

Investigating Document Triage on Paper and Electronic Media

George Buchanan, Fernando Loizides

Future Interaction Technology Laboratory
University of Wales, Swansea
{g.r.buchanan,csfernando}@swansea.ac.uk

Abstract. Document triage is the critical point in the information seeking process when the user first decides the relevance of a document to their information need. This complex process is not yet well understood, and subsequently we have undertaken a comparison of this task in both electronic and paper media. The results reveal that in each medium human judgement is influenced by different factors, and confirm some unproven hypotheses. How users claim they perform triage, and what they actually do, are often not the same.

Keywords: Digital Libraries, Interaction Design, Document Triage

1 Introduction

The activity of information seeking drives the use of digital libraries. Better understanding of the information seeking process unveils new aspects of user behaviour that can subsequently be used to improve the function and interaction of digital library services. This paper investigates the critical stage of information seeking called *document triage*. During document triage, a user examines a document and determines its relevance to their information need. This initial decision may subsequently be revised: the perceived relevance may rise or fall as the user comes to a fuller understanding of the piece. Our focus is purely upon the initial relevance decision that determines if this subsequent reading will occur.

Previous studies (e.g.[12, 13]) have identified which elements of documents are used to make relevance judgements during interactive information retrieval. These experiments have elicited subjective information from users as to which document properties they use when making relevance judgements. The findings of this foundational research have gradually accumulated to give us a consistent set of document features that play a key role in relevance decision-making: e.g., when searching academic papers, users often first refer to the document's abstract [12]. Researchers have also studied many differences between paper and electronic texts, particularly in the case of reading (e.g.[11], and how users scan search result lists [16] but document triage has received little attention.

Our goal is to improve the effectiveness of relevance decision-making in electronic environments. We already know what document properties provide the key

features for document triage. However, users explore documents interactively, not simply as passive abstractions with particular properties. During triage, documents are explored and evaluated in a short timespan, and we do not yet understand *how* users encounter and *process* this data.

There is good reason to scrutinise these interactive issues, particularly in digital systems. Experiments have explored the use of scrolling and interaction data for relevance feedback[6]. Further progress requires that we better understand the interaction itself to provide a foundation for progress with that approach. Likewise, users are increasingly using digital documents and computer-based reading[7]. Improvements to the document reader software through which this reading and judgement is being made can increase the total efficiency of digital information retrieval, by enhancing the human performance just as information retrieval research advances search engine capability.

This paper reports on the experimental investigation of the differences between the document triage process on paper and electronic media. By providing a contrast of the two forms, we gain a deeper knowledge both of what users do during information triage, and how they interact with the documents that they evaluate. The paper proceeds in five parts: first, we briefly examine some key features of the existing literature; second, we describe the experiment itself, before proceeding to a discussion of our findings, qualitative and quantitative; subsequently, we reflect on the findings in the light of existing work, before concluding with a summary of the key findings and their impact for future research.

2 Related Work

Information triage is the activity where a user determines the relevance of a piece of information for a particular information task [8]. This decision can be made at various stages of the information process: e.g. on first encounter, when it is selected for detailed reading; on review after closer reading; or subsequently when it is considered for long-term retention or re-interpreted in the light of other material. Document triage is encapsulated in the early stages of triage, when basic relevance is assessed. Though later reflection may result in its later omission from a user's final repertoire of documents, all this later activity is contingent upon the document's initial acceptance.

Models of document relevance have been considered, contested and discussed over many years, focussing on the properties of documents: for instance, the relative significance of titles or abstracts [12]. Such studies have been focussed upon relevance as a property of a document, to be assessed, rather than upon the method and means of that assessment. Likewise, many studies, such as [3], have viewed relevance and related document features within the framework of information retrieval: namely, how to operationalise relevance computationally.

Only recently has attention within computer science moved from document properties and information retrieval mechanisms to the corresponding human processes. Recent papers by Badi[1] and Bae[2], Wacholder[15] and others have

pointed the way to a more detailed consideration of the interaction that occurs during document triage, in electronic and physical information media.

