

# An Empirical Study of User Navigation during Document Triage

Fernando Loizides and George Buchanan

Future Interaction Technologies Laboratory, Swansea University  
Centre for HCI Design, City University  
csfernando@swan.ac.uk, george.buchanan.1@city.ac.uk

**Abstract.** Document triage is the moment in the information seeking process when the user first decides the relevance of a document to their information need[17]. This paper reports a study of user behaviour during document triage. The study reveals two main findings: first, that there is a small set of common navigational patterns; second, that certain document features strongly influence users' navigation.

## 1 Introduction and Motivation

When seeking for information in a digital library, users repeatedly perform the task of “document triage”, where they make an initial judgement of a document’s potential relevance to their information need. There is a wide body of associated literature that gives indirect evidence about document triage. However, the scientific understanding of document triage specifically is limited.

The consistency and accuracy of human relevance decision making is often poor compared to nominal ‘perfect’ choices for known tasks on closed sets of documents [4]. Human information seekers seem to rely at least in part on search engines to ‘predict’ the most accurate documents for their information needs. Following a search, there is minimal tool support for the relevance decision-making of users. Observed behaviour suggests that the user’s attention is narrowly focussed on a few parts of particular documents. For example, in a search result list, users mainly focus on documents presented on the first or second page [20]. Similarly, users choose only a few documents for further scrutiny, and appear to read little in detail beyond the first page [4]. This suggests that users are using only a limited pool of information to make relevance decisions.

Document triage can be viewed as a form of visual search. In computer-human interaction, the visibility principle asserts that the important aspects of a system must be visible [11]. Applying this principle to documents, one can argue that users will be influenced by what is displayed of a document. We wished to explore whether the factors suggested in the current literature (e.g. [7,19]) did indeed influence how users move and view actual documents.

Information seekers have their own mental models or “personal information infrastructures” for systematic information seeking [16] in which they follow certain visual cues. There is a healthy diversity of information seeking models

(e.g. [8,13,16]), however none of these describes document triage in detail. The current research into triage is fragmented, and this no doubt explains the lack of such a model. Available reports focus on very specific aspects or external aspects rather than document triage generally [1,2,3,5,18].

The research reported in this paper is a laboratory-based study that provides an initial assessment of the impact of common visual document features on on user behaviour during document triage. While some features focus user attention, the evidence is that conversely some relevant material is overlooked. When such issues are better understood, designers of document reader software can improve support for triage work.

We now report the study in three parts: design, results, and discussion. The paper concludes with a summary of key findings, and a view of future research.

## 2 Study Design

In general form our study was a laboratory-based observational experiment. Participants performed document triage on a closed corpus of documents, evaluating each for its suitability for a particular task. Log data of their interaction was captured. Pre- and post-study questionnaires plus an interview collected further data for comparison with the participant's actual behaviour during the experiment. This general structure was chosen to permit us to systematically assess the interaction between visible document content and visible document cues (e.g. headings, diagrams) and both user behaviour and relevance judgments.

### 2.1 Apparatus and Participant Selection

We developed a bespoke PDF reader software, that logs detailed information on the user's navigation in an XML file. Our study used a set of PDF files that were examined by the participants. To analyze the navigation log data accurately, we extracted the content from each PDF file and passed it through a custom-built parser. This permitted us to identify the exact content visible at any moment, using the view position and zooming information. The parser identifies heading, image and other visual feature information. An initial run on the document data was validated by human visual inspection to verify that the output was accurate for the set of documents we were using.

We recruited twenty participants, all having previous experience of PDF reader software. The participants were studying at postgraduate level in a computer science discipline, avoiding the specific CHI sub-disciplines that appeared in the study material. The participants' ages ranged from 21 to 38.

### 2.2 Study Format

Before the study, participants reported their perceived importance of common document features (title, heading, etc.). This data calibrated our findings with other work (e.g. [7,19]). Overall, our data aligned closely with earlier work.

Participant responses indicated that a few features (e.g. title and heading text) would strongly influence most users, whilst many features (e.g. body text) possessed only moderate influence. We will report individual results later, when we compare observed behaviours against subjective self-reporting. We also elicited previous experience with document reader software.

