Loizides, Fernando (2012). Understanding and conceptualising the document triage process through information seekers' visual and navigational attention. (Unpublished Doctoral thesis, City University London)

CITY UNIVERSITY LONDON
EST 1894

Original citation: Loizides, Fernando (2012). Understanding and conceptualising the document triage process through information seekers' visual and navigational attention. (Unpublished Doctoral thesis, City University London)

Permanent City Research Online URL: http://openaccess.city.ac.uk/1304/

Understanding and Conceptualising the Document Triage Process through Information Seekers' Visual and Navigational Attention

Fernando Loizides

This thesis is submitted for the degree Doctor

of Philosophy in Human-Computer Interaction

CITY UNIVERSITY
LONDON

Centre for Human-Computer Interaction Design

School of Informatics

May 2012

Contents

List of Figures

7

List of Tables

Acknowledgments

Declaration

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to me. This permission only covers single copies made for study purposes, subject to normal conditions of acknowledgment.

Abstract

Information, is a valuable commodity and its effective use is a vital part of everyday life. With the advancements of the internet and the increasing accessibility to it, the location of information is no longer the primary concern of information seekers. Digitisation techniques have made a wide variety of documents available on-line, and more and more publications are being published in electronic form simultaneously to their physical counterpart. The largest challenge currently facing information seekers is that of locating the correct information from the abundance available to them. Whenever a search query is made, the user is inundated with multiple options of documents to choose from. These documents are all deemed to have some relevance to the query produced by using an information retrieval algorithm. Thus far, automatic support has only been provided until the document retrieval level. The user is then left to search through the result set, mostly unaided, by the system he is using.

In order to facilitate support for the users, a solid understanding of the information seeker's behaviours during this triage process is vital. Thus far, research into the behaviour of information seekers during the specific triage behaviour is limited. Even more limited however, is the evidence reporting the visual attention of the users. Since the triage process is highly visual, this important element need to be thoroughly evidence before accurately conceptualising the entire process.

For this reason, this thesis aims to investigate the visual attention of information seekers during the document triage process. This will inform the modelling and conceptualisation of information seekers' behaviour during triage. In turn, this can be used to inform the design of supportive software. The thesis contains a review of related research and identifies the gaps that needs further investigation. From these, a series of user studies are then conducted on document triage. These in turn, facilitate the formulation and discussion of 2 document triage models and measurements to record the effectiveness of document triage.

We study the visual attention of information seekers in four lab based studies, eliciting their exact gaze and focus details. We expand current research in the information seeking domain by reporting on findings from users' triage activities on small screen devices and when under time constraints. Furthermore, a high level diary study, gives us richer data on participants' triage activities over a larger period of time in their natural surroundings. All the studies are brought together to elicit requirements and measurements to understand system and user efficiency during each stage of the triage process.

Chapter 1

Introduction

1.1 Background

Locating information is a constant need in our lives. Whether the information needed is an appropriate place to eat, train times for the day, or which book to read next, information needs appear constantly. These information needs are often accompanied by a further problem; having to choose between likely successful options.

It is not surprising then, that we are increasingly forced to tackle this problem whilst recognising that we are inundated with options and information; often more than we can process. This phenomenon has been dubbed and referred to in many of the literature, such as Bawden and Robinson [18], as "information overload". Bawden and Robinson, also report that information is sometimes so diverse in nature and its quantity so large that in some cases it can lead to pathological conditions. Too much information can also cause us to be less effective at locating relevant information. In fact, "researchers across various disciplines have found that the performance (i.e., the quality of decisions or reasoning in general) of an individual correlates positively with the amount of information he or she receives up to a certain point". If further information is provided beyond this point, the performance of the individual will rapidly decline [37, 53].

The generic term for looking for information is labelled **Information Seeking**. The process entails everything from the realisation that a gap in our knowledge exists, to the satisfying or abandoning of that need [153]. This process has been researched thoroughly and models representing

the conceptualisation of the process are in abundance [76, 100].

This thesis focuses on a sub-process of information seeking. The process is called **Document Triage**. Simply speaking, document triage is the process we go through to make relevance decisions on documents. It occurs after a document set is presented to us, for example after a query is executed on a search engine, and ends when the documents are either all rejected, when relevant information is found, or when the user abandons the triage. Anything from our physical actions to our cognitive workings are used to provide our own individual triage activities. Document triage is specifically defined later in the thesis by related work (See Chapter 2), as well as a current definition based on the current work (See Chapter 8).

Document triage takes place in physical as well as electronic environments. An information seeker can search for academic documents using a search engine, to be returned with suitable results in a list. Similarly, an information seeker can request for academic documents by using a physical library. The librarian can then present the individual with an appropriate shelf or section of the library, which might prove useful to the information need. Both of these scenarios have one thing in common; namely, that the information seeker is then liable to triage through the information in order to extract what he or she thinks is relevant.

Like any complex task, there are a number of different approaches users take to triage through information. A number of these will be discussed throughout the thesis. This has given rise to quite a volume of research, but although some of this search in information seeking has touched on document triage, the amount of research that focuses on document triage is much less. Therefore, a timely part of this thesis is to help develop this underdeveloped area. We can draw on the existing information seeking literature which gives a foundation for hypothesis and expectations of how users may behave or what their task is. The next section will identify the aims of the research which will inform the content and structure of the thesis.

1.2   Aims

One main objective of this research is to provide empirical results of studies directly investigating the document triage process.

Document triage is a complex task with many variables surrounding it. One such variable

is the user type. Is the user a minor looking for information on-line on his favourite band, or is it a retired older individual, searching for holiday destinations? Another such variable is the type of documents being triaged. The information seeker can be looking through web pages, or a folder on the computer with PDF's. The behaviour of all the above examples may vary. It would be impossible for all the variables to be covered by the research in one thesis. In the research undertaken, we simply present findings which provide a norm, allowing other researchers to think about the variables, constraints and behaviours illustrated and to compare and contrast these to their own findings.

In this thesis, a suitable starting point that is familiar to the expertise of the researcher and representative of a large user base was selected. The most suitable starting point which focuses primarily on those criteria would be a visual analysis of document triage. There is a substantial body of literature on visual attention during information seeking which can act as a foundation to the research (including the author's own previous work). Furthermore, visual attention is the centre of gravity of a visual task such as document triage (See Section 1.2.1).

In terms of document genre and user, academic documents are chosen. These, being semi-structured in nature, make them a suitable starting point for researching how a document's features are likely to affect the visual attention of an information seeker. These types of documents also correspond well with the existing array of literature. Interfaces are also a major factor in today's diverse technologies. The interaction of a user on a mobile device while performing document triage is therefore also a vital variable that is required to be tested.

Besides a review of the related literature, this thesis has two aims:

1. **To study and report on the visual attention of information seekers performing document triage on semi-structured documents using electronic means.**

2. **To produce conceptual models and theoretical frameworks from empirical observational and qualitative data.**

1.2.1   Researching Visual Attention of Information Seekers During Electronic Document Triage.

Document triage is a highly visual process. There is normally no audio content to support the triage procedure. Therefore, information seekers will scan up and down the document given in search of clues to inform their sense making process. An information seeker is then cognitively influenced by what they see and act based on their perceptions of the content or the structure of the document [28]. To date, there have not been any reports on the exact visual attention of users during document triage. Earlier research has reported on some of document features which may influence an information seeker's relevance decisions. The self-reported data focuses mainly on predominant features such as the main title and abstract [129], often missing the richness which the remaining features can offer to the seeker. The remaining features, such as emphasised text, figures and conclusions sections are reported on as "full document body" and are not individually scrutinised. We argue that since a document is reported to be read beyond the initial features during document triage, then each feature could play a decisive role in the decision making process and should be individually reported on. We therefore begin by identifying user's visual interactions with the document features in order to identify them. We thus set a foundation for further research into the actual influence of each feature in the decision making process.

The visual path through which the users move from one element to the other and the order in which users scan things visually has not been reported on. Given the variation in structure that documents may have there are a number of unanswered questions we can ask ourselves. For instance, do information seekers look at all the headings or just some? Do they spend an equal amount of time on all the length of the document? The main methodology and study design used for recording document triage to date, have been self-reported evidence [158]. One of the reasons for this is mostly technological availability and limitations in recording the actual visual attention of users. As eye-trackers are becoming more widespread and accurate, we are able to produce direct observational measures to add more detail to our current knowledge.

1.2.2 Producing Conceptual Models and Theoretical Frameworks

By analysing the empirical data we extract from our studies, we identify patterns of navigation and visual attention. We want to represent these common patterns in a conceptual form to allow for them to be summarised and understood more widely. Unlike similar models to date, we require our models, to be evidenced on statistic of actual observations human visual behaviour rather than self-reported evidence which can at time be inaccurate or lack detail. These, models may be of interest to groups such as researchers, psychologists, developers and designers as they help identify human intentions when seeking for information. Although we acknowledge that our models do not represent the document triage process in its entirety, we are able to use them as a representative of our research within the constraints that we have implemented and this should be acknowledged by ones using them. Further work, either alongside our data, or building from it, is also encouraged in order to refine and develop the models further.

In order to be able to provide models, we need to acquire the data upon which they will be based. There are several methodologies that were considered to be implemented. The first decision made was to actively involve users. Our goal, was to record visual attention and this could only be achieved by observing participants performing triage. Furthermore, the system used itself was not under scrutiny and the data was previously extracted using different systems with corresponding results [28]. We also employed a fully running system, which represented realistically how users would triage data electronically. Since we are attempting to capture the users' actual current interactions with the document, there is no need to simulate or prototype a variant to already existing systems that are in use. We should note however, that we did modify the system to record user navigation (See Chapter 3), but this function was invisible to our participants. To complete our methodological approach, we decided on a realistic context of use to be implemented. Our participants, would be given a realistic task to work on, which was domain specific, thus reflecting their behaviour as realistically as possible. Although this slightly limits the generalisability of our findings in that it may not apply to other types of search, it allows for he investigators to be in control of the experiment and create likely scenarios which the participants would undertake. Furthermore, the experiment is more robust, since our expertise in the domain gives us the knowledge to better understand the documents used and the material we are working with.

