation of the results in plenum following the afternoon coffee break.

3 The IRiX Results

We divide the analyses and results obtained by the Discussion Groups into three sub-sections below.

3.1 Discussion Group A: *Context, Situation, and Task: Implications for IR*

Discussion Group A focused on some central concepts concerned with IR in context: ‘Situation’ was “best conceived as the *totality of the values* held by the facet ‘variables’ where ‘facets’ are seen as users, systems and environment.

According to the Group the *User facet* at generic level consists of Motivation; Knowledge; History; and Individual differences. In more detail, the group viewed ‘Motivation’ as consisting of the following variables: Imposition (Intrinsic vs. Extrinsic); Task type; Task level; Goal; Salience; and Immediacy. ‘Knowledge’ was divided into: Domain knowledge; and Search knowledge. ‘History’ associated to Behaviours (Use behaviours; Viewing behaviours; and Saving behaviours); Queries; Judgements; and Strategies. Lastly, the ‘Individual differences’ concerns Demographics; and Cognitive styles.

The Group analyzed the *Systems facet* according to the following variables to be taken into account: Resource; Interface; Technology; and Retrieval Model.

The ‘Resource’ sub-facet consists of: Domain; Genre; Source; Format; and Media. The ‘Interface’ group of variables is, according to the Group: Query methods; Results display; Help support; Interaction affordances. ‘Technology’ demonstrates this handful: Device (PDA; PC); Architecture (P2P; Intranet; Distributed) and the ‘Retrieval model’ shows: Representation; Ranking; and Matching.

The final major facet, the *Environment*, relates to this range of variables, according to the Group: Temporal; Physical; and Social.

The ‘Temporal’ variables are: Clock-time; Search phase; Constraints; and Duration (e.g., short term, long term): The ‘Physical’ set of variables is seen as: Location; Temperature; and Mobility. Finally, the ‘Social’ parameters are Roles; Norms; and Community. There was some panel discussion on this list of variables, for instance seen in relation to a number of models discussed prior to the group discussions (e.g., Fig. 2) and associated with the process of interaction, not explicitly present.

3.2 Discussion Group B: Adding Context to the Cranfield Laboratory Model

Discussion Group B saw three lines of action that could lead to workable extensions of the Cranfield model and experimental research framework. First, one carries out the traditional batch IR experiment – BUT with ‘full’ context description. This implies to have (full) control of a limited set of contextual parameters to be studied during experimentation. This can be done via simulation of searcher’s situation (see e.g. Ryen White’s contribution in the Workshop Proceedings on the IRiX website).

A second line of action was named *interactive light*. Here the context parameters define the user population, from which the researchers then must elicit a set of sample test persons. The test persons issue requests/questions to two systems to be compared, using side-by-side comparison. The advantages are that interactive light experiments can be embedded in operational environments and demonstrate realism. The disadvantage is probably, that not all contexts are applicable in this framework.

The third line of experimental IR extending the laboratory model might be really ‘interactive’ investigations. Such experiments should incorporate the elicitation of user context features, information needs, the interaction process and the user response to interaction.

The Discussion Group put forward some recommendations.

One alternative seems to be to find context *sweet spots*:

- Run an experiment that measures benefit in an organization;
- Ask customers, find a sweet spot, and prove it;
- Look for solutions in enterprise/personal search, rather than web search;
- Look at current context successes and build system prototype;
- Look at current failures and resolve

An alternative road might be, according to the group:

- Run a user study in very constrained environment;
- Hypothesize approach;
- Optimize system, and run against canned model;
- Run interactive light experiment;
- Start with a canned model, find out what people do with it;
- Look at search failures where context was the key (be it location, ambiguity, doc. type etc.)

Lastly the group proposed this question: What sort of context do we explore?

- Delivery form?
- Context captured as text that can modify a query?
- Context captured as metadata that can modify structured queries?

The latter can be seen as an extension of the traditional topical IR experiments. The panel discussion focused on the conception of ‘interactive light’ methodology.

3.3 Discussion Group C: *Central variables in contextual IR*

Discussion Group C was concerned with central variables in contextual IR. The group had a long discussion on various kinds of variables, proposed by the participants, which the group tried to organize into a coherent framework. This turned difficult in the limited time frame allowed due to many perspectives present. One way of organizing the variables is to arrange them to contexts, see Fig. 2, of IR, information seeking, work task, and socio-organizational and cultural context (Ingwersen & Järvelin, 2005, p. 322).

