Contexts of relevance for information retrieval system design

Erica Cosijn[1]

and

Theo Bothma[2]
Department of Information Science, School of Information Technology,
University of Pretoria, Pretoria, 2001, South Africa
erica.cosijn@up.ac.za
theo.bothma@up.ac.za

Received: 22 September 2005

Users judge relevance in various dimensions, but systems traditionally only support matching of queries to documents or document representations on an algorithmic or topical level. We argue that systems should support users in order for them to make relevance judgements on the level of cognitive relevance, situational relevance, and socio-cognitive relevance as well. Current studies in the field of Information Retrieval and Seeking are discussed from a relevance point of view, in order to show how systems might be adapted to assist users in making multi-dimensional relevance judgements.

1 Introduction

Traditionally, the focus of IR research is on topicality as the deciding criterion for relevance. It is essential to understand the manner in which relevance is judged in order to improve the representation of, and access to information. A previous study has confirmed that users also judge relevance on levels other than topicality (Cosijn 2003). The purpose of this paper is to review the larger significance of these results regarding the implementation of the findings in terms of the possible applicability of the framework defined by the model that is briefly described in Section 2. The main question that will be discussed in Section 3 is: *How can systems be improved in order to help users to make relevance judgements on other levels as well?*

The analysis presented here should be seen as a possible contextualisation of the model within current research projects and provides a guideline for future research on relevance. The research in the field has been mapped to the model in order to expose the 'bigger picture' of what is being done within relevance research. Although the list of studies reviewed below cannot be regarded as being comprehensive, all the studies mentioned already have as underlying theme the understanding of various types of relevance judgements as made by users of IR systems.

2 Proposed relevance model

The different dimensions of relevance have been identified in a theoretical study by Cosijn and Ingwersen (2000), and subsequently modeled (Cosijn 2003) as a modification of the Ingwersen *Cognitive Model of Information Transfer* (Ingwersen 1996). This model is depicted in Figure 1.

In this model, the original elements of the Ingwersen model have been retained, but the dimensions of relevance are defined in terms of the relationships between the information objects (as perceived) and different elements in the information searching process. In this manner, situational relevance can, for example, be defined as the relation between the definition or perception of the work task in the user's mind, on the one hand, and the information objects as perceived by the user, on the other hand. Affective relevance has been shown to operate on a different level and affective relevance judgements can be associated with any of the subjective relevance types.

The model was then empirically tested according to the following issues: usefulness and viability of the model, the influence of the nature of the work task on the application or non-application of documents in work task fulfillment, the influence of work and search task execution on the type of relevance judgements made, whether some relevance dimensions necessarily include others, and the relationships between types of relevance judgements.

Thirty-three users performing three different types of work tasks (undergraduate students writing a guided research essay, masters and doctoral students, and researchers writing articles and conference papers) were asked to judge the relevance of the documents utilized in the execution of the work task. In total, 497 documents were judged, of which 320 were relevant to the works tasks and 147 were retrieved and at least partially read, but were not relevant to the extent that they were cited. The questionnaires and other measuring instruments are described in Cosijn (2004). The empirical

1. Erica Cosijn is a Senior Lecturer in the Department of Information Science, University of Pretoria, South Africa
2. Theo Bothma, PhD, is Professor and Head of the Department of Information Science, University of Pretoria, South Africa

testing showed that the model is a valid representation of the types of relevance judgments made by users (Cosijn 2003). Detailed results will be published elsewhere.