The research methodology implemented in this thesis is mostly empirical and lab-based. This was the most suitable approach to take due to the nature of the data we wanted to extract. This methodological approach primarily gains quantitative information about users performance but supplements with qualitative data to get a better insight into what users' behaviours actually are. Our models, are primarily based on probability statistics of users's actions. We chose this method over self-reported data for two reasons. The first is that the data may be biased and false, as the users may not describe accurately what their behaviours are. The second reason, is that we needed to provide a further richness to the data, which is not seen in previous studies. Reader and Payne [124] argue the validity of controlled lab-based experiments within a methodological viewpoint within the HCI field. They also argue that, beyond comparative studies, controlled lab-based experiments "fruitfully can be used in HCI to test hypotheses that derive from theories, where the theories have design relevance, even if the hypothesis or findings do not". This is important, since our chosen methodology for the majority of the experiments is lab-based and controlled. As they report, "experiments can have applied relevance even if the findings themselves are not immediately applicable".

The detailed reporting of the visual attention of users would enable us to identify more accurately where users are actually focusing their attention and if there are any elements we did not identify previously as visually attractive. This way, conceptualising a human behavioural process assists in the understanding of document triage. Upon having quantitative data as a foundation, models of the human process that is being observed can be formed. In turn, these generated models can help us inform the design and provide systems which are closer to what information seekers require.

Models for the entirety of the information seeking process already exist (See Chapter 2). Gary Marchionini describes one such behavioural model for information seeking as a whole [100]. Document triage, which is the focus of this research, lies within the last three stages of Marchionini's information seeking model; namely Extract Examine Results (from the corpus of documents), Extract Information (for relevance decisions), and Reflect and Stop (alter the query or stop if enough information has been identified). Other models, such as Ellis' [50] and Kuhlthau's [91] models also provide us with information on the general information seeking process. These models provide

us with a further view on information seekers' feelings and emotions during the seeking process. Kuhlthau reports on a low confidence level which increases as the information seeker has clearer goals and understanding of the search being undertaken. We can therefore extract that we should see a progressive confidence level in users as the triage process progresses.

There is however, a need to focus on the individual internal workings of the information seeking process in more detail. Bates [17] has critiqued the common information seeking models, agreeing with Kuhn [86] in claiming that "major models that are as central to a field as this one is, eventually begin to show inadequacies as testing leads to greater and greater understanding of the process being studied". Ellis also observes that "there is a need for more micro evaluation of the activities and environments of the users of information systems in order to develop an understanding of the relation of information services to those activities and environments" [51].

Two Document Triage models are created and presented in this thesis (see Chapter 8). The first, dubbed the **funnel model**, conceptualises the document triage process in its entirety. The second, covers the internal workings of within-document triage and reports on the visual perceptions of information seekers. We note that the models presented are not a representation of the behavioural and procedural stages of every scenario related to document triage. As mentioned earlier, they produce a norm and starting point based on the data we have acquired from the user studies, related work and previous work of the author. It is to the discretion of fellow researchers to identify the similarities and differences between the work presented here and their own experiments. The models help to guide specific scenarios and stimulate thinking. The models' generalisability is limited to the work discussed in the subsequent chapters and should be subjected to scrutiny when further variables are introduced. These models can then used by other researchers as a starting point to investigating document triage further. We also encourage developers to test the models by making use of them in the design of bespoke software. These can then be used to establish guidelines and heuristics for developers relating to information seeking.

1.2.3 Summary

In this thesis we are primarily researching the visual properties of document triage in the context of academic information seekers and the documents that they use. We use observational methods

rather than self-reporting evidence which was used until now. The information is then used to produce conceptual models of user behaviour during document triage. In order to extract the visual attention that is required by us, we choose to employ mostly lab-based controlled studies which fit our needs better than the alternative methodologies. In order to increase ecological validity[1] we focus on replicating, as close as possible, real world scenarios and software to study the participants' actions as non-intrusively as possible. We also report on a diary study to investigate the strengths and weaknesses of the studies, by providing richer data. Due to privacy and practicality issues, we were unable to conduct observational experimentation on the information seekers' own work environment.

1.3 Thesis Overview

Overall, this thesis reports 5 main user studies. From these, two conceptual models of document triage are produced. In this section a description of the main focus and contribution of each of the chapters within the thesis is provided, as well as how those chapters connect together to form a coherent argument.

1.3.1 Chapter Two

Literature Review. Following the introduction, the main related work is presented. Direct and indirect evidence which links to the document triage process is reported on. The aim is to inform the reader of the foundations by which inferences are made and the current stage of research in the area. Although this chapter provides the bulk of the analysis of related literature, each chapter expands on more detailed work relating to the specific subject discussed.

1.3.2 Chapter Three

Empirically Exploring Document Triage Behaviour. A lab-based user study is presented that investigates the patterns of information seekers' scrolling while performing triage. Participants were asked to perform triage on a set of documents given an information need and empirical data was collected. A bespoke software tool, aimed for supporting researchers was created, and is

---

[1]For a further discussion on ecological validity see Chapter 9.

presented, which recorded the contents of the screen in detail. The navigational patterns of the participants as well as the contents of their screen and time spent on each document section are also reported on.

1.3.3 Chapter Four

Eye-Tracking Document Triage. The navigational patterns identified in the lab-based in Chapter 3, made it evident that the visual focus of the participants required a more precise and detailed examination. For this reason, a lab-based eye-tracking study was conducted. This second lab-based study builds on the results of Chapter 3 to report on the exact visual attention and fixations of information seekers, while performing triage within-documents. The findings also reveal further granularity in user behaviour and we uncover parts of the document that are scrutinised during triage that previous studies did not report on.

1.3.4 Chapter Five

Small Screen Triage. Our previous studies, produced a foundation and clear understanding of the visual attention of information seekers during document triage. We therefore, proceeded to change external variables in order to get a better understanding of variations to the initial findings. One of the variables we needed to investigate, was the impact of screen real estate on triage visual attention. Thus far, we have focused on one screen size. In order to increase the ecological validity of the results, we investigate the impact of a small screen device to document triage. Using specialised portable hand-held testing equipment, a lab-based study was conducted, reporting on the navigations as well as screen contents of information seekers using a small screen real estate. Several behaviours are replicated from our previous studies. We are able to further report behaviour specific to information seekers performing triage on small screens.

1.3.5 Chapter Six

Applying Time Constraints to Document Triage. Document triage takes place outside of a controlled lab-based environment. Time-constraints are a commonly accepted limitation to a user's information seeking process. We also identify during our studies, and from related work,

that document triage is also recognised as a 'fast-paced' process. This chapter examines the effects of time constraints on the document triage process in a lab-based study. This is important because physical time constraints and even emotional frustration can cause users to have a low tolerance of time consuming information seeking tasks (See Chapter 7). The effects of three time constraints on a document triage process are reported on and contrasted with the studies without a time constraint.

1.3.6   Chapter Seven

Diary Study of Multidisciplinary Document Triage. Our research into within-document triage, identified how users will judge documents as to their relevance. Our studies however, have been lab-based and so we were unaware of any variations that may occur from a users natural setting. Furthermore, the disciplines reported on have been limited. This chapter presents a diary study which gives evidence of a diverse user base during their triage activities in their everyday lives. The findings are used to complement and confirm the validity of the data reported on so far, within other disciplines and outside the lab-based environment. We also uncover behaviours worth researching further and provide a richer data set for the understanding of the document triage process. One esxample of this are the feelings and confidence of users during the different stages of triage.

1.3.7   Chapter Eight

Modelling the Document Triage Process. Two models of Document Triage user Behaviour are presented. These models are built on information from the previous chapters (3, 4, 5, 6, 7). The first, looks at the document triage process in its entirety. The second model, focuses on conceptualising the behaviour of information seekers during within-document triage. Furthermore, the document triage process is defined lexically and mathematically. Heuristics are presented for a) an information seeker b) a software designer and c) document authors. These encompass our findings and further aim to inform the design of specialised document triage software; an example of which is presented in the chapter.

1.3.8 Chapter Nine

Discussion. Following our 5 user studies and our model presentation, we highlight and discuss the main findings and their contribution to the existing literature. We complement, expand on and sometimes contradict previous work by other researchers.

1.3.9 Chapter Ten

Future Work and Conclusions The work produced in this thesis makes a substantial contribution to the information science, human-computer interaction and digital library community. As we identified however, there are limitations to the research presented. Document triage is far more complex than we anticipated. For this reason, we focused on investigating a very specific domain within-document triage in detail, while explaining the process as a whole. In this chapter, we suggest future work to expand our understanding of the document triage process that we were unable to cover in this thesis. Finally, a short summary of the work of the thesis is presented.

Chapter 2

Literature Review

2.1 Introduction

Our main focus is the visual attention of information seekers during the document triage process. In particular, we are considering academic documents, which are semi-structured in nature. In order to familiarise ourselves with document triage and better understand the context within which our participants will be tested, we review research both in document triage, and within other directly related areas. Within the field of document triage it is necessary to identify and understand the facts on which strong scientific evidence points to gaps in the knowledge and require reporting on. Each subsequent chapter, that contains a user study, contains a further, more in depth, review of specialised work.