3 User Study

In order to uncover the gaps in existing understanding, we undertook a user-study to explore the differences between document triage behaviour in electronic and physical environments. Recent research[7] demonstrates that significant shifts in user behaviour emerge with a move from physical to digital media, thus a better understanding of the current differences is timely. Existing research provides triangulation data for new findings. This section first describes the experimental design, before discussing the quantitative and qualitative results.

3.1 Experimental Design

Information triage is a multi-stage process, as is the even more specific task of document triage. One common distinction in the digital environment is between reviewing the document through: its full-text; a descriptive overview page with title, abstract, etc.; or at the result list of document titles. We subsequently studied two electronic conditions: the first requiring the user to open the full document; the second giving a descriptive page that listed title, abstract, author and publication information. In all conditions, including paper, a printed list of document titles was given, and in both electronic forms a result-list format of document titles appeared at the commencement of the experiment.

The experiment was conducted in three conditions: first, paper-based with each participant being given a list of document titles and a set of printed documents; second, digital documents accessed via a results list and overview page; third, digital documents accessed as PDFs from a single file folder. In the digital folder condition, the files appeared with their titles, rather than filenames.

The study was conducted on a standardised task: judging the relevance of 20 documents from the ACM Digital Library on a set topic. The search for this study was standardised, and thus every participant viewed the same twenty documents. For the paper condition, these papers were printed double-sided on A4 paper. Digital documents were presented as PDFs, read using Adobe Acrobat.

Thirty participants were recruited: 10 for each condition. Participants completed a pre-study questionnaire, and a post-experimental interview was used to elicit particular details of their interaction with the documents. Video recordings were taken of all participants' sessions, and the on-screen activity captured by Blackberry Flashback for the digital sessions. Each participant was asked to score each of the twenty documents for relevance to the task, giving a rating out of 10 (completely relevant) to 0 (non-relevant).

Participants were computer science staff and post- and under-graduate students, aged from 18 to 52. Of the thirty participants, twenty one were in their final undergraduate year or were postgraduate. Participants were assigned to each condition to ensure balance between the different modes. The task topic

was of variable familiarity to the participants: from two who were research active in the area, to four first-year undergraduates with only limited knowledge.

A panel of three domain experts examined each paper in detail and rated each as relevant, part-relevant or non-relevant, and scored out of ten for relevance. Where disagreement occurred (re. two papers), a discussion was engaged with until consensus was found. This basic design reflects the common practice for assessing relevance for experimental corpora in the TREC conference series[14].

3.2 Findings: Process

Before studying the interaction of users with individual documents, we will first review the behaviour of users during their relevance judgement as a whole. Though there were some common patterns, in fact there were a number of critical divergences in the process followed by participants in the three different modes.

One key difference was total time taken for document triage: the ten paper-mode participants taking a mean of 23m 42s, overview-page mode 17m 11 s, and folder-mode 28m 12s. A t-test of these differences was significant at the 10% level for all cases, except between the two electronic modes, which was significant at the 5% level ($p=0.021$). One reason for the low times of overview-page participants was that few papers were read in their full-text form. Only two participants opened more than three papers for full viewing. In contrast, in both paper and folder modes 96% of all papers were read as full documents.

Process Strategy The most common approach to the triage process was a linear process of reading the first document and then reading and scoring every document in turn. This is the dominant case both with the participants in the paper and electronic groups (25 of 30 participants). Those five participants using a non-linear method were all found to be in the eleven participants who took the longest total time. One potential weakness of the linear approach, identified by six participants, is that an individual's use of relevance scores can change as further documents are read. On paper, however, the eight linear-method participants often paused to reconsider their current scoring, whilst this was not observed at all in the two electronic modes: one participant commenting that "I realised at some point that the titles are kinda wrong sometimes in the meaning, so some of my scores in the beginning might be wrong but there we go".