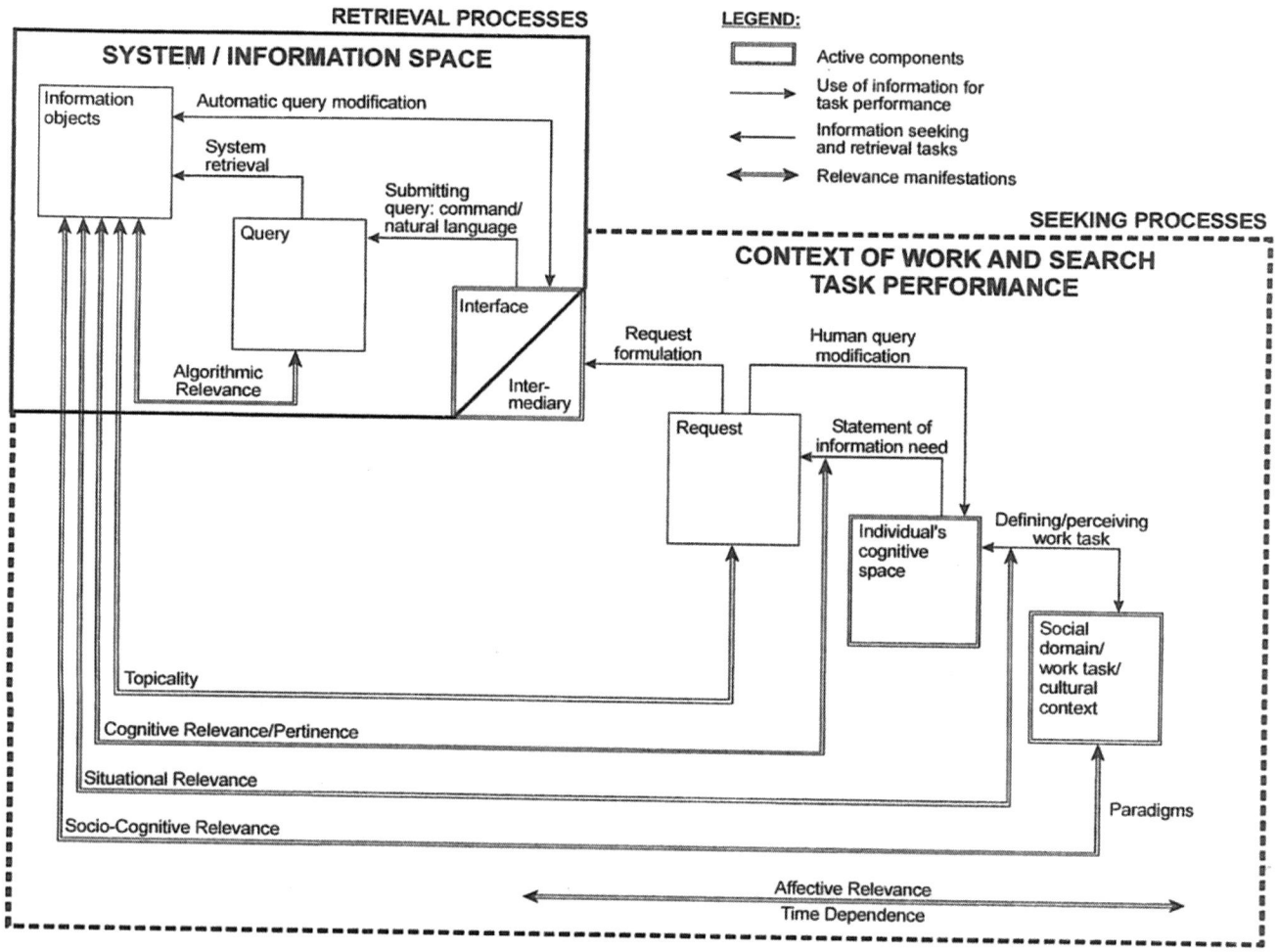

Figure 1 Interactive Information Retrieval: Work task performance, search task performance and relevance types

3 Dimensions of relevance for system design

In Section 3, each of the relevance types is discussed briefly in order to re-establish the parameters of the definitions of the relevance types. These definitions are important, because these are the parameters in which the argument will take place. It is acknowledged that relevance is a fuzzy concept and that definitions vary, but by defining each manifestation clearly and only arguing within those parameters, misunderstandings should be minimized.

For each of these studies, the recent and current research into facilitating these relations in the search process has been analysed, with the focus on the more subjective relevance types of cognitive, situational and socio-cognitive relevance.

3.1 Algorithmic relevance

In the model as depicted in Figure 1, the relation is defined as that between the query and the information objects. This relation is system-oriented to a very large extent, as it depends on the degree of similarity between the features of the query and the features of the information object. This type of relevance is by nature system-dependent. It is not influenced by the user, nor is it related to any subjective information need the user may have.