Participants were then given two information seeking tasks. The subject area of the tasks were not the speciality of the participants, but was sufficiently close to their domain to be intelligible. It was stressed to the participants that no prior knowledge of the subject was assumed, and their personal performance was not being assessed. The first information task was to find material on the interfaces of tablet PC's, the second was to find papers on specific CHI evaluation methods. The participant was then given a set of documents to evaluate against the given information need. They were instructed to give each document a rating out of 10 (1 for non-relevant documents, 10 for extremely relevant documents). The articles ranged from short papers (2 pages) to full journal papers (29 pages).

Users were given an open-ended timeframe to complete their tasks and could ask any question to clarify the task objectives. They were supplied with a dictionary should unknown words need clarification. No guidance was given regarding the papers themselves. The entire computer display was recorded using screen capturing software (BB Flashback<sup>1</sup>). A post-study interview was undertaken to obtain a better understanding of the participants' actions and thoughts. This interview was semi-structured with standardized questions about certain visual features (e.g. headings). Individual actions noted by the observer were also probed to elicit a better understanding of the user's intentions.

The corpus of documents was selected to permit a more detailed picture of the impact of specific document content and visual features. Whilst the majority had, broadly speaking, similar numbers of headings and other features per page, there were outliers that could reveal the impact of extreme cases. For example, one document contained text in only one typeface, size and weight.

### 3 Results

In this section we initially describe *navigational patterns* that we observed participants use during the document triage process. We then study what was actually visible during the triage process. Attention (i.e. viewing time) for different content is, naturally, not the same, and we report the relationship between document content and visual features and the time for which that content was visible. Finally, we report the participants' relevance judgments, and how these relate to the navigational behaviour identified in the study.

#### 3.1 Document Navigation

From the viewing log data, we extracted a number of high-level navigation patterns that recurred in the participants' viewing of documents. Despite being

<sup>1</sup> <http://www.bbsoftware.co.uk>

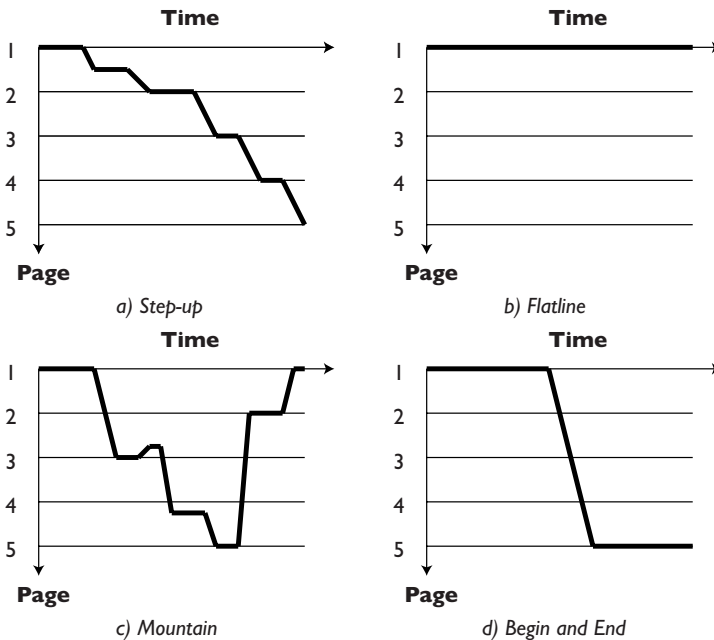


Fig. 1. Navigational Behaviours

given a free choice of action, all participants viewed documents in the same order, proceeding through the list of documents in a linear fashion.

We then studied the participants' viewing behaviours within documents. The first few seconds of every document log was linear, but from that point variations quickly emerged. These are now reported in descending frequency order. Figure 1 presents caricatured timelines to illuminate the patterns visually.

**Step Navigation** is our term for the predominant viewing pattern, found on 229 document passes, and in all 20 participants. As seen on Figure 1a, the initial part of the document is viewed for a prolonged period without any movement. In comparison, the other parts of the document receive periodic spans of attention, and many only viewed momentarily during rapid scrolling.

The initial view contains the title, abstract and, on most occasions, a large part of the introduction. (The precise figures of how long users remained on different parts of the document is discussed in subsection 3.2). The conclusion section of the document was often the only part of the end of the document that received any attention. Allowing for very small (< 10 pixel) backtracks that are probably a result of mouse manipulation errors by the user, the user travels linearly through the document, but with attention unequally spread.