Therefore, the Discussion Group decided to design a concrete study, which would exemplify the effects of traditional and contextual variables in IR interaction and allow the assessment of the contribution of an IR system in performing some work task type. The group chose the school exams environment as the focus. In this environment, see Fig. 3, one would have:

- A collection of high school textbooks under two IR systems: one could be, e.g., printed books with back-of-book indices and the other a CD-ROM collection equipped with a search system with/without support to browsing hypertext structure of the textbooks;
- National exam questions in Finnish high school subjects (e.g., religion, physics, biology, psychology, and history), say a total of 12 questions;
- A set of 2nd year high school students answering the exam questions (as a kind of pre-examination test) divided between the two systems; employing a Latin square design to control for possible interactions of variables;
- The students’ task would be, using one of the systems, to answer 12 exam questions within the regular time allowance for the exams (say 6 hours) as well as they possibly can - as in the real high school exams;

- The answers would be assessed by the teachers of the subjects using a standard mark scale of 0 – 6 points (a zero meaning a bad answer), as done regularly in the national exams.

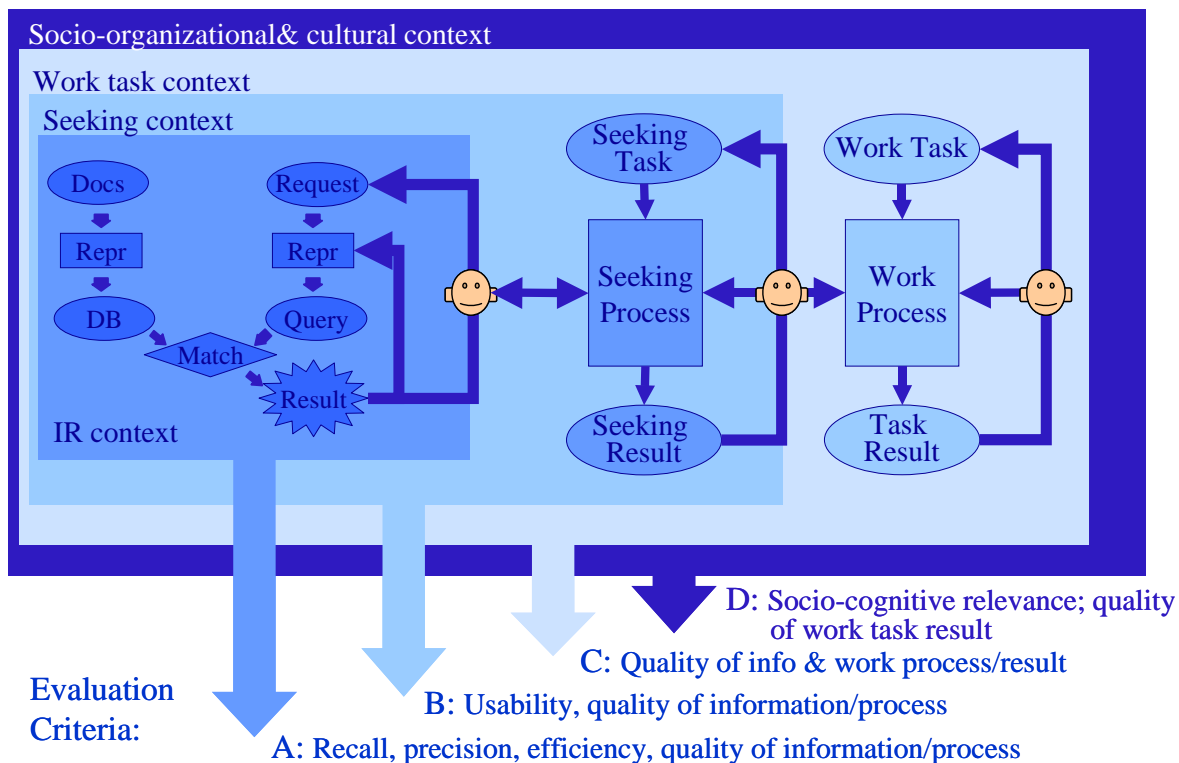


Fig. 2. A research framework for IR: Contextual extensions and relevance criteria. (From Ingwersen & Järvelin, 2005, p. 322)

The work task outcomes - high school exam question answers – would be assessed for their quality in a regular way. Since the task assignments are controlled, all other differences between the users and differences in the exam question answers and their quality would be contributed to the two different IR systems.