Document triage is the process by which an information seeker makes a relevance decision on a set of documents. An equivalent automated process is that of information retrieval. In turn, document triage and relevance decisions are part of the information seeking process. A conceptualisation of the relationships between these processes can be seen in Figure 2.1. Information seeking has been studied by many different disciplines. The interaction with electronic documents has also provided the need for human computer interaction and information retrieval based research. Therefore, the research will be based on the domain of human computer interaction. We provide a review of computer software which assists document triage that has been created and tested with users. Although we are interested in recording the document triage process when it is not influenced

by supportive technologies, we need to be aware of the effects these have.
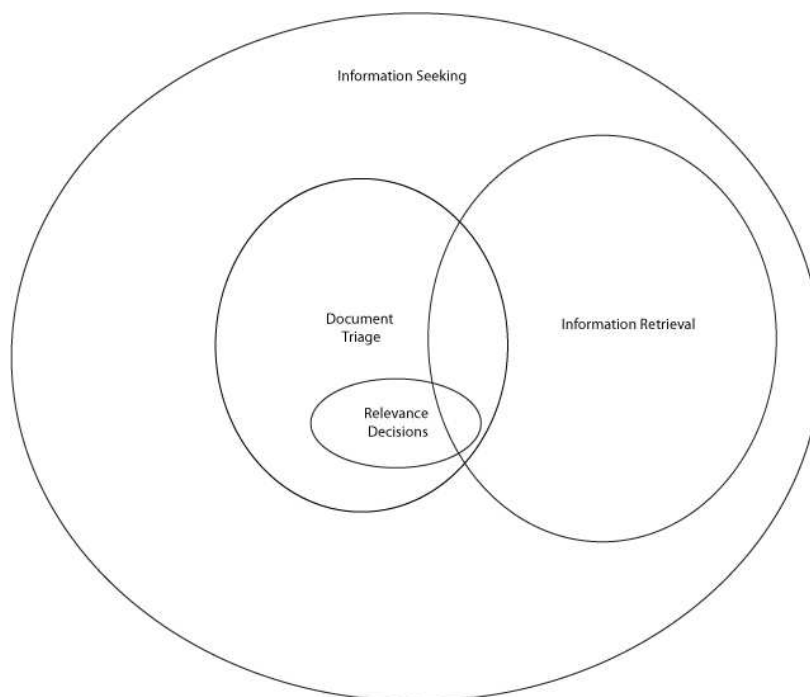
Figure 2.1: Conceptual representation of where the document triage process lies within the information seeking domain

Knowledge that exists and contributes to the field of document triage is identified for a holistic picture of information seeking to be built. In order to guide the reader through the literature we present the related work in a top down approach starting from an overview of information seeking. In summary, the following literature is discussed:

- **Information Seeking.** A review of the information seeking process as a whole is presented in section 2.2 of this chapter, through the different representative models. Our main motivation for researching the document triage process specifically, is the lack of data specialised to this part of information seeking. Although these models provide an overview of the information seeking process, there is little evidence to detail document triage. Bates [17] and Kuhn [86] report on how models, such as those conceptualising information seeking, which cover a process so central to a field and therefore many stages, begin to show inadequacies in terms of the details within the processes. Ellis agrees and stresses the need for more micro evaluation of specific environments of the users in information systems [51]. These works, testify to the

lack of detailed theoretical infrastructure within information seeking, such as the document triage process.

Thus far, information seeking has given overall patterns on how users search for information and interact with technology to achieve their information seeking goals. These models, guide our research by giving us a framework within which document triage occurs. We are able to understand the process that happens before document triage and the effects after document triage has taken place. We in turn focus on the micro analysis of a sub-process of information seeking (document triage) which has been criticised for its lack of accurate representation [86]. In order for us to be able to build on (or indeed contradict) the details of document triage as they are presented by more generic information seeking research, we need a clear understanding of these general foundations. We therefore present a series of models for the information seeking process and focus on the parts that relate to our research.

- **Relevance Decisions.** Within information seeking, users need to identify documents ideally as being 'relevant' or 'not relevant'. Document triage is a form of relevance decision making, directly relating to documents. More specifically, this can be described as the process by which these decisions are made. In this section, we focus on the relevance decision process (See Section 2.3). This takes place towards the end of the information seeking cycle. The related work gives more detailed descriptions of the behaviours that our users are likely to perform and uncovers some likely reasons for this. We are more interested in the actual behaviour rather than the cognitive workings, yet we gather together this data also to get a more accurate representation of information seekers.

- **Satisficing.** Satisficing is the theory that users will find an adequate solution rather than the optimal one during an information search. Since our users are making relevance decisions based on an information need, we study this theory in more depth (See Section 2.4) to set a theoretical foundation for our work. The theoretical framework is also enforced by the knowledge from related literature to information seeking.

- **Document/Information Triage.** The word *triage* comes from the medical term meaning

"sorting and allocating aid on the basis of need for or likely benefit from medical treatment" [1]. This process can be applied to sorting information and specifically, wading through documents to assess their relevance to an information need. Although the specific area of document triage has not been thoroughly researched, we present work that has been published in the area which contributes to our own (See Section 2.5). After reviewing the information seeking process as a whole, and having laid the foundations of the more specialised relevance decision making process, we are more equipped to understand this work.

- **Eye-Tracking.** Document triage is a highly visual task. Thus far, the vast majority of studies have focused on self-reported data. In section 2.6 we present eye-tracking studies that follow similar methodologies to our own and utilise the direst observation methodology, using an eye-tracker for accurate results. We also justify the reason we chose eye-tracking to be used in several of our studies.

- **Cognitive Psychology, Perceptions and HCI Principles.** The highly interactive process of document triage requires a grounding on interaction principles as well as the cognitive understanding of information seekers. Although we do not report on "why users look at the things they do" in detail (instead we focus on the 'what the users look at'), we need to be able to understand the cognitive limitations and natural interactions that users will have when presented with an information seeking task, in order to plan the experiments correctly. By looking into the principles and methodologies of earlier work (See Section 2.7), relating to these principles, we are able to more accurately structure our hypotheses.

- **Software Tools.** Software tools can assist users in locating relevant information from a large data set. There is limited research in the tools which assist information seekers in their triage process. Our aims are to understand and conceptualise the visual attention the users while performing document triage. In turn, the data we produce can better inform the design process for supportive triage tools. Interestingly, very little work has been done on tools specifically created for document triage (See Section 2.8). What is even more interesting is that the experiments undertaken include alterations of document views to participants;

---

[1] Princeton on-line dictionary 2010

for example, by presenting documents in a summarised view. In the review we infer that changing the presentation of documents can affect the user's behaviour and influence the document features being viewed. Researchers can compare the findings in our studies to assess the benefits of these tools and if a design can be altered to address the features that an information seeker requires.

By reviewing the research already conducted surrounding document triage we establish a theoretical foundation and methodological framework to help tie sections together with the research presented. We also identify work previously not reported on which we require in order to further enforce this foundation. We are specifically interested in the process by which relevance decisions are made; namely, document triage.

2.2 Information Seeking

2.2.1 Introduction

Information seeking is defined as "a process in which humans purposely engage in order to change their state of knowledge" [100]. The term information seeking is chosen over information retrieval (IR) because as Marchionini states: "it is more human oriented and open ended". IR findings are still reported on however since that the two processes can often overlap. Since the focus of this thesis surrounds the manual human process of document triage, information seeking is the most suitable starting point to report on. The process of document triage is a browsing technique rather than the extraction of a known item. We therefore look at the information seeking field for a foundation of triage activities.

As users of electronic libraries look for information they implement several strategies to help them find the information they require. These strategies may be analytical (more direct and structured) or browsing (more opportunistic and less structured) in nature [94, 101]. The information seeking process begins when one decides that there is a gap in his or her knowledge. This has been referred to as an anomalous state of knowledge [19]. The information seeking process will then continue to exist until the seeker is either satisfied with the information required, or abandons the process.

Information seeking is a process that has been thoroughly researched in its entirety, and is best explained by several models that have been created and adapted to represent it. These range from behavioural and process, to cognitive models. We know from both information seeking and information retrieval research that "information systems ought to be highly iterative and interactive" [20, 21]. Belkin et al comment also on models being under attack because they do not represent some of the user groups.

We will now review in sequence some of the most widely established information seeking models. A more comprehensive guide to information seeking models can be found by Wilson [160].

2.2.2   Information Seeking Models

2.2.2.1   Marchionini's Model

Figure 2.2 shows us Marchionini's procedural model for information seeking [100]. The process begins when a user understands that there is an information need, or as Belkin explains it "an anomalous state of knowledge" [19]. The information seeking process stops based on "external factors like the setting, or search system, or internal functions like motivation, task domain knowledge, and information seeking ability". Before a user can begin to search for information however, a process of understanding the problem is required, which attempts to define the problem. It is worth mentioning that this stage in the information seeker's process is always active, with a high probability that the user will transition to this stage to reformulate his or her understanding and definition.

The seeker has to then identify an appropriate search system which may provide the required information. An information seeker's experience in the domain determines the accuracy of predicting a suitable repository and search system. Interestingly, Marchionini states that "it is well known that information seekers prefer colleagues or human sources to formal sources and then proximate sources of information and easy-to-use systems." The advancements in search systems have given faster and more reliable technologies. This statement may or may not still hold.

Before the information seeker begins to triage through the returned document set, a query must be formulated in order for the user to communicate to the search system what information is being sought. Upon the formulation of a query, the user will execute the search.

In this thesis we look at the three last components of the model, namely: *Examine Results,*
Extract Info, Reflect and Stop.
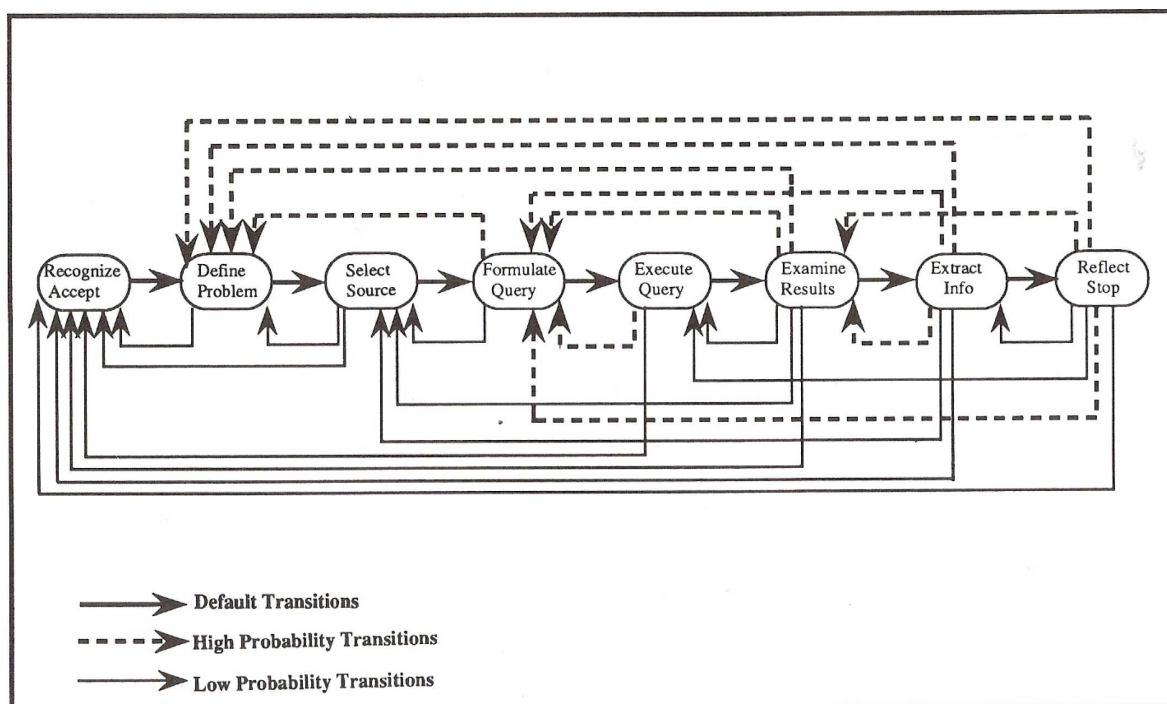
Figure 2.2: Gary Marchionini's Model for Information Seeking.

