EXTENDING INFORMATION SEEKING AND RETRIEVAL RESEARCH TOWARD CONTEXT

K. Järvelin[1] and P. Ingwersen[2]
Email: [1]kalervo.jarvelin@uta.fi - [2]pi@db.dk

[1]Department of Information Studies, University of Tampere – FINLAND
[2]Department of Information Studies, Royal School of Library and Information Science – DENMARK

ABSTRACT

The paper analyzes research in Information Seeking and Retrieval (IS&R) regarding its frameworks, focus areas and neglected areas. Based on the analysis, two action lines are proposed. On the one hand, IR research needs to be extended to capture more context but without totally sacrificing the laboratory experimentation approach. Only by this line of action one may approach real *IR engineering*. IR systems should be seen in context of their use. On the other hand, current information seeking research needs to be extended both toward the task context *and* the technology. The diversity of task/actor contexts is far from exhausted in information seeking. Therefore lots of research is needed exploring IS&R in further task/actor contexts. Moreover, the systems context in information seeking research has been limited and often nonexistent. This research should reach toward system and interaction features.

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.

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

9. The *access and interaction dimension*: strategies of information access, interaction between the actor and the interface (both in social and in system contexts).

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.

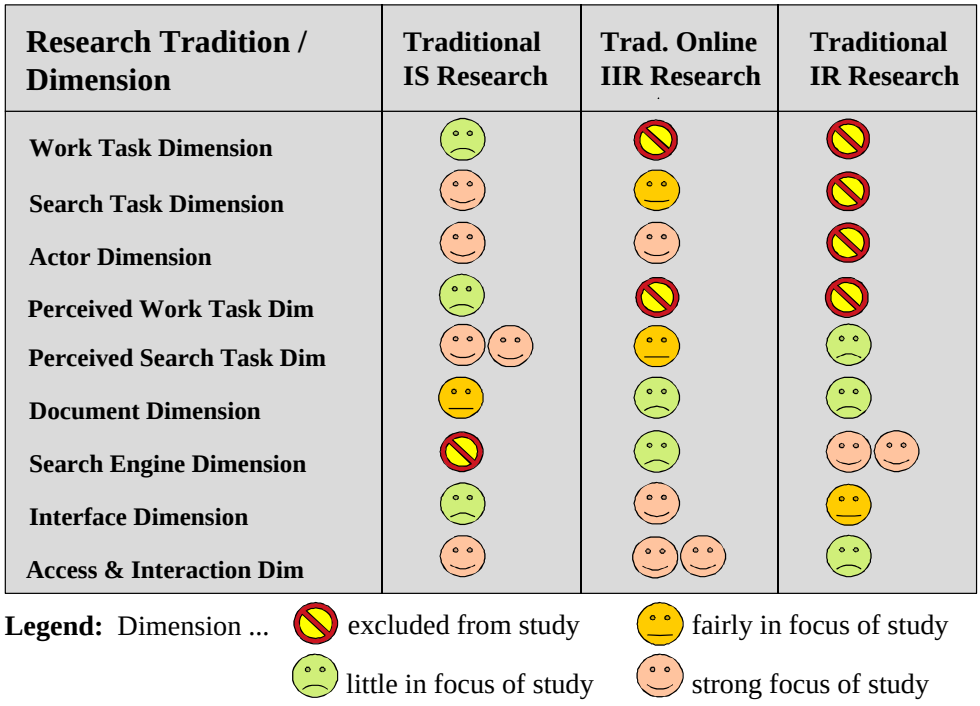

Figure 1. Foci of traditional IS&R research

Section 2 discusses design and evaluation frameworks for IS&R with a starting point in the nine broad dimensions presented above. Section 3 presents the space for IS&R research and the elements research so

far has covered. Section 4 gives the conclusions. This article is based on the forthcoming book by the authors (Ingwersen & Järvelin, 2005).

2 DESIGN AND EVALUATION FRAMEWORKS FOR IS&R

In this section we discuss the design and evaluation frameworks (theory) of IS&R in context. The nine dimensions presented in the introduction are essential dimensions of this context and discussed first. Thereafter we present a design and evaluation framework based on three embedded layers of the task dimension: work tasks, seeking tasks and retrieval tasks.