The second most common behaviour was the **flatline** pattern (see Figure 1b). Users simply scrutinize the first page or visible part of the document. In contrast to the step pattern, there is no subsequent scrolling to later document content. This was observed on 73 documents and in 18 users.

The third pattern was for the user to scrutinize the document from the beginning to the end – in the same manner as the step-up pattern – and then backtrack through the document towards the beginning, producing a **mountain** pattern (see Figure 1c). In our specific example we see the user returning to an already viewed part of the document to view it for a longer timespan (and, presumably, in greater detail). In eighteen cases, the user returned to the very beginning of the document for a final view. Overall, the mountain pattern was observed 23 times across 12 individual participants.

A last behaviour observed is the **begin-and-end** pattern (Fig. 1d). This is when the user scrutinizes the beginning of the document before directly scrolling to the conclusion of the document without any significant pause to viewing other document parts. Variations arise in backtracking to locate the actual conclusion, but in essence only two parts of the document are viewed. This behaviour was observed 33 times across all 20 users.

The models discussed above are nominal perfect models. Users often did not follow one of the models perfectly, but rather incorporate behaviours from two or more models to create hybrids. Furthermore, all the models for navigation, apart from the flatline pattern, share a common variant: where users, after reaching a specific point later on in the document, return to the beginning of the document for an extended period of time. In total, 77 of the 320 document triage runs returned to the beginning of the document, though 23 of these views were for less than one second.

### 3.2 Document Features and Navigation

We now report the impact of document content and visual features on the users' viewing behaviours. We commence by characterising user behaviour on pages that lack any of the features we are investigating, before turning to each feature in turn. Summarized figures are found in Figure 3.

**Document Length.** Document length does result in – on average – increased viewing times. The general view can be seen in Fig. 2, where documents are ranked by length and with total viewing time presented. However, this picture is complicated by users using navigation tools to bypass particular pages (e.g. “go to” the penultimate page). We therefore evaluated the impact on page number on viewing time and the number of participants who viewed the page, as well as the simplified aggregate view corrected. Taking the entire viewing history of all documents, we discovered that the Pearson's correlation between the page number and the average viewing time of that page was -0.411. The correlation between the page number and the number of participants who viewed that page was -0.785. Hence, as the number of a page increases, fewer participants viewed it, and those who viewed it did so for a shorter time. This characterizes a global pattern of sharply falling attention being paid to succeeding pages.

**Featureless views.** Areas without any noticeable properties (i.e. plain text only) were typically scrolled past at a fast rate. This brief viewing time makes it unlikely that even superficial skimming of the page occurred that could contribute

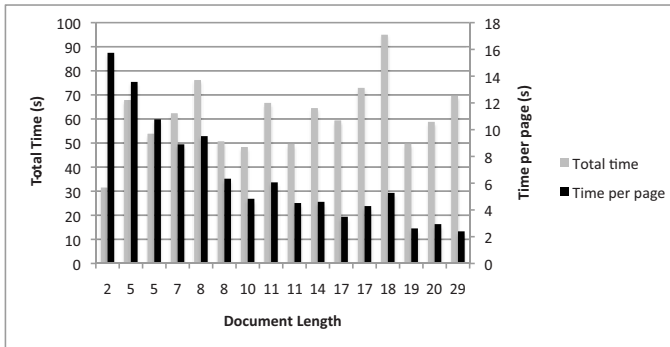


Fig. 2. Timings per document

to their relevance judgment. The average viewing time for a page view that contained plain text only was over one second (1.38s). Excluding one featureless first page with a corresponding long view time, lowers the average to below 0.5s. In contrast, the average time in view of all pages is some 5.28s. Unsurprisingly, applying Student's  $t$ -test at either the view or page level returns  $p < 0.001$  ( $t = 4.96$ ,  $df = 4018$ ). This is strong evidence that both pages and displayed views with minimal distinguishing features will receive only scant attention from the average user.