A user examines results in order to "assess progress towards completing the information seeking
task". Marchionini reports on how "surrogates (e.g. titles, abstracts, thumbnail images) can be
scanned to identify those that suggest a more comprehensive relevance assessment". The examina-
tion process is currently aided by highlighting, organisational views and specialised tools. More on
specialised tools is is discussed in Section 2.8.

An information seeker extracts information from a document that he has deemed to be relevant.
Marchionini reports of how this is done during the information seeking process. He goes on to say
that the document may be re-examined eventually. The triage process encompasses the subsequent
further reading stage (as we present in Chapter 8) and the model presented produces a holistic
procedural conceptualisation of the entirety of the process.

After the seeker has examined the results from each individual query then he may choose to
iterate back to find more information or to stop the process completely. Likely reasons for ending
an information seeking session can be external factors like the setting or internal factors such as

motivation. These last three steps of Marchionini's model encompass the behaviour of the information seeker during document triage. The model presented is a procedural one and so does not include users' cognitive or behavioural contribution. Furthermore, the level at which the model reports on the stages is limited. A more detailed, specific approach is needed to investigate the document triage process specifically. We are given specific guidance into the details that need to be scrutinised further. Marchionini reports on the use of surrogates (e.g. titles, abstracts, thumbnail images) that can be scanned to identify those that suggest a more comprehensive relevance assessment. These are therefore the main features which our research focuses on.

2.2.2.2 Kuhlthau's Model

Kuhlthau takes a more cognitive and emotional model approach the Information Search (Seeking) Process [90]. Kuhlthau reports on user experience during six stages in the process. There are three areas that are explained: Feelings, Thoughts and Actions. The Model can be seen in Figure 2.3. This thesis mainly focuses on user behaviour and visual attention. However, it is acknowledged that these are not the only factors which affect an information seeker. We therefore examine, the *Exploration*, *Formulation* and *Collection Stage*, which are directly related to the document triage process.

Model of the Information Search Process						
	Initiation	Selection	Exploration	Formulation	Collection	Presentation	Assessment
Feelings (Affective)	Uncertainty	Optimism	Confusion Frustration Doubt	Clarity	Sense of direction / Confidence	Satisfaction or Disappointment	Sense of accomplishment
Thoughts (Cognitive)	vague →			focused	increased	interest	Increased self-awareness
Actions (Physical)	seeking	relevant Exploring	information	seeking	pertinent Documenting	information	

Figure 2.3: Carol Collier Kuhlthau's Model for Information Searching.

These steps begin with feelings of uncertainty and confusion from the Exploration stage. These feelings and thoughts "diminish as confidence begins to increase" in the Formulation stage while

"uncertainty subsides as interest and involvement deepens" at the Collection stage. It is important to take these emotions into account while performing user testing and investigate their relevance with data at the triage level, to identify inconsistencies or report on possible reasons for user behaviour.

Document triage is a subset of the information seeking process. By interpreting the findings by Kuhlthau, we predict that information seekers will begin with a low confidence level and progressively become more and more confident during their triage process. The stage of exploration is the most complex step of the triage process. The seekers, however, seem to be tolerable during the early stages of triage at which point the knowledge on the problem and the results should be low. It is not uncommon though, for users to find the initial exploration stage discouraging, especially as the ability to express the need to a search system can be difficult.

As the focus formulation stage is reached, "feelings of uncertainty diminish and the seekers' confidence increases. Interaction with the search system and the information at this stage is very high. It is common for a user to experience difficulty in the remainder of the seeking process, if a focus is not established at this stage.

During our investigation we primarily focus on the visual attention and behaviour of the participants. Kuhlthau's model investigates the feelings and thoughts of users. The detailed analysis of the thoughts and feelings of users during document triage is not thoroughly reported on. We do however report some data to compliment our models (see Chapter 7).

2.2.2.3 Ellis's Model

David Ellis focuses on the behaviours involved during the information seeking process [50, 52]. There is no structure or flow to the original model in that the behaviours reported on are descriptive in nature to a higher level of abstraction to individual actions. Ellis does not use the term stages but uses the term 'features' to represent the behaviours. These features are named and defined here: This model has been iterated and extended during the years. The main patterns of behaviour identified by Ellis are:

1. **Starting**: Identifying relevant sources of interest.

2. **Chaining**: Following and connecting new leads in an initial source.

3. **Browsing**: Scanning content of identified sources for subject affinity.

4. **Differentiating**: Filtering and assessing information sources for their usefulness by their known differences.

5. **Monitoring**: Keeping up to date of developments in a given subject area.

6. **Extracting**: Selectively identifying relevant material.

7. **Verifying**: Checking the accuracy of the Information.

8. **Ending**: "Tying up loose ends" through a final search [160].

We identify the stages which document triage takes place and are therefore more informative to our research. An information seeker during the triage process is likely to **browse** the contents for likely suitable sources of information. This will be done by **differentiating** information sources based on the meta data available and visible to the user. The information seeker then **extracts** what he or she identifies as relevant information. **Verification** of the information is done subsequently by within-document searching or further reading (See Chapter 8).

The behaviour explained and modelled by Ellis reports on several key features of document triage. What the model lacks, is the detail of each of the patterns explained. We expand on these patterns by providing a more in depth analysis of the visual patterns of information seekers during the four stages of Browsing, Differentiating, Extracting and Verification.

2.2.2.4   Bates's Model

Marcia Bates explains how there are different modes of information seeking that users employ during their search [17] (see Figure 2.4). These modes can also be employed during document triage. As Bates discusses, "Directed and Undirected refer, respectively, to whether an individual seeks particular information that can be specified to some degree, or is more or less randomly exposing themselves to information. Active and Passive refer, respectively, to whether the individual does anything actively to acquire information, or is passively available to absorb information, but does not seek it out". Document triage can be characterised as an "Active" process and usually "Directed" dubbing it "Directed Searching". Bates agrees with the collective research, in that "people

use the principle of least effort in their information seeking, even to the point that they will accept information they know to be of lower quality (less reliable), if it is more readily available or easier to use" [123]. This enforces the theoretical framework and background we have regarding satisficing (explained in more detail in section 2.4). We can therefore be more confident in extracting the principles of satisficing to account for the behaviour of our participants.
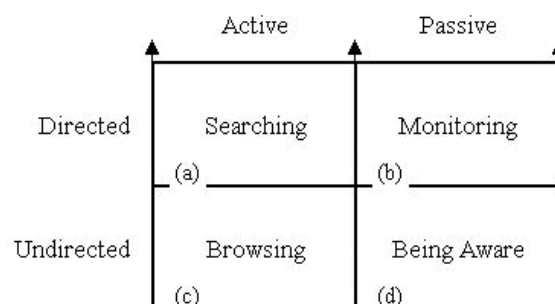
Figure 2.4: Bates' Model for Information Searching.

Bates continues to address the problem of information overload and reports on the advancement of active search systems. These force "people accustomed to mostly passive ways of learning new information not only have to search actively for the information, but also have to master a fair amount of ancillary skills and knowledge just to be able to search for the information, with no guarantee that that effort will actually lead to an answer". Clearly there is a need for usable, document triage software to assist the information seeker.

2.2.2.5   Ingwersen's Framework

Peter Ingwersen approaches the information seeking process from a cognitive perspective [71]. Figure 2.5 shows Ingwersen's cognitive model for information search. It integrates information retrieval system design with information seeking user behaviour to form a hybrid view of the internal cognitive workings of an information seeker. In effect, this model integrates the models discussed above to produce a high level abstract view of the information seeking process. All features of the document are vital in the document triage process; from the objects a user will interact with (e.g. pictures, text) to the interface and social environment. As noted by Saracevic and Wilson however, "The weakness is in that it does not provide for testability and even less for application to evaluation of IR systems"[160, 131].
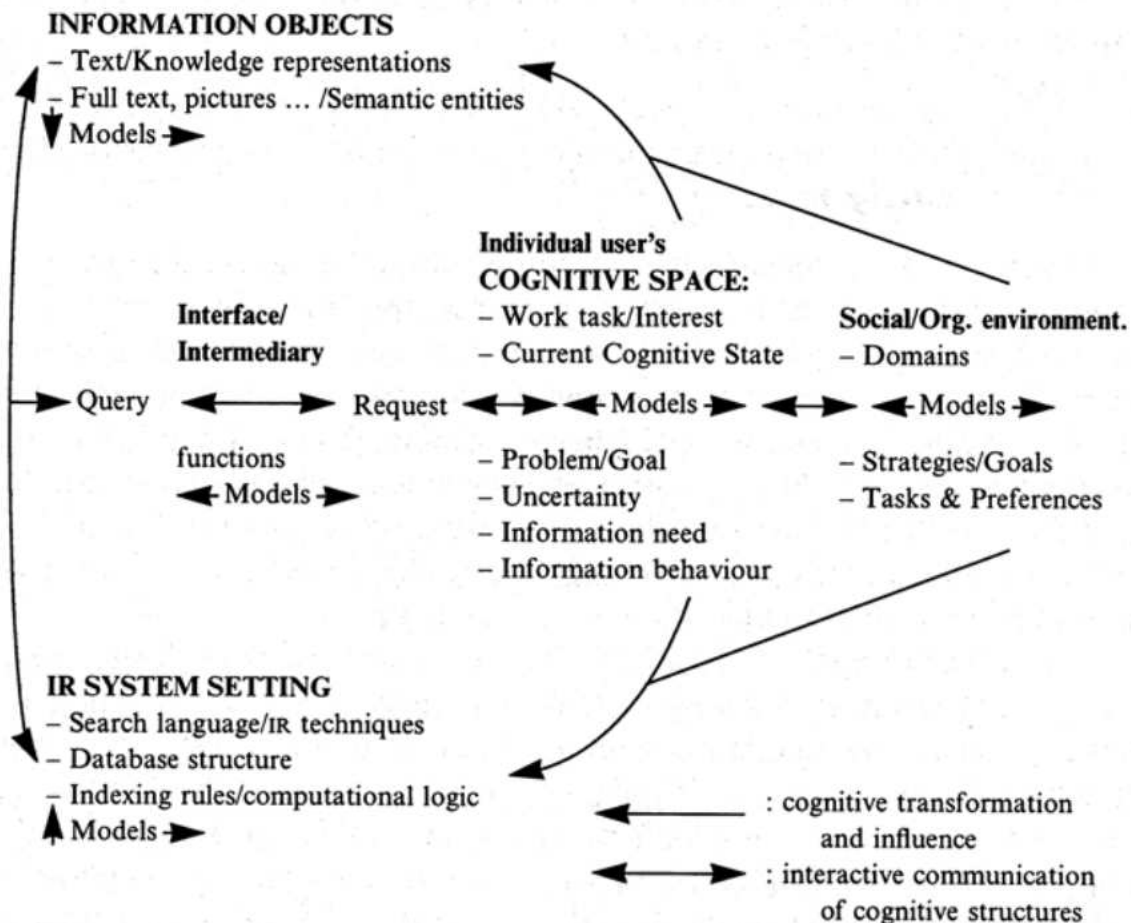
Figure 2.5: Peter Ingwersen's Cognitive Model.