Organising Documents Our paper-mode participants frequently used the available deskpace to organise the scored documents in piles. One participant organised the documents at the start of the process, spreading them over the desk in order to gain an overview of the whole list: "If I can see the titles all at once, then I can judge more relevant documents and save time". Most participants, however, categorised documents as they were read in turn, reflecting the perceived relevance (see Figure 1, left). Seven out of ten participants divided the (tentatively) scored documents into piles, each pile being a set of documents with the same ranking category. The most common practice (5 users) was three

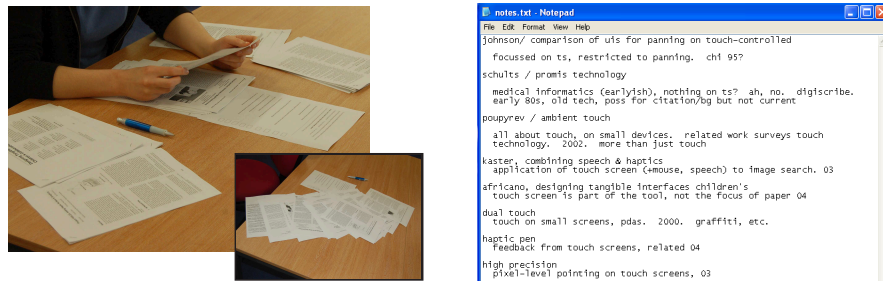


Fig. 1. Users organising pages for triage (left), fanning document (inset left) and electronic note taking (right)

piles: relevant, non-relevant and part-relevant documents. Six participants used these piles at the end of the task, placed across the desk as they confirmed their final relevance scores. Four of these reviewed all documents at this stage, and subsequently gave each a revised score in light of the scores given to other papers. One participant used a notable fanning effect (see Fig.1, inset) produced by the relevance of the documents. When asked why he did it he commented that "they sort of fanned out without me realising". This fanning effect was not reproduced in any way by any of the electronic or folder participants.

The same set-creating activity was seen in two participants in the electronic modes. One marked potential relevance by placing a mark next to the paper title on the scoring sheet, whilst another kept score on a Notepad document (Fig.1, right). These participant thus created virtual document sets and proceeded to go back to the relevant documents in more depth to verify their judgements.

Average relevance scores were very similar across the different modes (6.356 and 6.512 in the digital and physical conditions respectively), and show no statistically significant difference ($p=0.74$).

Triage Effectiveness The 20 documents used in the experiment were rated by our expert panel as containing 7 relevant, 5 partially relevant and 8 non-relevant documents. We evaluated the performance of our participants by considering the number of relevant documents in their 7 most highly ranked choices. This score showed little deviation across all modes, with 128 of the 210 documents (61%) in question being considered relevant by our experts. Paper and folder modes scored a little higher (63% each) and the overview-page mode lower (57%). These differences are not statistically significant within our sample size.

However, this picture hides considerable variation between modes, and we shall meet a number of these differences as we study how users viewed and evaluated the document content. Overview-page users highly rated short papers: three their four top-rated papers were 4 pages or less, whilst all top-ranked document in folder or paper modes were more than four pages long. Folder participants favoured documents with more statistical tables and technical diagrams.

3.3 Document Content

Abstracts and Titles Previous research has indicated that document abstracts and titles play a key role in the relevance judgement of human readers[12]. This was true of all participants, particularly those in the two electronic modes.

Overview-page participants frequently (157 of 200 documents) viewed *only* the overview page of document title, abstract etc. provided after clicking on a document in the list. One participant opened all 20 documents, another ten, five opened documents that did not provide an abstract on the overview page, and the remaining three viewed no full documents at all. When documents were opened, the users performed only limited scrolling: suggesting that only the abstract and title from the top of the documents was read in detail, even with two page documents. Folder-mode participants also spent a high proportion of their time on the first page of the document, as we shall see below. Differences also appeared regarding the conclusion pages of documents: electronic participants rarely displaying them (total 17 of 243 opened documents), whilst reading was common in the paper mode (89 from 173 read documents).

Subjective feedback reinforced this observational data: e.g. twenty six participants reported that the title was important, and they favoured a “title including the search words”. Imprecise abstracts were criticised by most of these participants, but this was not always acted on: eight overview-page participants noted specific abstracts that they found vague, but we observed that subsequently they did not chose to open any of these documents in question for reading.