**Initial Page.** Abstracts and titles are considered to play a key role in user relevance judgments [19]. Both are most often located on the first page with much of the introduction. Across all documents, the first page accounted for 48.2% ( $SD = 15.24$ ,  $n = 320$ ) of all viewing time. This suggests the first-page content plays a central role in the relevance decision making of users. The average time a user spent on the first page was 23.21 seconds, in contrast with an average for all pages of 5.28s. As mentioned in Sec. 3.1, users relied solely on these three parts for making their relevance decision on 43 documents, and a further 30 documents were judged with less than 10% of the user's time spent outside the first page. Sixteen users relied solely on the first page for at least one document, and two users displayed this behaviour on 11 documents each.

Participants returned to the beginning of a document in the triage of 77 documents. In 23 instances the last view of the first page was extremely brief ( $< 1$ s). However, the average time for this second viewing for the remaining 54 cases was some 13.94s ( $SD = 13.45$ ), compared to an initial viewing time of 29.15s ( $SD = 30.41$ ). Second viewings can thus be more than cursory, though the difference between first and final view times is significant ( $p < 0.0001$ ,  $t = 4.46$ ).

Given the high impact of the first page itself, for the content features that follow, we discount the first page in all cases.

**Headings.** A simple correlation of the number of headings per page with the time spent with the page in view gives  $r = 0.286$  ( $df = 4018$ ,  $p < 0.01$ ). Considering the more focussed data of views, this correlation increases slightly to  $r = 0.322$ .

This suggests that headings have a modest impact on navigation behaviour. The time in view of an individual heading was 6.57s (SD=2.69), while for views without a heading the average was 4.20s (SD=2.05). The t-test yields  $p < 0.0001$  ( $t=4.438$ ). Furthermore, there is a correlation between the number of headings on a page and the number of participants who viewed that page for  $\geq 0.3s$  of  $r=0.354$  ( $df=199, p < 0.01$ ). Thus, one can conclude that the presence of a heading increases the time for which user will study the page, and it also increase the likelihood that they will halt at that view rather than scroll continuously.

**Pictures.** In total 46 separate pictures occupied 11.9% of the total surface area of the documents, and were in view for an average of 17.73% of the participants' time. Testing the time in view for a picture (5.86s) results in Pearson's  $r=0.242$  ( $p < 0.05, df=720$ ), indicating a weaker impact than headings. This is also reflected in other statistical comparisons: images fail to change the number of participants who viewed the page for  $\geq 0.3s$  ( $r=0.042$ , not significant)

**Tables.** Tables and other statistical figures occupied 7.6% of the visible space and 9.2% of total triage time. A Pearson's coefficient of a mere 0.024 confirms that there is minimal likelihood of an effect, confirmed by other tests failing to obtain significance. It seems likely that this content plays at best a marginal role in users' information triage work.

**Emphasized Text.** Emphasized text includes bullet points, bold text, italic text and underlined text. When the documents included emphasized text, these areas were visible on average 7.05% of the participants' time (SD=3.56). On average participants took 3.35 seconds to triage each area containing emphasized text. In total participants spent 938 seconds of their triage time on these areas, accounting for 4.8% of their total triage time.

**Conclusions Section.** Of the twenty documents used in this study, 14 contained a "conclusions section". The average time on display was some 7.72 seconds when viewed (SD=2.60s), and applying the t-test against all other pages results in  $p < 0.0001$  ( $t=4.36$ ), confirming an effect, and Pearson's  $r$  is 0.432<sup>2</sup>. No correlation or effect is discernable when testing for either views of  $\geq 0.3s$  (e.g. Pearson's=0.04) or any view. However, it must be considered that conclusions are typically late in the document, and we earlier noted a negative correlation between page numbers, viewing rates and times. Put simply, many participants fail to navigate as far as the conclusion on any given document.

Testing the conclusion page against the three previous pages, to mitigate this placement effect, strengthens all time outcomes (the average viewing time of the prior 3 pages is a low 3.56s). Applying Student's t-test to determine an effect on the number of participants who viewed the page results in  $p=0.051$  (one-tailed,  $t=1.75, df=14$ ), but the difference is small (average number of viewing participants being 11.05 for conclusions pages, and 9.82 for neighboring ones). Our observations of the study suggested, and the navigation patterns reported

---

<sup>2</sup> N.B. we would consider this test inappropriate on a binary value, but include its result for completeness.