Ingwersen's model provides a framework by which he connects and provides an understanding between an information seekers' cognitive process, actions and physical representations such as information objects. Although this thesis is not focused on the cognitive workings of information seekers performing document triage, it is made evident that the interaction as a whole affects the behaviour of a user. Ingwersen provides an abstract overview of the information seeking process which provides certain limitations. The level of detail which is presented does not cover the internal workings of the triage process. Our focus will be users' visual attention during document triage. This contribution can work alongside Ingwersen's model as practical examples to the theoretical foundation it has set.

2.2.3 Summary

We have so far covered an introduction to the start and internal working of the information seeking process through the study of established models. During each model we identify the individual stages that relate to the document triage process and the gaps that are in need of further research. We also report on which parts of these we cover in our research and which we lay foundations for.

A crucial part of an information seeking process is that of relevance decisions. This was discussed in Ellis's model (Differentiating), Kuhlthau's model (Formulation) and Marchioinini's model (Examine and Extract Results). Document triage is the process by which these decisions are made. We now produce a more detailed description of the literature of this topic.

2.3 Relevance Decisions

2.3.1 Introduction

Document triage is the process by which relevance decisions are made. The research we discuss here identifies behaviours which contribute to a theoretical foundation for our work. We are able to identify where in the literature the gaps are and which methodologies were employed to gather data. We now report on key findings from the relevance decision literature relating to document triage.

2.3.2   What is a Relevance Decision?

The concept of relevance and relevance decisions is a difficult one to define [109, 134]. Saracevic argues how the measure of relevance is "of the effectiveness of a contact between a source and a destination in a communication process" [130]. Saracevic also states that "Relevance does not have to be explained; it is universally understood. It is an intuitive, primitive, yknow notion", in other words, "people understand relevance intuitively" [132]. Saracevic's definitions, although not very technical, summarise the problem faced by systems in interpreting the documents retrieved as relevant or not for the information seeker. This is one of the major reasons why information retrieval systems, can never substitute a human in making *subjectively perfect relevance decisions.*

More definitions and explanations have been produced related to relevance decisions. One important distinction made, was that of separating a system's relevance decisions with that of a user [133]. Schamber et al also give a concise discussion on the meaning of relevance [134]. There is also literature reporting on relevance from the system's perspective. As Park, building on Cooper's work on logical relevance [40], states, "topical relevance is context-free and is based on fixed assumptions about the relationship between a topic of a document and a search question, ignoring an individuals particular context and state of needs" [120]. The focus can be set on the actual query terms which are used as a representation of the need and therefore, the interpretation of the query [68, 147]. It was then Froehlich who suggested that there is no need to have one definition for relevance, but that information scientists "can determine the diverse criteria that people bring to systems by which to judge its output" [55]. In this thesis, we are interested in the human decision making process rather than the information retrieval side of making decisions. The transition from a human information need to a search engine query surrogate is beyond the scope of the thesis. The focus rather, is on the relevance decision making process. This is achieved by investigating the triage of information seekers after the query has been executed and the document set has been presented to the users.

2.3.3   Relevance Through Electronic Means

Relevance decisions (also called relevance ratings and relevance judgements) "are based on information seekers' judgements about a document with respect to an actual problem" [100]. Most

searches are mediated using electronic means. An information seeker, will have in mind an information need, which he or she needs to satisfy. In order to express this need to an electronic system, the information seeker must provide a representation of that need in a way that the system understands. This is usually done in the form of a query term surrogate which is executed by a search engine. Information Retrieval algorithms using weighting (such as Term Frequency x Inverse Document Frequency [82]) attempt to use probabilistic modelling to predict relevance of the query from the user to the documents in a document corpus. Likely results are ranked and presented to the information seeker. In this way an information retrieval process is used to retrieve information which belongs to a fuzzy set [46], in that there are often not well defined boundaries regarding the relevance and information need.

The Text Retrieval conference (TREC) identifies that "people have information needs; they express these needs using some query interface, then evaluate and interpret the results, and are either satisfied with the results or continue the process" [29]. TREC used the Interactive Track which is a task aimed at exploring interactive information retrieval. The "interpretation and evaluation" (document triage) of the resulting documents is a human process which can be fallible. Information seekers may neglect relevant documents and conversely accept documents which upon further scrutiny provide very little information [28]. When performing document triage, relevance decisions are a direct result of the information seeker's visual search of a document. Cool et al identified that seekers believe that 'Format Facets' affect their relevance decisions [39]. In this study, Cool et al used feedback from 300 freshmen participants in order to formulate the data with regards to facets affecting relevance decisions.

Furthermore, task complexity can also influence information seeking behaviour and therefore the relevance decision process [33]. Bystrom and Jarvelin classified the types of information sources as:

1. **Fact-oriented:** registers (manual and computerised catalogues and files) commercial databases.

2. **Problem-oriented:** the people concerned (for example, people proposing, or affected by, administrative actions) official documents (for example, agendas, meeting minutes, letters, applications, memoranda, maps, unpublished planning documents).

3. **General-purpose:** experts (including knowledgeable colleagues), literature (for example, books, reports, journals, newspapers) personal collections (personal notes, calculations, etc.)

They also classified the sources as being either internal or external to the organisation in which the user works. The information seeking process however is not yet complete. The information seeker then has the final judgement on the perceived relevance of the documents to the information need. The accurate definition of relevance has been subject to much debate and interpretation.

2.3.4   Visual Attention in Relevance Decisions

Document triage is a highly visual process. This thesis investigates the information seeker's visual attention in particular, with special focus on academic semi-structured documents. Digital document triage has also been reviewed, often in the context of web searching [150]. We are therefore required to look at work which identifies characteristics of documents, with regards to the relevance decision making process.

Early work from Saracevic pioneered this research into the features which affect the decisions of users [128]. In this work, Saracevic investigates the effects of titles, abstracts and the full text has on participants' ratings of documents. Participants, after searching for documents, are returned first with the title, then the abstract and subsequently the full text of the document. They, at each point, provide a rating for the document based on the information they have. Interestingly, he reports how 78% of the evaluations had the same relevance rating.

Marcus et al, presented a subsequent study in which participants were shown four document representations for each document, in a random order and asked to rate the document as to its perceived usefulness [102]. The participants were also asked to make the same judgement based on the full text of the document. In their findings Marcus et al present the length hypothesis, in which the there is a positive correlation of a document representation with its length. 95% of the participants indicated how titles and abstracts were useful to evaluating the documents. 74% predicted matching subject terms.

Joseph Janes presented document representations to participants incrementally, using a different order each time to different groups, and asking them to assess the document on each representation [75]. The information presented was, the Main Title, Abstract, Bibliographic Citation and Indexing

Terms. The abstract was reported as the most important feature, followed by the main title. The remaining two elements were reported as much less significant.

Cool et al revisited factors on documents which have an effect on users' relevance decisions [39]. The main aim of the research was to "see whether factors other than topical relevance are significant in such decisions". One factor that they found was the format (or Formal characteristics of the document). Examples of these were: Lists, Diagrams, Statistics, Pictures; Class text; Book review; Title; Introduction; Division into topics. This information was extracted from the participants by use of questionnaires and suitable for the task. The visual attention by the participants here is subjective and can therefore neglect important elements of a document which are not exposed. Furthermore, the importance of each item is not specifically reported on.

Carol L. Barry, in a later paper, investigates the reasons "for pursuing or not pursuing documents based on information contained within representations of those documents (i.e., titles, abstracts, indexing terms, etc.)"[16]. Barry investigates three main document features, namely, "document traits", "source traits", and "document representation". *Document traits* are defined as "characteristics of the document that pertained to the physical format or actual publication of the document". *Source traits* are defined as document characteristics that pertained to the intellectual source of the document. If a document trait or source trait was not explicitly mentioned, the response was simply coded for the document representation or part of the *document representation*. Barry found that "only two characteristics, abstracts and titles, were marked by every respondent". Furthermore, "Only three document characteristics co-occurred with all three categories of responses: Abstracts, titles, and the full text of documents".

2.3.5 Summary

This section has shown that relevance eludes ready definition. Relevance is adjusted based on the need of the user or the system which returns it. Several studies have shown specific features which are reported by information seekers as important. Two of these features which are predominant in the studies are the title and abstracts. Most of the studies reported used academic documents which are semi-structured to produce their results. Furthermore, these experiments were mostly done by using representations of documents. Although several studies report on some features of

the documents which are utilised by an information seeker to make relevance decisions, several features seem to be missing. Examples of these are the headings within a document, the images, emphasised text and the conclusions section. These have thus far, been unreported on regarding document triage and relevance decision making.

2.4 Satisficing

Users, when making relevance decisions on documents, unlike automatic information retrieval systems, often do not attempt to make an optimal choice of documents. They do not read every word on every document and, after comparing all the documents, select the most mathematically probable relevant documents. This is evident in their behaviour, which is not always the same between users. This effect of making a decision between satisfying an information need in a way that suffices rather than attempting to find the optimal solution manually, has been dubbed 'satisficing' [141].