Headings and Emphasis Existing literature reports a common focus upon section heading text, which we also observed in all modes. However. participants reported seeking all emphasised text: e.g. italicised paragraphs, bullet points and figure captions. Participants commented that though they first look for key words in titles, later attention is given to headings in general. Post-experimental feedback and talk-aloud comments revealed that though title text was an initial attractor during triage, in all participating groups the initial importance of titles declined as heading and caption text was scrutinised. Participants from every group observed that titles can be misleading, and as the task progressed they therefore stopped relying so heavily on the titles alone. Yet again, few participants subsequently reconsidered the documents that were already ranked. Only two paper-mode participants who made this observation returned to review the score they had given earlier documents. Participants frequently paused at length on pages containing large amounts of bold text, whilst few pauses of over half a second were recorded on pages with no emphasised content.

Images Images and figures play a very important role in general to both the paper and the electronic folder participants. Sixteen of these twenty participants rated images as “useful” or “extremely useful” when deciding the relevance of a document. One participant commented that “A picture is a thousand words they say. I think it is much more”. Subjective feedback also revealed a favourable

emphasis on concise, easy to read diagrams, when these support the relevance-judgement decision: e.g. “Pictures are the main thing I look at. You can pick up the meaning and get the general idea of the paper I think from them”. Participants also reported scanning neighbouring and caption text. In contrast, the seven of the ten overview-page participants rated images “not important” for relevance judgements. This difference in reaction could well be a product of the readers’ behaviour or received stimuli from the different experimental modes.

We subsequently evaluated the relationships between images, tables and technical figures and users’ relevance judgements. Photographic images seem to have no discernable effect: indeed, there is a moderate negative correlation ($r=-0.31$) between the number of images in a document and its relevance score. This relationship is explicable by document length effects (see below). Technical figures and tables, on the other hand, have a positive correlation between the number of figures and relevance rating ($r=0.34$), which is particularly noticeable in the case of those in the electronic folder mode ($r=0.57$), and statistically significant at the 5% level in both cases. Statistical analysis thus concurs with the subjective feedback: when reading full- electronic documents, technical illustrations and tabulated data play significant roles in relevance judgement.

Document Length Document length played an important part in document review. Even in the paper mode, where participants would often glance at most, if not all, document pages there was a declining likelihood of text being read when it fell later in the document. For example, in the case of one 24-page document, only one reader was observed as having viewed every page in paper mode, whilst not one reader in the electronic mode read beyond the second page.

A comparison of the paper-only against the two digital-mode groups illuminates differences in relevance judgements. From our experimental data, there are statistically significant effects from document length: shorter papers being favoured in the digital domain, when compared to paper ($p=0.05$). Further scrutiny between the two digital conditions reveals further detail: those using the overview-page interface were most likely to score longer documents lower – often without actually reading them. Individual papers demonstrate these general differences: e.g. paper no. 16 (4 pages) moved from rank 13 (paper) to 4 (digital), and its mean score average rose from 5.4 to 7.5.

Within-Document Navigation Document navigation methods varied notably between the physical and electronic modes. In electronic modes, many documents (34%) were never scrolled, and 64% not read beyond the first page. Subsequent to this initial view, users would often rapidly the document downwards before unpredictably and apparently randomly skimming the document. Our participants explained that they were searching for relevant-looking headings. In contrast, those searching on paper repeated a superficially similar pattern, reading the first page in some detail, before skimming linearly through the document, often with a particular pause on the page containing the paper’s con-

clusions. This repeated the emphasis on the beginning and end of the document, with the modification that the body was typically read before the conclusions.

One method of within-document navigation only available to electronic documents is the search function. Subjectively, this feature was rated highly by all our participants, twenty four of whom cited “Ctrl-F” as a key advantage of reading on the PC. The search facility of a reader application could be used to identify the specific parts of the document that contained the search terms used in the original query. Fifteen digital-mode users reported the advantages of search, and all claimed to use it. However, the observational data is at odds with these claims. Out of a total of 243 documents read in full PDF form, search was only used on 11 occasions by 4 users. Furthermore, on the majority (8) of these occasions, the first match only was inspected.