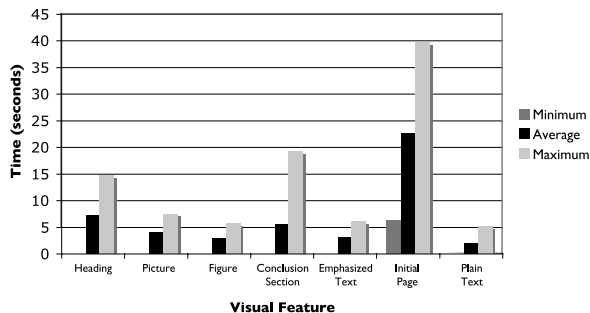


Fig. 3. Mean visible time per visual feature (in seconds)

above, both suggested that some users attempt to guess the location of the conclusions section, and this data gives some further support to that hypothesis.

Participants viewed an average of 8.5 conclusions ( $SD=4.73$ ) from the 14 documents that possessed one: a rate of just over 60%. Not all were scrutinized in detail: just over 30% were seen for less than a second.

**Summary.** The outcome of these observations is a clear set of effects on user navigation behaviour. The initial page holds a key place in the user’s triage reading. Subsequent pages receive less and less attention, with falling viewing times and declining likelihoods of a page either being viewed at all or read in detail. Headings, the conclusion section and some graphical content do result in higher display times and, in the case of headings particularly, an increasing likelihood of the page being displayed for more than a brief overview.

## 4 Discussion

We now recap the main impacts of the findings reported in the body of this paper. In principle, the documents presented to the participants were all rated relevant to the tasks by three subject experts. This should encourage engaged triage behaviour, as opposed to very brief rejections. Our participants reported similar subjective ratings for the importance of document features as reported in the current literature (e.g. [7]). That data is universally self-reported, and we now can compare this against actual user behaviour. In this section, we will supplement our main data with other information from the study, including subjective ratings of document features.

### 4.1 Navigational Behaviour

We identified four distinct types of navigational behaviour that occur during document triage. By far the most common is the *step up* pattern (see Section 3.1). In principle, this navigational behaviour allows the information seeker to



skim all parts of the document, and therefore observe all of a document's features (described in Section 3.2). Variants also allow for this full document skimming. The *flatline* and *begin and end* patterns of navigation are less common than the *step up* approach, but were observed in most participants (see Section 3.1). These do not allow the user to interact with the entire content of the document and may result in the information seeker making poor relevance decisions.

## 4.2 Reported and Actual Behaviour

**Plain text** received a subjective rating of 4.8 out of 10 in the pre-study questionnaire. Plain text was, of course, visible for almost every moment of viewing time. However, when the document display contained only plain text, there was minimal time spent viewing that part of the document (Sec. 3.2). As we shall see, while content text is important, the evidence from other visual cues leads the reader's attention to a particular point. User 15 explained in the post-study interview that "If there is no text that is bold or stands out then I just scroll without reading". This seems consistent with the observed pattern of pages with no visual features being viewed for much less time than other content.

**Document Length.** Liu [14] has suggested that electronic reading is used less for in-depth reading. The data from our study suggests that document length has a minimal effect on users' relevance ratings. However, there is a clear effect where each succeeding page in a document receives less attention. This suggests that if a search engine determines that the relevant content for a search falls in later pages, and is absent from the first page, users may well fail to identify that material in their visual scan of the document. Byrd [6] and Harper [9] have both endeavored to make later content more visible in the user's interaction with documents. However, this earlier work failed to arrive at proven effective solutions. This topic of research appears, therefore, to be worth further study.

**Main Title, Abstract and Introduction.** In the pre-study questionnaire, participants viewed the main title as one of the two most important features of the document, rating it with a subjective importance of 8 out of 10. The Abstract and the Introduction ranked 3<sup>rd</sup> and 4<sup>th</sup> in importance. As all this material typically falls on the first page of a document, one would expect it to receive a considerable amount of scrutiny – and hence viewing time.

This is confirmed by the behaviour of the participants, who on average spent 48% of their triage time on the initial page of a document. Two documents had no formal abstract or introduction, and for these documents the average share of time drops to 43%. We therefore are able to add concrete data for viewing times of visual data to the findings of Cool, and others [4,12].