System or algorithmic relevance is measured in terms of the comparative effectiveness of logical or statistical similarity of features inferring relevance. There are various models of matching the query (as a representation of the user's need) to the information objects (whether as full-text or as representations). Systems may be Boolean (exact match) or best-match (for example vector space, probabilistic, etc.) in nature, or a combination of both. Although this study limits its scope to the more subjective types of relevance judgements, the concept of algorithmic relevance is, nevertheless, included in the model and therefore a brief review of recent projects aiming to increase the comparative effectiveness of the relation between the query and the information objects has been given.

on systems that selectively weed out the irrelevant information based on the user's preferences (Quiroga & Mostafa 2001). Although this has been said in a different context, it is clear that *cognitive relevance* is implied.

3.4. Situational relevance

Situational relevance describes the relationship between the perceived situation, work task or problem at hand and the usefulness of the information objects as perceived by the user. The criteria by which situational relevance is inferred are usefulness in decision-making, appropriateness of information in problem solving and the reduction of uncertainty.

According to Borlund (2000:42) '... the judgement of situational relevance embraces not only the user's evaluation of whether a given information object is capable of satisfying the information need, it offers also the potential of creating new knowledge which may motivate change in the decision maker's cognitive structures. The change may further lead to a modification of the perception of the situation and the succeeding relevance judgement, and in an update of the information need'.

Subjective relevance types, including situational relevance, are generally accepted to be both dynamic and multidimensional in nature. In the information seeking process, these relevance types are continually and interactively assessed. This assessment is not binary, but rather judged as degrees of relevance. In order for systems to support the searching behaviour of users in this context, it must allow for interactive information retrieval[See Borlund (2000) for the evaluation of such systems, and Savage-Knepshield and Belkin (1999) for a historical overview of trends in interactive IR (IIR)]

Situational relevance in a previous study (Cosijn 2003) was empirically found to be more strongly associated with work task execution than with search task execution. Therefore, interactive IR should also support searching over more than one session, and complex profiling should be able to dynamically include changing situational factors as well.

3.5. Socio-cognitive relevance

Socio-cognitive relevance is, together with cognitive, situational and affective relevance, regarded as a subjective relevance type. Socio-cognitive relevance describes the relationship between the situation, the work-task or problem at hand in a given socio-cultural context on the one hand, and the information objects on the other, as perceived by one or more cognitive agents. The social or organizational domain, or cultural context in which the individual finds himself is defined by a paradigm, which dictates what problem explanations may be found to be acceptable.

Retrieval of information limited to particular paradigms or socio-cultural or socio-cognitive domains may not be easily solved by improvement to systems. Facilitating serendipity or IIR may yield somewhat improved results, but in general the nature of socio-cognitive relevance is such that metadata would probably be the best solution to this particular problem.

The purpose of metadata is to describe the structure of the content data, and more importantly, to capture any additional properties that may characterise it. Metadata formats are divided into three categories: simple, rich and structured (Hakala 2004):

- *Simple formats* are proprietary and based on full text indexing. Search engine crawlers create this type of data. They are easy to use, but are weak for information retrieval purposes, as they do not support field searching.
- *Rich formats* are associated with research and scholarly activity, and require specialist subject knowledge to create and maintain. These formats are usually based on international standards, e.g. MARC (Machine-Readable Cataloguing), FGDC (Federal Geographic Data Committee), ICPSR (Interuniversity Consortium for Political and Social Research – an SGML codebook initiative describing social societies), CIMI (Computer Interchange of Museum Information), EAD (Encoded Archival Description) and CERIF (Common European Research Information Format).
- *Structured formats* are a compromise between simple and rich formats, specially developed for Internet usage. These include data that contain a detailed enough description to allow a user to assess the potential utility or interest of a resource without having to retrieve it. The data are structured and support field searching, but are still domain specific. Some structured formats are the IAFA (Internet Anonymous FTP Archive) templates; RFC (Internet Request for Comments) 1807 (format for bibliographic records); SOIF (Summary Object Interchange Format); and LDAP (Lightweight Directory Access Protocol) Data Interchange Format (LDIF). However, the Dublin Core Metadata Element Set (http://dublincore.org) is one of the first truly universal formats. This metadata element set is intended to facilitate the finding of electronic resources, originally conceived for author-generated descriptions of web resources.