Information seekers will employ different searching strategies in order to obtain the required information. Satisficing has been well established within the information seeking field as a representative behaviour for users while searching for information, due to the time constraints and cognitive limitations [153]. It is indeed logical that a "user could not enter a library and carefully evaluate every book, periodical, and multimedia resource before selecting an item" [8]. We should note here that alternative approaches that support satisficing exist. For example, the model of a Generalised Information System [163] is also a theoretical construct which assumes that information seekers calculate the probability of the outcome of each decision to make an optimal decision. We do not follow these theoretical hypotheses and models since they do not accurately represent the process which we are studying. The GIS model, for example, would not be realistic in a situation which has thousands or perhaps millions of possible decision outcomes. For this reason, we follow the satisficing literature as a foundation framework of this research, which is also within the domain of information foraging and scent behavioural theories [2] [121].

Reader and Payne [124] experiment with satisficing by giving a limited time to users in order to observe the effect of trying to obtain information that is 'good enough' rather than optimal. They

---

[2]Foraging and information scent is a theory which states that humans will employ "built-in" searching strategies to seek out useful information just like searching for food for earlier humans.

hypothesise that as long as a document is "good enough to meet the learning needs, it will be read". They argue that satisficing would be more suitable when there are more than one document which is 'good enough' and if the differences among these texts are significant and perceptible. From these two experiments they discovered that participants would use satisficing when the texts were not presented in an outline. It is interesting to also note that there was no evidence to show that the participants showed more efficient learning, in either of the searching strategies which may imply that satisficing is just as efficient as other methods such as sampling [142]. We note a possible gap in the research findings of Reader and Payne that could be expanded on. In their studies, they used texts between 1533 and 1672 words. This does not allow for reporting on behaviour which is necessarily valid in longer (or shorter) text conditions.

For us, one of the most important finding in their experimentation is that subtle design changes to the user interface can have a large impact on user behaviour. For this reason, we choose to use a document reader which is not visually or functionally different to what information seekers would generally use, thus increasing the ecological validity of our findings. This way we are able to identify the visual attention of the user relative to the actual document content without biasing their behaviour. Further research can then be carried out to identify how different interfaces affect visual behaviour. In that case, our data can act as a base model for comparison.

Duggan and Payne continue experiments on satisficing, reporting on skimming through texts with time constraints. They ask the question "are readers able to allocate attention effectively when faced with a longer document than they can read in the time available"[47]? They base their experimental motivation to the fact that "there is little evidence that skim readers are able to devote their limited time to the most valuable parts of a single document". They also present a hypothesis based on the Pirolli and Cards' earlier work on foraging [121] which states that "if readers are likely to begin to read linearly, then they may continue until the rate of information gain drops below the acceptable threshold. At this point they will leave the current patch of text and skip to the beginning of the next patch".

In [47], the authors focus their findings on understanding and memory rather than visual attention, which is our particular interest. Their data reported that memory for reading and efficiency, was not significantly worse in any of the two given conditions. Furthermore, their methodological

approach allowed half of the users to read only half the paragraph contents rather than giving them the entire document. Behaviours however, such as the readers' "tendency to glance across an entire page before settling to read" were reported using an eye-tracker. What is particularly interesting to our research is that the experiment uncovered that skim readers where more likely to allocate their time to more important material within the text, even with time constraints were in place. We are able to further their findings by studying how skimming within the entirety of the document using the same time constraint approach (See Chapter 6).

Duggan and Payne presented further findings [48] from their eye-tracking study in a subsequent paper which highlighted some behavioural elements of the participants. They concluded that "Locating information at the start of the paragraphs and pages increases the likelihood that it will be read. We were able to further expand this data by reporting on which specific text is visually scrutinised by information seekers during triage (See Chapter 4). In the same experiment they also report on how "Information that is later in a text is still relevant to time-pressured readers". In our research, we introduce further time constraint scenarios which add to the behavioural patterns reported here (See Chapter 6). We find that when skimming time is limited even further, visual and scrolling patterns vary even further.

In summary, experimentation on satisficing reports on information seekers' navigational behaviours and patterns which act as guides in our experimental methodology data analysis. Furthermore, these assist us in that they often provide likely explanations for the visual behaviours we observe. We do not explicitly examine the question of why users focus on the features they do. As we recognise gaps in the satisficing experiments, we are able to expand on them and thus indirectly contribute further to the satisficing theory. We are also able to report qualitative and qualitative data that can be used alongside other related research or as foundations for better understanding document triage.

2.5 Document/Information Triage

2.5.1 Introduction

Research related to document triage as early as the 1960's has been reported on; previous to the terms' appearance. In order for us to have a consistent characterization of document triage in our research, we state our own definition. This does not differ greatly from already existing definitions, as we shall see, but gives the reader a clearer understanding of our work.

2.5.2 Lexical Definition and Related Terminology.

In this thesis the definition of the term ***document triage*** is as follows:

```
"Document Triage is the fast process by which information seekers,
go through a set of potentially relevant documents to establish
relevance to their information need."
```

The document triage process has been described from related research in many ways. This literature either looks at the document triage process directly or provides indirect evidence through related studies. One example of this is Wang's cognitive model of document use [158]. Wang and Soergel describe document selection as "the endpoint of a bibliographic search". In this work the process takes a much wider approach than the one suggested in this paper, leaning towards the information seeking model as a whole.

A narrower approach taken is that of Wacholder et al where they describe the "book selection process" as beginning "when an individual has an information need that may be satisfied by a book (whether printed or electronic)" and ends with "the user's decision as to whether the book contains enough information on the topic at hand that it merits more in-depth reading" [157]. This definition looks at the interaction of only one document (specifically a book type) from a within-document perspective.

2.5.3   Existing Definitions

The term 'document triage' is new to the digital library field and has been previously used in a limited number of work[3]. Related terms however have been used to describe the process. We highlight some recurring phrases in the existing definitions of document triage which guide us in formulating our own definition.

The term 'document triage' was first introduced in a broader scope as "information triage" [103]. Here, Marshall et al describe information triage as *"the process of sorting through relevant materials, and organizing them to meet the needs of the task at hand"*. In later work Marshall et al defined 'document triage' as *"the practice of quickly determining the merit and disposition of relevant documents"* [14]. Another definition of document triage, defined the process as one *"determining important, time-sensitive documents from unimportant ones"* [92]. Marshall, in later work, mentions document triage as the process of sorting *"through large sets of documents to identify useful materials..."* [13]. In our earlier work, we had defined document triage as *"the critical point in the information seeking process when the user first decides the relevance of a document to their information need."* [28]. As the process's workings became clearer, the term was refined to include the new findings. The definition for document triage was then defined as: *"the moment in the information seeking process when the user first decides the relevance of a document to their information need"* [97].

The definitions above are analogous and complementary, since each is adapted to suit the individual research being performed. We have identified and highlighted key words that are dominant in these terms in order for us to understand the general meaning of 'document triage', and formulate a suitable universal term. The definitions[4] of the key words summarizing document triage are:

Document: A piece of written, printed, or electronic matter that provides information or evidence or that serves as an official record.

Triage: (Originates from a medical term): The process of determining the most important people or things from amongst a large number that require attention.

---

[3]Here, the term information/document triage is NOT used in the way used by Lankes' model [122]. Lankes uses the term triage in the sense of directing users to expert sources which will later guide the information seeker and recommend likely useful material.

[4]Taken from the Online Oxford Dictionary. http://www.oxforddictionaries.com/

Relevant: See Section 2.3.2.

2.5.4 Document Triage Levels

In Chapter 8 the concept of the three triage levels is introduced. The model presents three stages at which document triage can occur. The first stage, is that of the surrogate level (usually a results list). The second stage is that of within-document triage. A third stage, that is often neglected, is that of further reading triage. In this thesis the main focus is on within document triage.

Surrogate Level After an information seeker executes a query on a search system, he is returned with a set of results. These are usually summarised in the form of surrogates (See Figure 2.6). Surrogates are representations of the actual documents, and can give vital summarised information about the contents of the documents. These surrogates are usually informed by metadata (such as the Dublin Core Standard [5]).

Spink et al, presents research into how information seekers triage results lists [144]. Their findings show how 70% of users will look at 2 pages of results lists or less before moving on to another query, thus dismissing the remaining document set. They also report on how "users continue to have low tollerance for wading through large retrievals". It is hypothesised that due to this low tollerance, users would find shortcuts to information seeking. Nickolas et al, reported that if an abstract webpage was accessed before the document itself could be downloaded, users would likely not download the remaining document [113]. They were not able to prove this thory however, do to subscription and access issues. Buchanan and Loizides however, were able to justify Nickolas et al's claims [28].

Beyond simply scanning through results, information seekers need appropriate interactions to facilitate their triage activities. Marshall and Shipman researched the effects of comparing triage on paper with that of using an electronic support tool (VIKI) in order to organise and extract perceived relevance from documents [103]. They found that users rarely took notes of their thoughts. Furthermore, participants used the table or software space to create semantic categories, in which the documents were disseminated.

Later research showed how categories (or piles) were used as a filing system for organisational

---

[5]http://dublincore.org/ - 2010

Figure 2.6: ACM Results Set After an Executed Query, Displaying a Surrogate View of Potentially Relevant Documents.

document triage activities [28]. We realise how important the task of document triage support is since users "are inundated with so much information that they spend the majority of their time sifting through documents rather than focusing on the task itself" [13]. From this evidence on organising material during document triage, researchers have attempted to create predictive models on user interest.

Badi et al, report on the organising and reading activities of information seekers during document triage, and how these activities can be used to detect user interest. The role organisational activities have on information seekers' triage process as well as relevance decisions is not a new concept. Furnas and Rauch developed "design considerations for the construction of advanced information environments" [56]. These were used to create a prototype system called NaviQue, which allows the user to manipulate a large information work surface, to add and manipulate "objects" within it.

Within Document Level

Within document triage is the point in an information seeker's triage activity when the full text of the document is being scrutinised (See Figure 2.7).