2.1 Broad 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 components. 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.

2.2 IR Research in Isolation

With a view on the nine 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 relationship 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.

2.3 Information Seeking Research in Isolation

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.

Developing *theoretical understanding* of a domain is a necessary task for any discipline. An essential issue is the definition of the domain. It should cover a meaningful system of phenomena that supports explanation and understanding. The theoretical understanding of information seeking clearly has advanced in the 1990's as several new models suggest. Taken together they suggest a perspective covering phenomena from information systems and their design, through information access by various processes to work tasks (or other activities). The focus of theoretical analysis has however been in the *seeking process*: its stages, actors, access strategies, and sources. Work tasks and information (retrieval) systems have received less theoretical attention.

Developing *empirical understanding* of phenomena within the domain is also necessary for a discipline. Theoretical understanding must be grounded on observables. Otherwise it turns into speculation. Information seeking phenomena in various contexts are understood, explained and predicted by having theoretically justified findings on work, seeking and search tasks and their context. With a few exceptions, the empirical findings concentrate on the seeking processes, with less attention to work tasks and information (retrieval) systems. They were often quite descriptive. The empirical studies provided only few answers

to research questions relating characteristics of contexts and situations to characteristics of tasks, actors, and information, seeking processes, sources, systems and use of information. The process oriented modern approach in Information Seeking has covered several empirical domains in, e.g., Social Science and Engineering, and some work task contexts, e.g., student information seeking for a term paper or research proposal. However, many remain unexplored. This is only healthy for a research area.

Supporting information management and *information systems design* may be the weakest contribution of Information Seeking. This may be understood through Figure 1 – studies in Information Seeking rarely include information (retrieval) system design features in their study settings – features that the information (retrieval) system designers find relevant and deal with. In such a situation the research results cannot communicate to systems design, the worlds do not touch.[2] While our understanding of work task requirements and effects on information seeking has advanced, the understanding on how to derive and apply design criteria for information (retrieval) systems has not advanced correspondingly.

These considerations suggest that research in Information Seeking should be extended both toward work tasks and toward information (retrieval) systems (or technology) – see Figure 2. Having its roots in Library Science user studies, Information Seeking has come a long way toward research that is no more revolving around the users of a single institution. However, information seeking as such is the study of *something-in-between* and not a theoretically justified area in isolation. Paying due attention to the goal of augmenting work task performance and alike lay interests (Figure 3) and the available technologies turn Information Seeking much alike the disciplines Information Management, Information Systems, Organizational Design etc. One may loose one's independence but gain a better ability to communicate across disciplinary boundaries.

2.4 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 IR experiment as well as real interactive IR over several short-term interactions. Possible evaluation criteria in each context are given: A – D. The nine 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.

[2] In principle, it may also be that IR system designers are busy with wrong variables or features.

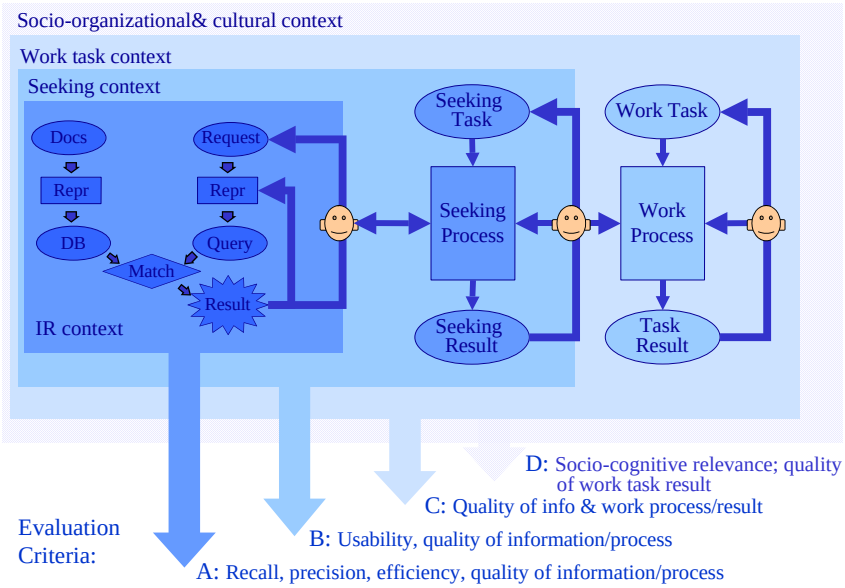

Figure 2. Nested contexts and evaluation criteria for task-based IS&R

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.