The *de facto* standard for metadata, especially on the Web, is Dublin Core (DC). Dublin Core is a general set of metadata elements and is often enriched by application domain-dependent additions, such as the NDLTD (Networked Digital Library of Theses and Dissertations) and the LOM (Learning Object Metadata). The elements and definitions of DC are based on the official standard for the element set of DC (ANSI/NISO Z39.85-2001). The elements can be seen as describing three different dimensions of metadata, i.e. describing the content or *data*, describing the *source*, and

describing the *collection process* to collect the content. This subdivision is very important, since it describes the reality of the aboutness, isness and processing of the information objects Cosijn *et al.* 2002).

It is especially the data elements that are related to the source that may be of importance for improving access to *socio-cognitively relevant* information objects. Metadata elements such as the following DC elements have great potential to help users to judge the relevance of retrieved information objects with regard to a particular situation, or within a particular socio-organizational domain during the search task:

- Type: Nature or genre of the content of the resource
- Format: Physical or digital manifestation of the resource
- Identifier: Unambiguous reference to the resource within a given context
- Source: Reference to a resource from which the present resource is derived
- Language: Language of the intellectual content of the resource
- Relation: Reference to a related resource, and
- Coverage: Extent or scope of the content of the resource.

Another technique that may be used to facilitate socio-cognitive relevance judgements is that of co-citation analysis. Patterns of co-citation can help a searcher to understand which publications and authors may be grouped together in terms of their approach to a subject. This may then give an indication of acceptability within a particular socio-organizational domain.

An interesting study by Yuan and Meadow (1999) showed another possibility of improving access to socio-cognitively relevant documents. Authors in different fields use different words to describe concepts, for example *data* and *information* is used differently in the fields of computer science and information science. Yuan and Meadow (1999) found that when two individual papers, or two authors over several works, use the same variables (or terms), it indicates a *similarity in approach* to the subject. According to them, if authors use the same variables, 'such usage may be a stronger indication of similarity than co-citation because it represents what the authors did, rather than what they say' (Yuan & Meadow 1999: 147).

In traditional systems, both topicality and socio-cognitive relevance types were facilitated purely by human input. However, by using technologies such as described above, both these relevance types may be partially facilitated at a systems level.

3.6. Affective relevance

Affective relevance is described in terms of the relation between the goals, intents and motivations of the user and the information objects. Affective relevance should not be seen as the ultimate subjective relevance in a scale of relevances, but rather as another dimension of relevance judgments that may be associated with the other subjective types of relevance.

At this point it would be prudent to add a note on the time dimension encountered in the judgments of relevance by users. The phenomenon that relevance judgements changes over time has little bearing on algorithmic relevance, but as the relevance judgements become more subjective, changes in cognition over time have an increasingly profound influence on the dynamic process of interpretation, and are especially individualized in affective relevance.

As such, it is probably not possible to improve systems (other than profiling) or information representation to expressly facilitate this manifestation of relevance.

4 Conclusions

This study has aimed to improve our understanding of relevance by providing a model for understanding the concept of relevance in terms of relations between information objects on the one hand and the various aspects of the information seeking and retrieval process on the other.

In the historic development of IR as a field of study, three main research paradigms can be clearly identified – the systems approach, the user approach and the cognitive approach (Ingwersen 1999). Recently the emergence of a (tentative) fourth approach has become evident – the socio-cognitive or epistemological approach (Hjørland 2001).

Relevance may be regarded as the central and most fundamental concept within the field of information science (Froelich 1994; Saracevic 1996; Saracevic 1999; Schamber *et al.* 1990). We are studying relevant information, not just any information. As such, relevance should not be studied from a limited perspective. Systems may be improved by making their algorithmic relevance scores better correlate with the subject, but users judge relevance from a much broader perspective – not only from a cognitive perspective, but also within an epistemological framework.