Tefko Saracevic, began exploring the effect of document features, to an information seeker's relevance decisions [129]. He found that decisions from the users do not change as often as one might think, when presented with a title, an abstract and the full text. As Saracevic concludes:

```
''In general, then, a title based index file cannot be expected to perform
any higher than .60-.75 sensitivity  and abstract index files any higher
than .75-.85 sensitivity--because the users do not recognise relevant
answers from comparable formats at any higher rate. These generalizations
assume, of course that the users' judgements of full texts is final and
standard, which in any case seems a reasonable assumption''.
```

Of course, his work was not performed using contemporary technology and electronic readers. Users are increasingly using digital documents and computerbased reading [95]. Reading technologies have changed, and therefore these findings need to be cross-referenced with further studies to convey their validity to the digital domain. Studies between electronic and paper document triage have indicated similarities and differences between the two mediums [28, 103]. Furthermore,
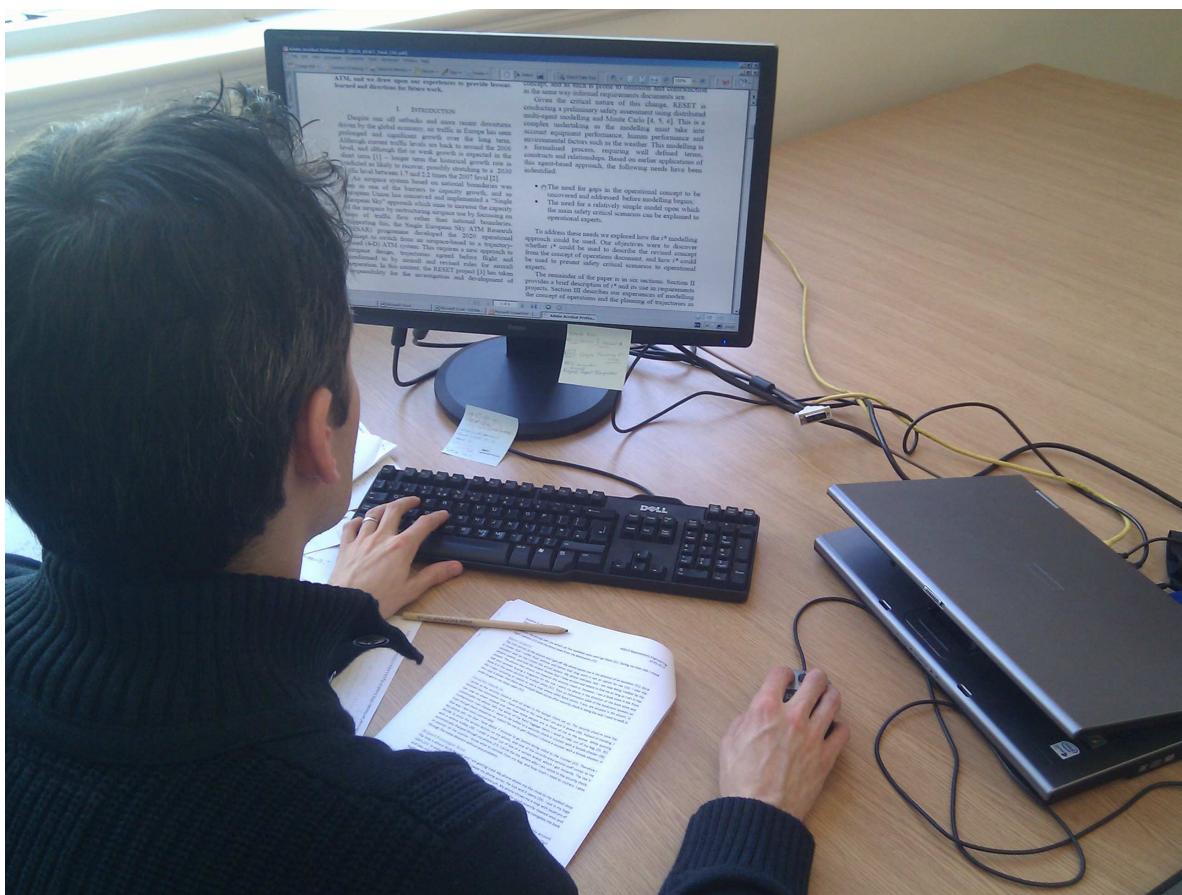
Figure 2.7: An Information Seeker Performing Within Document Triage. The User is Engaged Heavily in an Interactive Process.

whether the user learns the central element, which in our case and Saracevic's is the main title, or the external elements (the remaining features) will be a matter of "featural salience and personal strategy" [106]. We wish to understand how this matter of salience, used predominantly for images and icons, applies to our work in triage.

Beyond these scrutinies, hypotheses have been formulated into what areas and features of a document play the most part in the selection criteria of an information seeker. Wang and Soergel, suggest a model to show the document selection process and criteria (See Figure 2.8) [158]. One of the model elements is DIEs (Document Information Elements) which are used to predict relevance by the information seeker. The criteria they use is limited in terms of within-document features. The factors reported on are Title, Abstract, Journal, Author, Geographic Location, Publication date, Document type, Authors affiliation, Descriptors, Language and Others (publisher, document length, volume and issue, subfile, edition, authors expertise, table of contents, and citation status). The attention is not reported on, neither is any data on the further document features however.
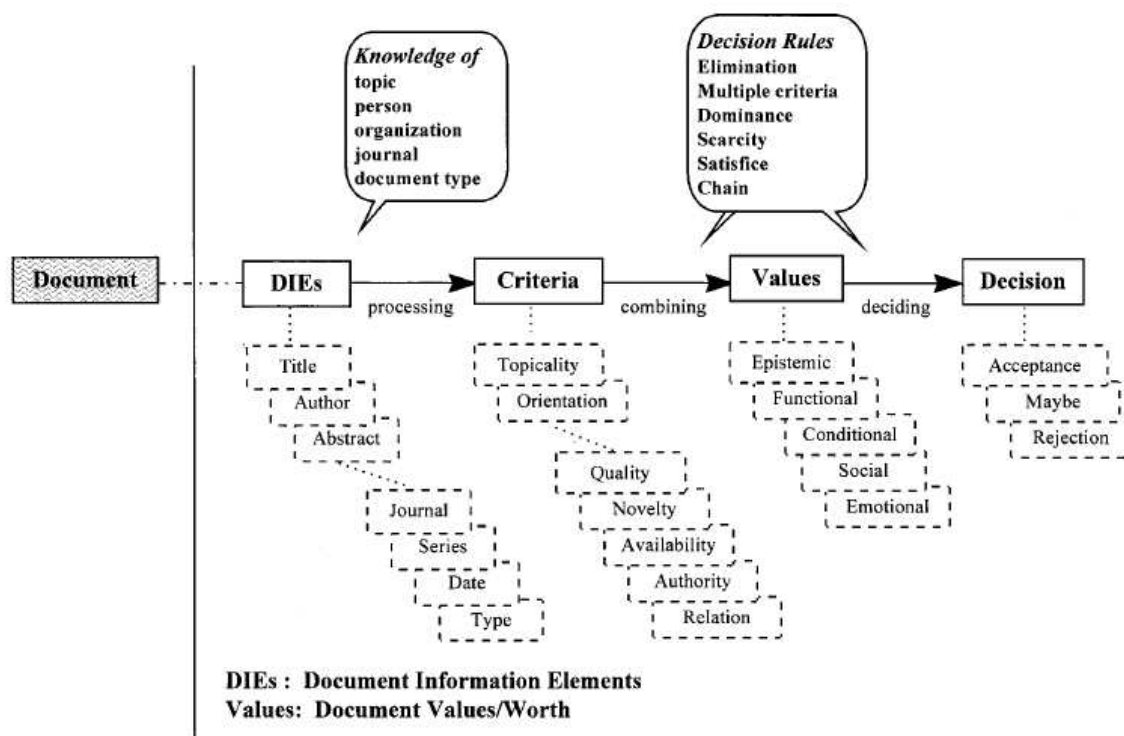
Figure 2.8: Wang's Document Selection Model

Beyond the generic feedback on elements of a document which are considered important for

relevance judgements, experiments have also looked at the role of interaction and scrolling. This research used scrolling to identify user behaviour during information seeking [9, 84]. We do not however focus on the visual features of a document, which again produces a gap in the research area.

Further Reading Level Upon an information seeker making a decision to read a document in more detail, further triage occurs (see Chapter 8). This is the further reading stage and refers to what is known as active reading [7]. Figure 2.9 shows a user reading a document within a book. The reading stage is more relaxed and less interactive than the other two stages (See Chapter 8). Triage however, can still take effect during reading.
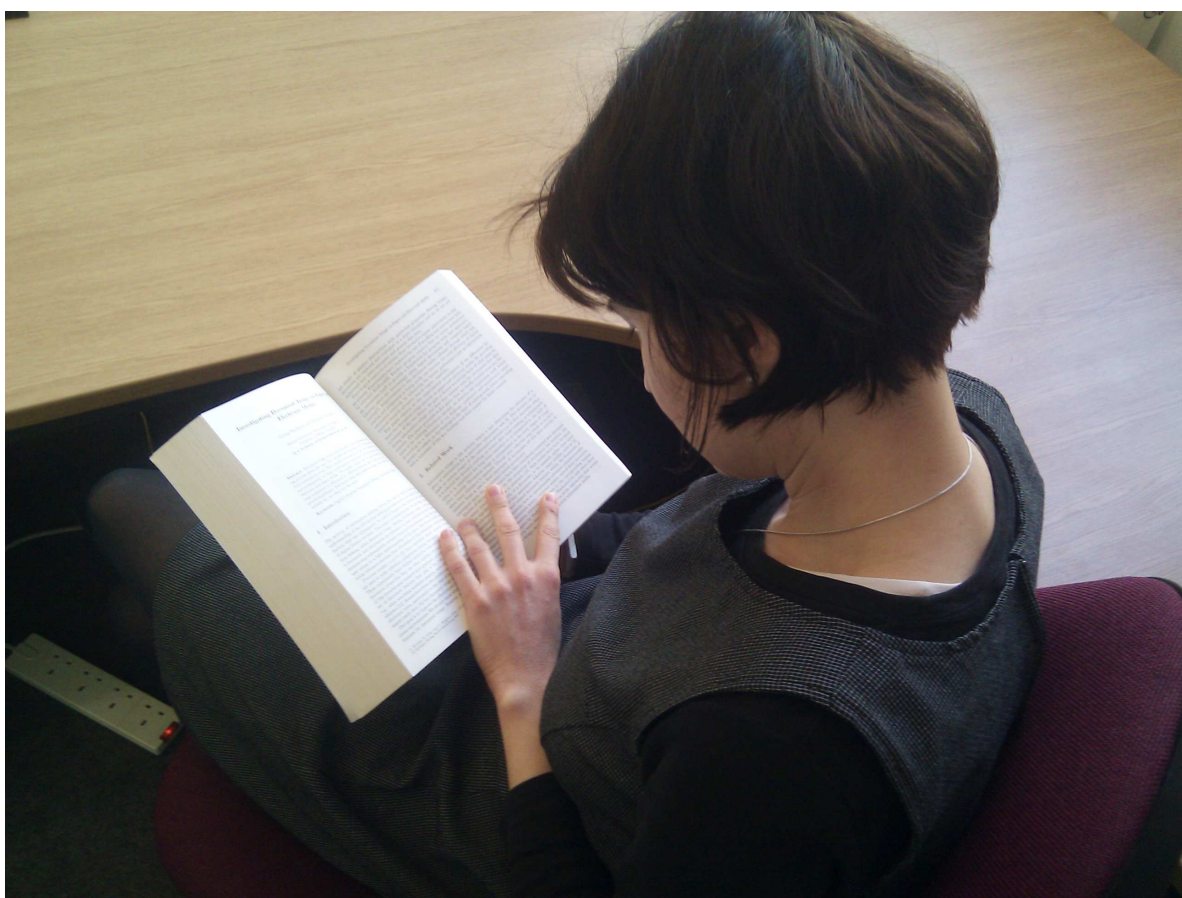
Figure 2.9: A User Reading a Document Within A Book.