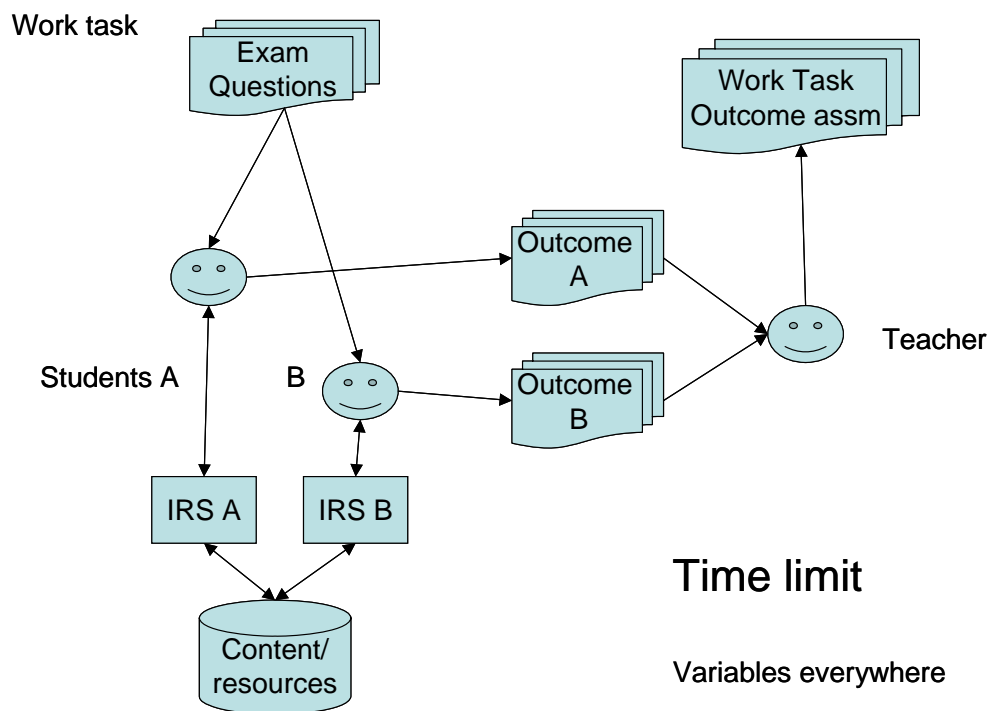


Fig. 3. A research setting example for IR, given a specific work task and environment.

The systems would be introduced to the students prior to the test. Multiple data sets would be collected: search logs, student interviews, students' search notes, exam question answers, the teacher's assessments and the collections. One could also analyze search strategies, query formulation and expansion or use of operators, material found and actually used, etc., and their connections to the outcomes.

The group also discussed other possible work task contexts that would be amenable to a similar analysis, but did not come by good examples. Clearly, one needs to have an intellectual task that is routinely executed by several people in a similar context. The panel discussion centred on how to make more generalizable the research setting example, Fig. 3.

4 Conclusions

In conclusion one may emphasize the Workshop's serious wish to produce and advance workable methodologies for IR in context research. The structured list from Discussion Group A may not be final and some important variables may have been overlooked. However, it provides ideas of what actually to look for as independent, controlled and dependent variables (and values) when performing field studies as well as in connection to laboratory tests. The list is related to the research framework discussed in (Ingwersen & Järvelin, 2005). In that respect the methodological suggestions put forward by Discussion Group B suit IR in context very well. Discussion Group C's contribution became primarily to attempt to provide an exemplary research setting which, in a controlled manner, demonstrates a sensible realism involving real

work tasks, human interaction with IR systems, and adequate performance measures. This succeeded but also did demonstrate the difficulties associated with IRiX.

In connection to the SIGIR 2006 in Seattle, Elaine Toms, Canada, has agreed to make the workshop proposal for the continuation of IRiX.

We wish to thank all the participants and the review committee members for their efforts. We are also indebted to the Brazilian staff that organized our work in a comfortable way.

5 References

Cool, C. (2001) The concept of situation in information science. *Annual Review of Information Science and Technology*, 35, 5-42.

Cool, C. & Spink, A. (2002) Issues of context in information retrieval (IR): an introduction to the special issue. *Information Processing and Management*, 38, 605-611.

Ingwersen, P. & Järvelin, K. *The Turn: Integration of Information Seeking and Retrieval in Context*. Springer, 2005. 460 p. (The Kluwer/Springer International Series on Information Retrieval).