An obvious counterargument is that there are too many seeking contexts with too many possible combinations of systems and tools – the design and evaluation of IR systems becomes unmanageable – therefore it is best to stick to the tradition of design and evaluation. If one does not know more than one's own unsystematic recollection of personal IR system use, such design and evaluation demands may be of tall order, indeed. However, even limited knowledge on real IS&R may reveal typical uses, strengths and weaknesses of various tools and systems – and how their users perceive them. This provides a better basis for design than de-contextualized standard assumptions and measures. If automobile designers would behave alike, they would focus on the engines (e.g., horsepower, acceleration) no matter whether they design a sports car, pick-up or a truck! A nice parallel may be observed in the critique of Information Seeking research by Dervin & Nilan (1986) – *mutatis mutandis*.

Finally, information seeking seldom is an end in itself but rather serves a work task (or other interest). The real impact of information seeking and retrieval is its contribution to the work task process (e.g., effort, time) and the quality of the result. Therefore, in the end, IS&R should be designed and evaluated for their utility in the work task context. Again, an obvious counterargument is there are too many work task contexts that are too weakly related to IR. The design and evaluation of IR systems thus becomes unmanageable and cannot learn from all too remote task requirements. Therefore, the counterargument goes, it is best to close one's eyes and stick to the tradition of design and evaluation. However, even limited knowledge on real work tasks may reveal typical uses, strengths and weaknesses of various tools and systems – and how their users perceive them. Moreover, many work task requirements are relevant to IR design.

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.

2.5. 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 (1962) framework for knowledge work augmentation, where a human is augmented by language, artifacts and methods in which (s)he has been trained. [3]

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.

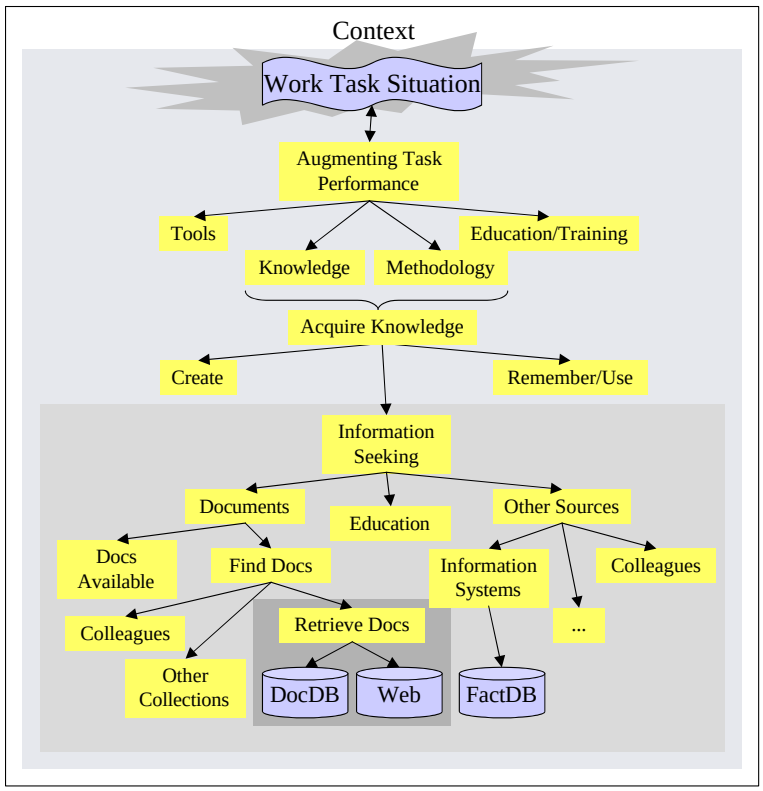

Figure 3. Augmenting work task performance – perhaps by IS&R (based on Järvelin, 1986)

[3] 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.