Subjective feedback reported that obscure and misleading titles are a negative influence in ranking, and time consuming as the reader has to work harder to establish the relevance of a document. Similarly, some abstracts were criticized for being imprecise. Participant 6 commented that he would "prefer a short abstract which doesn't over-analyze the paper". Participant 13 reinforced these points, saying "I love a short simple abstract which gets to the point, and a good

main title”. Due to the format of the study, we cannot discriminate between the impact of these different factors in detail.

**Headings.** have long been held as a significant factor in document triage[7]. The participants rated headings as the second highest visual factor when making a relevance decision (see Sec. 3). Participants repeatedly observed during interview that “headings are really important”. However, we doubt that headings alone are the issue: rather, the manner in which they attract the reader’s attention to body content is key. The impact of a lack of headings is often affective, comments made during the sessions included “I don’t like this, there are no headings or anything” and “what is this? It doesn’t even have any headings in it”.

**Pictures and Tables.** Pictures attracted participant attention and had a mild effect on subjective relevance ratings. Participant 1 explained that “pictures give you a better idea and makes a paper more appealing to someone”. As for abstracts and other content, form appears to have had an effect on the final impact of illustrations. Obscure diagrams attracted considerable criticism. Participant 8 noted that reading some of the figures was simply “too much effort”. Figures had virtually no impact on user behaviour.

**Conclusions Section.** Conclusions received significantly more viewing time than nearby material. Only 60% of the conclusions were reached and viewing times were very brief in a third of cases. It is arguable that conclusion sections have not had, as a whole, much impact on subjective relevance ratings. The observed behaviour is consistent with reading of conclusions being informative and summative, rather than a key point in a user’s triage decision.

**Summary.** Liu [14] characterised digital reading as hasty and incomplete. It seems that this characterisation may be found in an even more extreme form during document triage. Much of each document was simply scrolled by with minimal observation, whilst considerable viewing time was given to the first page alone. Viewing behaviours elsewhere seem influenced by highly visual content. The previous portraying of information seeking as a scent-following or berrypicking behaviour seem relevant, but in a much higher paced form. Content that does not appear on the first page is unlikely to gain attention unless near to “honeytrap” features such as images, or a heading that reflects the user’s information goal. Our data is not yet sufficiently conclusive to be certain of this simple cartoon of user behaviour. Further work is needed to finally prove or disprove this hypothesis.

## 5 Conclusion and Future Work

We are developing a prototype document reader that provides improved user support for triage reading. This software will provide improved within-document query support using IR methods [9] and approximate matching. Most importantly, we aim to improve the attention given to content with low visibility but a good match to the user’s query. There are a number of methods for achieving such

focus support; e.g. keyword highlighting, interactive feedback (e.g. non-linear scrolling) and simple visualization techniques (as already explored in [9,10,15]). We aim to compare these systematically.

User studies to gain further insight on the information triage process are also needed. Eye-tracking provides an opportunity to study visual focus in more detail. However, the centre of gaze has an imperfect correspondence with the user's centre of attention [18].

One major concern, for both user studies and system design, is that the problems of overlooked relevant material are even greater in large documents, such as PDF books of hundreds of pages. Users thus require much better support than we observe at present. It may be that user behaviour changes on longer texts to compensate for the rising cognitive demands of both visual search and precise scrolling. A related issue is that triage behaviours will certainly vary between the media concerned and the information needs and skills of the users. Further user studies with different communities (e.g. humanities researchers, lawyers) and media types are needed to provide sufficient data to provide reliable generic knowledge. Finally, as we refine and understand the process of document triage we are closer to formulating a conceptual model to represent patterns and suggest "relationships that might be fruitful to explore or test" [21].

Through a controlled laboratory-based observational study, we collected empirical data of the navigation patterns of users when viewing documents for triage. Combining self-reported significances of visual document features with the detailed navigational logs from our instrumented PDF reader, we arrived at a detailed picture of document triage as a visual search task. We uncovered a number of differences between subjective ratings and actual behaviour, and were able to characterise a set of navigational patterns that frequently occurred during triage. The study data confirms the importance of known visual cues, whilst illuminating some distinctions between them. We are also able to provide solid data to confirm the intuition of previous researchers that in digital documents little attempt is made to engage in detailed reading. This has significant repercussions for future research: e.g. user cognition may require better support to target skimming activity, and documents that lack the visual infrastructure (e.g. headings) to assist user navigation will receive low subjective relevance scores. We look forward to a more detailed assessment using eye-tracking technology to observe finer detail within the patterns we report here.