Differences also appear if we evaluate the amount of time spent in different parts of the document. Taking data from the screen-captures for the participants in the electronic folder mode, and the video capture of the paper mode, we can compare the displayed time for different parts of the document. Those in the electronic mode spent 68% of their viewing time on the first page of the PDF, without scrolling (i.e. on the top, visible portion of the first page). In contrast, only 32% of their time was spent on the remainder of the document, nearly equally divided between scrolling activity (15%) and the stationary reading of content (17%). Interviews revealed that most of the content time was focussed on larger visible elements such as headings and images. This contrasts markedly with the user behaviour when interacting with paper, with an average of 47% of time spent on the entire first page, with the remainder spent on the rest of the document. However, between-page movement accounted for a small percentage of time (< 5%) and reading for nearly 50% of total document viewing time.

Users in overview-page and folder modes often continually scrolled documents up and down without pausing. It is hard to compare this mode of reading with the serial reading of pages seen in our paper-based participants, but the data above shows that the paper mode participants spent longer on the second and subsequent pages of documents. The conclusion page of papers was a common navigational destination in full-text reading; however, this was viewed for 52% of read papers in the paper mode, and for only 17% in the folder mode. Of the twenty participants in these two modes, only five stated that the conclusion played a notable part in their relevance judgement. In the overview-page mode only 43 papers were read, and only six were navigated to their conclusion page.

In electronic modes, following the skimming review of a document, users usually (72% of documents) return again to the top of the document. On the other hand, paper participants pause at several points in a document. Pages containing diagrams or illustrations are viewed for a longer period of time (average per viewed image page of 10.3s compared to 4.8s in electronic modes). Users’ subjective feedback reported stopping to read the text relating to the images, headings, and sentences from paragraphs that are considered important. Readers seldom returned to the abstract page (24% of papers) at the conclusion of reading. Four paper participants reported that they would even stop to read

something of interest that is not necessarily related to what they were looking for. Twenty two of all participants reported that recalling location of material within a document was easier on paper, and reading within documents seems more systematic (linear) in our observations of participants' reading.

Subjective Feedback Participants were asked whether they would prefer to use the other medium (e.g. electronic/PDF for paper participants) to perform a search, or if they would prefer to use both physical paper and electronic media simultaneously. Only 3 (10%) of participants would chose to swap medium, whilst 10 (33.3%) would choose to use both media at the same time.

In order to discover the different affordances of the two media, participants were asked to list the features of each medium they find useful. In addition to those already mentioned, 14 participants reported greater comfort for paper, 11 noted annotation was easier, 10 said it was easy to organise (e.g. into piles) and nine cited it as being more portable. In comparison, few advantages were cited for electronic media did not have so many points in its favour – the only advantage reported by more than two people was the search facility.

When asked specifically about annotations twenty three participants considered them very important for triage but only one of the twenty digital participants took annotations of any kind. Participants commented that they could “write their own annotations for future searches” and that annotations help in organisation or, as one participant put it, “collecting my thoughts”.

4 Discussion

As noted in Section 2, there is a gathering body of evidence on user behaviour in document triage. Some of our findings strongly triangulate with this existing literature: e.g. the significance of titles and the dependence of readers upon abstracts of academic texts is already well documented [12, 4]. Our participants were entirely consistent with this recorded focus upon these key document features.

Digital document triage has also been reviewed, often in the context of web searching[13]. However, such studies have either focussed upon generalised web searching, or validated electronic information seeking criteria against the previous literature. This paper has focussed upon how these known items are actually used during the triage process itself.

Document length is one property that carries across physical and electronic environments, if experienced in differing ways: e.g. download times – a known concern on the web – does not translate to paper. Reviews of relevance literature[9] make clear that longer documents are less likely to be accepted for casual information seeking, whilst shorter documents are less likely to be accepted when “authoritative” data is being sought. Overview–page participants scored longer documents lower than participants using paper. Given that the participants were given the same task in both modes, and the experiment balanced for experience and other effects, it is reasonable to assert that this difference is a property of the electronic environment.