The model developed and tested in a previous study (Cosijn 2003) and represented here defines the various relevance types and their interconnectivity. From the additional information provided on the various manifestations in this paper it

One of the most enduring debates within the systems approach to IR is the use of natural language versus controlled vocabulary to improve retrieval. A recent study in this field was done by Tomaiuolo and Packer (1998). A subset of this type of research is the work of researchers such as Sanderson (2000) on sense disambiguation. Other researchers concentrate on improving relevance feedback methods, for example the research by Voorhees (1998) on the role of assessors in measuring relevance feedback, Lee (1998) on multiple evidence from relevance feedback methods, Lam-Adesina and Jones (2001) on summarization techniques for term selection in relevance feedback, Voorhees (2000) on the validity of TREC for using relevance as a measurement of retrieval effectiveness and Voorhees (2001) on the role of highly relevant documents in system evaluation. Another recent area of research within the systems relevance is that of partial or graded relevant assessments, for instance the work of Järvelin and Kekäläinen (2000) on discounted cumulative gain which incorporates multiple relevance levels into a single measure and Kekäläinen and Järvelin (2002) on graded relevance assessments in IR evaluation.

The focus of the studies mentioned above is algorithmic relevance in the model derived in this study – the relation between the query and the information objects. Traditional Boolean systems facilitate binary relevance judgements, whereas best match systems, or a combination of best match and Boolean systems, are able to rank retrieved information by relevance. It is clear that even in systems relevance research there has been a move away from the traditional binary relevance judgements and a greater appreciation for the fuzziness of relevance judgements made by users and the need for interactive information retrieval (IIR). Therefore, research on retrieval systems improvement should focus more on facilitating fuzzy relevance judgements.

3.2 Topicality

Topical relevance is defined as the relation between the topic of the query and the topic of the assessed information objects. The finding of focus during the formulation of the request by the user, which is then transformed into a query by the system, is the criterion whereby topicality is inferred. The assumption is that both request and the objects may be assessed by a cognitive agent as being about the same or a similar topic, which implies a degree of subjectivity. The assessment is even less reliable if the information objects are represented by human-indexed terms.

Improving the relationship between the request and the information objects in terms of topicality is the focus of IR systems. Interesting new developments in the field of information representation, might prove to be useful in assisting users to judge potentially useful documents on a topical level.

Although not empirically supported, Ford's (1999) discussion of the possibilities offered by machine processing of similarities through high order knowledge representation and fuzzy (or parallel) IR is summarised here as a case in point.

3.2.1 High order knowledge representation

Relatively high order knowledge representations may be facilitated by linguistic analysis whereby similarity relationships at a relatively high level of abstraction can be made. A system such as DR-LINK '... can retrieve related articles that would not be found in a Boolean search because they contain the ideas, not the precise words, that were requested' (Feldman as quoted in Ford (1999:533)). This is still not enough, for current research, according to Ford (1999), is focused (*within narrow subject domains*) on:

- The computation of argumentation (components and structures of arguments are represented in such a way that patterns of argument and counter-argument may be mapped onto each other and compared for similarities and differences).
- Analogy-based representations and processing to support case-based reasoning (similarities are represented and then matched between stored cases of solved problems so that solution structures of known problems may be applied to new ones).
- The direct modeling of analogical reasoning (attempts to model human analogical reasoning to computers as well as commercial analogical problem-solving systems).

Information abstraction (structured knowledge representation of complex events, situations or relationships are created and then populated with text extracted from unstructured natural language texts).

The common thread in these studies is the specification of relationships between structural components at a level of abstraction higher than mere morphological or syntactical analysis, 'and of more complex semantic patterns than relative simple thesaural links' (Ford 1999:533).

3.2.2 Fuzzy and parallel IR

Often neural networks (employing fuzzy, parallel processing though pattern matching), focus on sub-semantic levels (e.g. image processing). Some systems, however, also use nodes to represent keywords and documents on a semantic level. Examples of these representations are taxonomies and ontologies (Welty & Guarino 2003) and topic maps (Pepper

2003). Knowledge of the relationship between query and documents is then stored in the pattern of links between the nodes (Ford 1999).