## Acknowledgments

This research is supported by EPSRC Grant GR/S84798.

## References

1. Badi, R., Bae, S., Moore, J.M., Meintanis, K., Zacchi, A., Hsieh, H., Shipman, F., Marshall, C.C.: Recognizing user interest and document value from reading and organizing activities in document triage. In: *IUI 2006: Procs. 11th Int. Conf. on Intelligent user interfaces*, pp. 218–225. ACM, New York (2006)

2. Bae, S., Badi, R., Meintanis, K., Moore, J.M., Zachhi, A., Hsieh, H., Marshall, C.C., Shipman, F.M., Costabile, M.F., Paterno, F.: Effects of display configurations on document triage. In: Costabile, M.F., Paternó, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 130–143. Springer, Heidelberg (2005)
3. Buchanan, G.: Rapid document navigation for information triage support. In: Procs. ACM/IEEE Joint conference on Digital libraries, p. 503. ACM Press, New York (2007)
4. Buchanan, G., Loizides, F.: Investigating document triage on paper and electronic media. In: Proceedings of the European Conference on Research and Advanced Technology for Digital Libraries, vol. (35), pp. 416–427 (2007)
5. Buchanan, G., Owen, T.: Improving skim reading for document triage. In: Procs. Int. Symp. on Information Interaction in Context, pp. 83–88. ACM, New York (2008)
6. Byrd, D.: A scrollbar-based visualization for document navigation. In: Procs. ACM Conference on Digital Libraries, pp. 122–129. ACM, New York (1999)
7. Cool, C., Belkin, N.J., Kantor, P.B.: Characteristics of texts affecting relevance judgments. In: 14th National Online Meeting, pp. 77–84 (1993)
8. Ellis, D.: A behavioral approach to information retrieval system design. *Journal of Documentation* 45(3), 171–212 (1989)
9. Harper, D.J., Koychev, I., Sun, Y., Pirie, I.: Within-document retrieval: A user-centred evaluation of relevance profiling. *Information Retrieval* 7(3-4), 265–290 (2004)
10. Harper, D.J., Koychev, I., Sun, Y.: Query-based document skimming: A user-centred evaluation of relevance profiling. In: European Conference on Information Retrieval, pp. 377–392 (2003)
11. Jones, S., Jones, M., Deo, S.: Using keyphrases as search result surrogates on small screen devices. *Personal Ubiquitous Comput.* 8(1), 55–68 (2004)
12. Joseph, J.W.: Relevance judgments and the incremental presentation of document representations. *Information Processing and Management* 27(6), 629–646 (2005)
13. Kuhlthau, C.C.: Seeking Meaning. A Process Approach to Library and Information Services. Ablex Publishing, Greenwich (1996)
14. Liu, Z.: Reading behavior in the digital environment. *Journal of Documentation* 61(6), 700–712 (2005)
15. Loizides, F., Buchanan, G.R.: The myth of find: user behaviour and attitudes towards the basic search feature. In: Procs. ACM/IEEE-CS joint conference on Digital libraries, pp. 48–51. ACM, New York (2008)
16. Marchionini, G.: Information Seeking in Electronic Environments. Cambridge University Press, Cambridge (1995)
17. Marshall, C.C., Frank, I., Shipman, M.: Spatial hypertext and the practice of information triage. In: HYPERTEXT 1997: Proceedings of the eighth ACM conference on Hypertext, pp. 124–133. ACM Press, New York (1997)
18. Mosconi, M., Porta, M., Ravarelli, A.: On-line newspapers and multimedia content: an eye tracking study. In: SIGDOC 2008: Procs. 26th ACM international conference on Design of communication, pp. 55–64. ACM Press, New York (2008)
19. Saracevic, T.: Comparative effect of titles, abstracts and full texts on relevance judgments. In: Procs. American Society for Information Science, pp. 293–299 (1969)
20. Spink, A., Jansen, B.J., Wolfram, D., Saracevic, T.: From e-sex to e-commerce: Web search changes. *Computer* 35(3), 107–109 (2002)
21. Wilson, T.D.: Models in information behaviour research. *Journal of Documentation* 55(3), 249–270 (1999)