Other recent work (e.g. [7]) suggests that electronic reading is associated with more casual styles of reading, when compared to paper. Our results suggest a similar difference in terms of triage: a preference for shorter documents, and a high proportion of skimming activity are more commonplace in the digital domain. When given a summary page of a document, readers in the digital domain seldom move beyond it, and even when reading full-texts, readers seldom scroll beyond the first screen of text. In comparison, over 80% of paper documents were read beyond the first page. What is not clear is the cause of this variance: is it a direct consequence of the interactive affordances of reader software, or indirectly caused by different expectation of digital resources?

4.1 Factors Affecting Relevance Choices

Users in the electronic modes made lower relevance decisions for longer documents. Whilst classic relevance ratings have viewed relevance in purely semantic terms, it would appear that in practice users adjust their relevance judgements when considering other factors. One model for this is to consider that a user's perceived relevance for a document is factored by the perceived cost of reading the document. In our case, wider navigation of paper documents suggests a lower cost than in electronic modes. In line with other research[5], our participants behaved as if electronic reading were a higher cost than reading on paper. Our earlier research has demonstrated that interactive costs for within-document navigation and reading can be reduced through changes to interaction design[5], though triage reading is substantially different to deep reading[1].

Our experiment confirms again the importance of document abstract and titles, but adds that interaction affects the impact of these elements. Participants, even when obliged to download the document, spent a high proportion of their time on first view provided by the document reader, and additional reading had little impact on user's relevance precision. Janes[4] used different presentations to evaluate the significance of abstracts and other elements of documents for relevance judgements. His conclusion was that abstracts formed the most important element of a user's decision, but his experiment used pre-defined formats, rather than allowing users to explore at will. What we observed during the interactive retrieval process was that when insufficient data was gleaned from an abstract, a higher volume of detailed reading proceeded. Furthermore, readers would often refer back to the abstract after linearly reading the paper for further detail. User confidence in the abstract is therefore tested by interaction, and doubt in the abstract drives users to higher interaction.

Time is also critical: initial response to a document (e.g. of the abstract or images) appears to be dominant. Our folder-mode participants were strongly influenced by images, but we have no evidence that this judgment is any more open to revision than the enduring impact of abstracts.

Paper-based participants showed more annotational and organisational behaviour than our overview-page and folder users – suggesting that these activities are easier on paper than in Acrobat.

Nicholas et al[10] observed, through a deep log analysis of major online journal repositories, that when an abstract webpage was accessed before the document itself could be downloaded, users often did not continue to the download stage. Their observation was clouded by issues of subscription – i.e. not all visitors *could* download full documents, and many could only view the abstract page. We can confirm Nicholas et al’s hypothesis.

4.2 Future Work

Our experiment identified further details regarding user interaction with documents during document triage. However, there are some natural limitations to the experiment that mean many issues require future study. First, we were not able to track the actual parts of the screen used by the users. Eye-tracking may uncover further details that give a more certain insight into the fine-grained viewing processes of readers. Secondly, the focussed task used here is clearly only one of a nearly infinite range of tasks that could be used. Other tasks on different topics, using different user groups, and studying different stages of the information seeking process – from initial topic investigation to the double-checking of fine detail – may reveal different patterns and behaviours.

In particular, this experiment suggests that full-document review is more time consuming in electronic forms than on paper, and that any electronic form yields a lower level of precision than is the case for paper. Further experimentation is certainly required to identify where these differences do and do not apply. The experimental design also meant that we did not gain any insight into the effect upon the number of documents selected in any mode, and this is clearly an issue deserving of further study. The long history of research into relevance judgements[9] clearly demonstrates that this complex and important task cannot hope to be resolved by any one study.