By using higher order knowledge representation and fuzzy and parallel IR, systems tend to become more intelligent. Although this type of research is relatively new, it is quite feasible that technologies such as those described above, may aid users in the judging of topically relevant information, by supplying wider information content than simply that which was requested through the query. Typical projects on these matters are for instance those related to sense disambiguation (Sanderson 2000), Park's (1996) work on inferential representation of documents within subject fields and Choi and Rasmussen's (2001) work on image retrieval based on topicality.

3.3. Cognitive relevance / pertinence

Pertinence is measured in terms of the relation between the state of knowledge, or cognitive information need of the user, and the information objects as interpreted by that user. The criteria by which pertinence are inferred are cognitive correspondence, informativeness, novelty and information preferences. For instance, a paper may be topically relevant but repeating what the user already knows. Cognitive relevance is clearly a very subjective judgment, as opposed to algorithmic and topical relevance as discussed above. The question on how to induce and facilitate the novelty value of information to users, must therefore by addressed on an entirely different level.

Traditional IR systems allow users to modify queries according to their own understanding of the problem. This, in turn, depends on the user's conceptual knowledge background and his understanding or perception of his information need. Toms (2002a) uses an interesting set of analogies to describe this aspect of seeking: 'Sometimes people seek a target with the precision of a cruise missile. Sometimes they seek a target with the imprecision of a Christmas shopper.'

The fact that the success of a query to retrieve cognitively relevant information depends on the user's understanding of both the system and the user's own problem space, tends to limit the possibility of the user finding relevant information. In recent research, however, there has been attempts to induce and facilitate serendipitous information retrieval. To continue with Toms' (2002a) analogy: 'Sometimes a target appears — unexpected and unsought, such as the five dollar bill fluttering in the fall leaves.'

According to Toms (2002a) there are essentially three ways to acquire information:

- Searching for information about well-defined and known objects.
- Searching for information about an object that cannot be described, but which will be recognized on sight.
- Accidental, incidental or serendipitous discovery of an object.

She contends that current information retrieval systems are based on the assumption that users know (or partially know) the object of their search, and that serendipitous information retrieval is largely ignored in information system development and research (Toms 2002a).

According to Figueiredo (2002), classic problem solving first requires a recognition of the problem, then some sort of divergence taking place and ultimately converging into a novel solution for the problem. Serendipity, on the other hand, is a creative process, whereby an attempt to solve a problem leads first to a divergence, and then to a new problem or a solution to a problem that was not known to exist. Serendipity is also defined by Quéau (quoted in Figueiredo (2002)) as 'the art of finding what we are not looking for by looking for what we are not finding'.

It is generally acknowledged that qualitative research sometimes contains 'good fortune', but according to Fine and Deegan (2002), serendipity consists in how this fortune is transformed into substantive discovery. Serendipity is therefore not only a 'chance encounter' (Toms 2002a), but more than that – it is the 'unique and contingent mix of insight coupled with chance' (Fine & Deegan 2002). Furthermore, Spink and Greisdorf (1997) found that highly relevant documents do not often change the user's cognitive or information space, but partially relevant documents do.

Serendipity rests on the three principles of insight, chance and discovery (Fine & Deegan 2002). The principles of chance and discovery could be built into systems, for example though improved browsing facilities (see Toms (2002b) for an example of such a system). However, the first principle, that of insight, rests solely with the user. To quote Louis Pasteur: 'Chance favours only the prepared mind' (Oxford Dictionary of Quotations, 1979).

Although the research focus of serendipitous retrieval is not necessarily that of helping users that cannot formulate their own information need satisfactorily, it is plausible that it may be utilised as an aid to users who cannot express their query to a sufficient degree. Research, such as that of Toms (2002a, 2002b), is very important in terms of the improvement of IR systems in order to assist users to judge relevance on a cognitive (personal) level.

Another important contribution within this focus of cognitive relevance judgements, is the research on profile building for information filtering. Coupled with browsing, personalization of information retrieval can help people to find information with potential value to their information needs. With regard to the Internet, Bowman *et al.* (1994:99) note 'at least 99% of the available data is of no interest to at least 99% of the users'. Personalization of information delivery relies

should be clear that these relevance judgements, either individually or jointly, may be and indeed need to be facilitated in some way by improving systems to make intelligent, interactive IR possible.