3. THE MULTIDIMENSIONAL RESEARCH DESIGN CUBE

The nine broad dimensions – work task (and organization), search task, actor, perceived work task, perceived search task, document, algorithmic search engine, algorithmic interface, and access and interaction – form a multidimensional research design cube with a lot of variables for each dimension. In order to approach a research program for IS&R, it is useful to look at the research done so far in the light of this research design cube. We shall use the approach of Figure 1 and consider the research done so far in a table format by main dimensions (Table 1 below).

Table 1 contains, in the case of IR, broad types of studies sharing the same strong model, the Laboratory Model. Because many studies subscribe to the same model, it is quite straightforward to find three representative study types for IR: the basic laboratory IR, the laboratory-based interactive IR and the online interactive IR. However, such a single model is not available for IS studies, and therefore we need to exemplify the area by more study types of less popularity (in fact, sometimes nearly individual studies). White space in the table denotes dimensions receiving no attention in the studies and progressively darker shading denotes the dimensions receiving marginal, fair or high level of attention.

Table 9.1. The research design cube and examples of the research done in IS&R. Shading represents focus areas of each study type from none to high level.

	DIMENSIONS EMPLOYED							
STUDY TYPE	**Work Tasks & Org**	**Search Tasks**	**Actors**	**Information Needs**	**Documents**	**IR Engines**	**IR Interfaces**	**Access**
1. Lab IR	-	Fixed, static	-	-	News full text	Exact and best match	-	Well-defined, topical
2. Lab IIR	-	Fixed, static	Searchers	-	News full text	Exact and best match	Query modification	Well-defined, topical
3. Online IIR behavior	-	Natural, short term	Searchers	Natural, short term	Doc meta data	-	Operational, Boolean	Natural, full bib access
4. Web searcher groups	-	Varied & natural	Actor groups	Assigned and natural	-?	-	Operat. ranked	Natural strategies & moves
5. Term Paper IR	Fixed	Natural, longitudinal	Students	Natural, longitudinal	Bib references and full text	-	Operational, Boolean	Natural strategies & moves
6. Web log surveys	-	-	-	-	-	-	Operat. ranked	Random query sample
7. Term Paper IS	Fixed	Natural, longitudinal	Students	Natural, longitudinal	Natural, all kinds	-	-	Natural, broad
8. Job level IS	Controlled	-	Actor groups	Natural, summary level	Types of sources	-	-	-
9. Discourse Analysis	-	-	Actor groups	Natural	Discourse sample	-	-	-

The table shows that laboratory oriented IR studies (1-2) focus on IR engines and query modification in best match environments, everything else by and large fixed. The human oriented IR studies (3-4) focus on searchers and access strategies in Boolean or Web search environments. All IR study types (1-4) neglect work tasks (and interests) and longitudinal information needs. The study type (5) spans over all dimensions but IR engines. It informs on how Boolean IR systems are used to support a given work task type. IR study types (3-5) employ dynamic information needs / search tasks either at session level or over longer time.

The Information Seeking studies (7-9) ignore IR systems and focus on actors and information needs, which are natural and dynamic (7) or summary (8) – independent of specific work task situations. In the case of Discourse Analysis (9) the information needs are certainly natural and may evolve while the actor discusses them. However, they are not associated with concrete search situations.

In each study type, white space on its row indicates excluded dimensions. Thus they cannot be used to analyze variation in the other dimensions. Variation in the other dimensions cannot either be seen to affect the excluded dimensions at all. This is obvious but deserves being stressed. For example, this renders most IR research irrelevant for task performance (knowledge work augmentation) and most Information Seeking research irrelevant for IR systems design. White space also generates hidden variables. For example, in study type (6) specific types of searchers produce specific types of web queries, which the overall survey cannot report.

4. 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.

REFERENCES

Dervin, B & Nilan, M. 1986. Information needs and uses. In: Williams, ME (ed.). *Annual Review of Information Science and Technology*, Vol. 21. White Plains, NY: Knowledge Industry Publications: 3-33.

Engelbart, D. 1962. *Augmenting Human Intellect: A conceptual Framework.* Menlo Park, CA: Stanford Research Institute.

Ingwersen, P & Järvelin, K. 2005. *The Turn: Integration of Information Seeking and Retrieval in Context.* Research monograph. Springer. 500 p.

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

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