5 Conclusion

This study revealed that many stages of the triage process differ between paper and electronic environments. Annotation and organising behaviour is rare during the electronic triage process, yet is commonplace in the medium of paper. Navigational behaviours also vary, with those reading on paper showing a broader reading of the content of papers, evidenced by a longer viewing time of content beyond the first page. In contrast, where a summative page of basic information is given in an electronic environment, document content is seldom used to shape the initial relevance decision. When full documents are read in an electronic environment, the larger proportion of time is spent on the first page, and much less on viewing of the subsequent content. This had significant impacts on the relevance judgments of our users, with certain features, such as greater length, leading to erroneously low relevance ratings. Users also seem to place high value on their first impression of a document, and further reading of the wider text has limited impact. For effective triage, initial reading needs to be concentrated on the most salient parts of the document.

6 Acknowledgements

This research is supported by EPSRC Grant GR/S84798.

References

1. R. Badi, S. Bae, J. M. Moore, K. Meintanis, A. Zacchi, H. Hsieh, F. Shipman, and C. C. Marshall. Recognizing user interest and document value from reading and organizing activities in document triage. In *Procs. IUI '06*, pages 218–225, New York, NY, USA, 2006. ACM Press.
2. S. Bae, R. Badi, K. Meintanis, J. M. Moore, A. Zacchi, H. Hsieh, C. C. Marshall, and F. M. Shipman. Effects of display configurations on document triage. In *Human-Computer Interaction - INTERACT 2005*, pages 130–143. Springer, 2005.
3. C. Cool, N. J. Belkin, and P. B. Kantor. Characteristics of text affecting relevance judgments. In *Procs. 14th National Online Meeting*, pages 77–84. Learned Society, 1993.
4. J. W. Janes. Relevance judgments and the incremental presentation of document representations. *Information Processing & Management*, 27:629–646, 1991.
5. M. Jones, G. Buchanan, and N. Mohd-Nasir. An evaluation of webtwig - a site outliner for handheld web access. In *Procs. International Symposium on Handheld and Ubiquitous Computing (HUC)*, pages 343–345. Springer-Verlag, 1999.
6. D. Kelly and N. J. Belkin. Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. In *Procs. 24th ACM SIGIR Conference*, pages 408–409, New York, NY, USA, 2001. ACM Press.
7. Z. Liu. Reading behavior in the digital environment. *Journal of Documentation*, 60:700–712, 2005.
8. C. C. Marshall and I. Frank M. Shipman. Spatial hypertext and the practice of information triage. In *HYPertext '97: Proceedings of the eighth ACM conference on Hypertext*, pages 124–133, New York, NY, USA, 1997. ACM Press.
9. S. Mizzaro. Relevance: The whole history. *Journal of the American Society of Information Science*, 48(9):810–832, 1997.
10. D. Nicholas, P. Huntington, H. R. Jamali, and A. Watkinson. The information seeking behaviour of the users of digital scholarly journals. *Inf. Process. Manage.*, 42(5):1345–1365, 2006.
11. K. O'Hara and A. Sellen. A comparison of reading paper and on-line documents. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 335–342, New York, NY, USA, 1997. ACM Press.
12. T. Saracevic. Comparative effects of titles, abstracts and full text on relevance judgments. *Journal of the American Society for Inf. Science*, 22:126–139, 1969.
13. A. Tombros, I. Ruthven, and J. M. Jose. How users assess web pages for information seeking. *Journal of the American Society for Information Science and Technology*, 56(4):327–344, 2005.
14. E. M. Voorhees and D. K. Harman. *TREC: Experiment and Evaluation in Information Retrieval*. MIT Press, 2005.
15. N. Wacholder, L. Liu, and Y.-H. Liu. Selecting books: a performance-based study. In *Procs. 6th ACM/IEEE-CS Joint Conference on Digital libraries*, pages 337–337, New York, NY, USA, 2006. ACM Press.
16. A. Woodruff, R. Rosenholtz, J. B. Morrison, A. Faulring, and P. Pirolli. A comparison of the use of text summaries, plain thumbnails, and enhanced thumbnails for web search tasks. *JASIST*, 53(2):172–185, 2002.