Through a literature review and meta-analysis, this paper is an effort to show how research in different areas of IR research is already moving towards improving access to information through facilitating users' relevance judgements when searching for information.

Relevance should be the one issue connecting the various approaches within information science. No single research paradigm should claim relevance for its own. In order to understand relevance, it is necessary to view the concept from a holistic perspective, taking into account the systems, the users, the cognitive overlaps of the role players within IR as well as the influence of the epistemological framework in which IR takes place. It is critical that future research in the field of IR should take all these factors into account.

Acknowledgements

This paper was first published in F Crestani and I Ruthven (Eds.): Context: Nature, Impact and Role. *CoLIS 2005, Lecture Notes in Computer Science (LNCS) 3507*, pp. 47-58 2005. Springer-Verlag Berlin Heidelberg 2005. It is reprinted here with permission from Springer.

References

Borlund, P. (2000) *Evaluation of interactive information retrieval systems*. Doctoral dissertation. Åbo: Åbo Akademi University Press.

Bowman, C.M., Danzig, P.B., Manber, U. & Schwartz, F. (1994). Scalable internet resources discovery: research problems and approaches. *Communications of the ACM* 37(8): 98-107

Choi, Y. & Rasmussen, E. (2001) User's relevance criteria in image retrieval in American history. *Information Processing and Management* 38: 695-726

Cosijn, E. & Ingwersen, P. (2000) Dimensions of Relevance. *Information Processing and Management* 6(4): 533-550

Cosijn, E. (2004) A methodology for testing dimensions of relevance. In: Bothma, T. & Kaniki, A. eds. *Progress in Library and Information Science in Southern Africa. Proceedings of the third biennial DISSAnet Conference (ProLISSA3)*. Pretoria. Infuse: 279-299.[Online] Available: http://www.dissanet.com . Accessed 21 February 2005

Cosijn, E. (2003) *Relevance Judgements in Information Retrieval*. Unpublished DPhil Thesis. University of Pretoria, South Africa

Cosijn, E., Pirkola, A., Bothma, T. & Järvelin, K. (2002) Information access in indigenous languages. In: Bruce, H. *et al*. eds. *Emerging frameworks and methods*. Proceedings of the Fourth International Conference on Conceptions of Library and Information Science. Greenwood Village, Colorado: Libraries Unlimited: 221-238

Figueiredo, A.D. (2002) The serendipity equations. In: Webber, R. and Gresse, C. eds. *Proceedings of the Workshop Program at the 4th International Conference on Case-based Reasoning, ICCBR01, Technical Note AIC-01-003*. Washington, DC: Naval Research Laboratory, Navy Centre for Applied Research in Artificial Intelligence. [Online:PDF]. Available: *http://www.aic.nrl.navy.mil/papers/2001/AIC-01-003*/ws4/ws4toc2.pdf . Accessed 21 November 2002

Fine, G. & Deegan, J. (2002) *Three principles of serendip: insight, chance and discovery in qualitative research*. [Online]. Available: *http://www.ul.ie/~philos/vol2/Deegan.html*. Accessed 21 November 2002

Ford, N. (1999) Information retrieval and creativity: towards support for the original thinker. *Journal of Documentation* 55(5): 528-542

Froelich, T.J. (1994) Relevance reconsidered – towards an agenda for the 21st century: Introduction to special topic issue on relevance research. *Journal of the American Society for Information Science* 45: 124-133

Hakala, J. (2004) *Internet metadata and library cataloguing*. [Online]. Available *http://www.lnb.lt/events/ifla/hakala.html*. Accessed 21November 2004

Hjørland, B. (2001) Brief communication: Towards a theory of aboutness, topicality, theme, domain, field, content and relevance. *Journal of the American Society for Information Science* 52(9): 774-778

Ingwersen, P. (1999). Cognitive information retrieval. In: Williams, M.E. ed. *Annual Review of Information Science and Technology (ARIST)*. Medford, NJ: Learned Information, INC. 34: 3-52

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

Järvelin, K. & Kekäläinen, J. (2000) IR evaluation methods for retrieving highly relevant documents. In: Belkin, N.J . *et al*. eds. *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Athens, Greece. New York: ACM Press: 41-48

Kekäläinen, J. & Järvelin, K. (2002) Using graded relevance assessments in IR evaluation. *Journal of the American Society for Information Science and Technology* 53(13): 1120-1129

Lam-Adesina, A.M. & Jones, G.J.F. (2001) Applying summarization techniques for term selection in relevance feedback. In: Croft, W.B. *et al*. eds. *Proceedings of the 24th Annual International ACM SIGIR conference on Research and Development in Information Retrieval*. New Orleans, Louisiana, USA. ACM Press: 1-9

Lee, J.H. (1998) Combining the evidence of different relevance feedback methods for information retrieval. *Information Processing and Management* 34(6): 681-691

Park, H. (1996) Inferential representations of science documents. *Information Processing and Management* 32(4): 419-429

Pepper, S. (2003) *The TAO of topic maps: finding the way in the age of infoglut*. [Online]. Available *http://www.gca.org/papers/xmleurope2000/ papers/s11-01.html*. Accessed 13 March 2003

Quiroga, L.M. & Mostafa, J. (2001). An experiment in building profiles in information filtering: the role of context of user relevance feedback. *Information Processing and Management* 38(5): 671-694

Sanderson, M. (2000) Retrieving with good sense. *Information Retrieval* 2: 47-67

Saracevic, T. (1999) Information science. *Journal of the American Society for Information Science* 50(12): 1051-1063

Saracevic, T. (1996) Relevance reconsidered '96. In: Ingwersen, P. & Pors, N.O. eds. *Information Science: Integration in Perspective*. Copenhagen, Denmark: Royal School of Library and Information Science: 201-218

Savage-Knepshield, P.A. & Belkin, N. (1999). Information retrieval interaction: trends over time. *Journal of the American Society for Information Science* 50(12): 1067-1082

Schamber, L., Eisenberg, M.B. & Nilan, M.S. (1990) A re-examination of relevance: Toward a dynamic, situational definition. *Information Processing and Management* 26(6): 755-776

Spink, A. & Greisdorf, H. (1997) User's partial relevance judgements during online searching. *Online and CDROM Review* 21(5): 271-280

Tomaiuolo, N.G. & Packer, J. (1998) Keyword and natural language searching. *Online*, November/December: 57-60

Toms, E. (2002a) *Serendipitous information retrieval*. [Online:PDF]. Available: *http://www.ercim.org/publication/ws-proceedings/DelNoe01/3-Toms.pdf*. Accessed 21 November 2002

Toms, E. (2002b) *Information exploration of the third kind: the concept of chance encounters*. A position paper for the CHI98 Workshop on Information exploration. [Online]. Available: *http://www.fxpal.com/ConferencesWorkshops/CHI98IE/submissions/long/toms/Default.htm*. Accessed 21 November 2002

Voorhees, E.M. (2001) Evaluation by highly relevant documents. In: Croft, W.B *et al*. eds. *Proceedings of the 24th Annual International ACM SIGIR conference on Research and Development in Information Retrieval*. New Orleans, Louisiana, USA. ACM Press: 74-82

Voorhees, E.M. (1998) Variations in relevance judgements and the measurement of retrieval effectiveness. In: Croft, B.C. *et al*. eds. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Melbourne, Australia. ACM Press: 315-323

Voorhees, E.M. (2000) Variations in relevance judgements and the measurement of retrieval effectiveness. *Information Processing and Management* 36: 697-716

Welty, C. & Guarino, N. (2003) *Supporting ontological analysis of taxonomic relationships. Institute for systems science* and *biomedical engineering of the Italian Research Council (CNR)*. [Online:PDF]. Available: *http://www.labseb.pd.cnr.it/infor/ontology/papers/DKE-2001.pdf*. Accessed 13 March 2003

Yuang, W. & Meadow, C.T. (1999) A study of the use of variables in information retrieval user studies. *Journal of the American Society for Information Science* 50(2): 140-150