It is difficult to elicit information on the document triage process at this level due to the complexity and location of the users when this stage is reached. Research on document triage during in depth reading is not reported on, neither is it within the scope of this thesis to investigate

it. We do here however review related work regarding reading activities, which help us to better understand the document triage process.

Adler et al, piloted studies into the different reading styles of users by eliciting their reading behaviours using diary logs [5]. They found that there are there are different styles of reading, some of which are related directly to the document triage behaviour. Examples are: "Reading in order to identify, Skimming, Reading to search/answer question,Reading to learn, Reading for cross-referencing. We know from early work that users are more likely to consider reading on a screen as less affordable than reading on paper [1, 28]. Interestingly, even earlier work contrasts the difference in affordance of reading on screens and paper. Some of the work shows results reporting on there being no difference between the two mediums [11, 2]. Mills and Weldon present early work which reports on characteristics of the documents on electronic text and their effect on users' reading behaviour [107]. It is reported that reading from electronic means has developed into a more "shallower, fragmented, and less concentrated reading" [93]. This gives the hypothesis that when a user can spend less time and effort triaging through documents in an electronic environment he will do so. Reading on computer/electronic displays has been thoroughly researched.

More recent research, has suggested methods and design guidelines which can support the creation of effective electronic reading technologies [115]. A similar study was conducted on graduate students, which investigated "on how they extract and record information as they read "[116]. In their work a model for document related activities is presented which identifies a read state and an information review state. These two states allow us to see that document triage can take place during the further scrutiny of a document.

There are also other factors which look into the formulation of relevance decisions, both at a cognitive as well as a behavioural level. One such factor is topic familiarity. The level at which an information seeker is familiar with the subject area he or she is researching will affect the seeking and triage process [85]. Shipman et al looks at annotations as criteria for deciding the relevance and usefulness of a passage within a document [138]. In this thesis, the focus is visual attention. The lessons learned from these studies are utilised when considering the methodology and data analysis, but are not further researched in depth.

2.5.5   Summary

We have looked at three stages where triage takes place. These stages are more thoroughly explained and examined in Chapter 8. We begin by looking at the initial contact with documents (surrogate level), where an information seeker is faced with a set of documents in summarised form. Currently, research has focused on the triage activities of web based results lists. We continue to compare and contrast our findings to validate their integrity in the academic document domain. The main focus of this thesis is in the within-document triage stage. The main reason for this is the high visual attention given by the participants at this stage and the lack of understanding and research reporting on this. Finally, a reference to the reading stage is made, which is often a neglected part of the triage process. This phase, although being harder to investigate, still plays a vital role in the ultimate decision of the relevance decisions of information seekers.

2.6   Eye-tracking

In the sections above we introduced a theoretical foundation through the satisficing strategy approach used by researchers to describe the decision making strategies of information seekers. During the review we have also described some limitations of the studies in the literature which we address in our own research. One of the main constraints in research thus far, has been extracting visual attention by observation rather than by self-reported evidence. Chapter 3 describes a method for acquiring visual attention from participants without using self-reporting. After the experiment is conducted, we recognise an even greater detail of visual recording is required.

To the best of our knowledge, there have not been eye-tracking studies that focus on the exact user's attention throughout the triage process of an academic document. Furthermore, specific document features, such as individual headings and emphasized text have not been covered by eye-tracking studies. This produces gaps in the research which we aim to contribute to with our research.

We consider eye-tracking to be suitable for our research. "Eye-movement recordings can provide a dynamic trace of where a persons attention is being directed in relation to a visual display"[80]. Using either a top-down or a bottom up approach, we can use eye-tracking to evidence "how users

make decisions and solve problems in search" and to identify which areas were "read rather than scanned" [34]. We are then able to inform our model with this data by reporting on "the sequence that users look at the visual elements in order to come to a decision"[45].

We record mainly, participants' fixation locations and timings. A "higher fixation frequency on a particular area can be indicative of greater interest in the target, such as a photograph in a news report, or it can be a sign that the target is complex in some way and more difficult to encode" [81]. In our research, we record these fixations but do not report on their effects to an information seeker's attention. The duration of a fixation is also linked to the processing-time applied to the object being fixated [80].

Besides fixations, eye-tracking techniques record saccades and scanpaths. We do not focus on these techniques as they do not provide us with the type of data we require. Saccades "cannot tell us anything about the complexity or salience of an object in the interface" [99]. Regressive saccades [6] can show us "a measure of processing difficulty during encoding" [99]. Methodological development has been mostly tested on reading which takes place on a steady page (unlike our scrolling reader) and reports well on "very small regressions, only skipping back two or three letters in reading tasks". Equally, scan paths during a search can measure the time efficiency of a participant to reach a target [57]. In document triage, we are not aware of the user having a specific target and therefore cannot use this method to extract meaningful data at this point. Unfortunately, we did not have the capability to record blink rate and pupil size which "can be used as an index of cognitive workload and effort"[25].

Eye tracking can also detect other parts, than the direct focus [66], but for the purposes of this research, we are interested in what is actually focused on by users. The only surrounding elements which can affect a users attention in a negative way, such as banners, have been reported on [30], but due to the nature of the documents we are interested in, these types of distractions should not be influential for our user pool.

We also recognise the capabilities and benefits of using eye-tracking from other studies which have been conducted successfully on similar documents, such as newspapers [111] and results lists [49]. In a recent study, memory was tested base on the visual attention of the users on academic

---

[6]backtracking eye-movements

documents [47]. The details of where the users were focusing on however, were only reported to half a paragraph of detail and therefore, we are unable to process the results to integrate them into our model.

2.7   Cognitive Psychology, Perceptions and HCI Principles

2.7.1   Introduction

Document triage is a human centred activity. Therefore, the internal workings of an individual's cognitive thinking needs to be acknowledged as a participatory factor in the decision making process. Marchionini addresses cognitive influence in the information seeking domain in that "people construct and then draw on mental models to predict the effects of contemplated actions; that is, they make inferences based on running particular mental models" [100]. Our work draws on procedural and behavioural models which work together with users' cognitive mental models to achieve the triage goal.

In this section, we examine some basic cognitive psychology principles to help understand information seekers and users' behaviour in general. Users are also engaged with information and communication systems, while performing triage. Human computer interaction provides a solid foundation by which to help identify problems and weaknesses in the systems being used. An overview of general HCI principles is also found here.

2.7.2   Cognitive Psychology and HCI Perceptions

Cognitive psychology deals with the "internal process involved in making sense of the environment" [54]. There are a series of categories which cognitive psychology studies. We are specifically interested in two of these: *Experimental Cognitive Psychology* and *Computational Cognitive Science.* The first form, looks into experiments with healthy individuals within laboratory based scenarios. The second, deals with the formulation of models to understand the human cognition.

The main area in which this research reports on is users' behaviour in terms of visual attention. Although this is the main theme and focus of the thesis, it is important to remind the reader that there are also cognitive factors that can influence or constrain the document triage and decision

making process. We therefore give a brief introduction and insight into these, since part of the results reported back from the users in our studies also contain subjective feedback which reflect these. Human cognition is an important part of the triage process. Decisions are guided by principles of cognition which are influenced by a user's external world. One example of this is the user's working (short term) memory. An information seeker needs to devote attention, hold data in thought and also process visual data while performing triage. Baddeley introduces the components of memory which allow for these operations to take place [12]. Interestingly, when one of the components is required simultaneously for two tasks, the tasks will not be performed efficiently [126]. This principle is required for the effective design of software systems. A user should thus not be required to carry out tasks which require the same memory component simultaneously. Supporting systems can be designed to support the user in such activities.

The ultimate purpose of the triage process is for the formulation of relevance decisions on documents. Decision making is a well researched area in terms of the cognitive psychology of an individual. There are several theories into how decisions are formed and models which try to predict the decision outcomes of people [155]. As Hastie states however, "Most current decision theories are designed to account for the choice of one action at one point in time" [64]. It is difficult to accept the decision making principles found in these publications as the results comparing several studies have been controversial. The models that are therefore used to support the research into document triage, are well established specialised models which have been tested and remain robust as to their conceptualisation of human behaviour (See Section 2.2).

Human Computer Interaction principles are guided heavily upon these cognitive guidelines. One example is the passing of information from sensory memory to short term memory (echoic memory). Dix uses these principles to ascertain that "we are able to focus our attention selectively, choosing to attend to one thing rather than another" [155]. By using these principles, we can apply logical and semantic constraints [44] to better user interactions with prototype applications. This, coupled with the creation of visible affordances [44] and building "simple, explicit interaction frameworks" [149] is aimed in this thesis.

2.8   Triage Software Tools

2.8.1   Introduction

Most of the work on document triage has been using manual searching, without specialised support from software. Research reports how the libraries themselves lack "to better support users overall information work in context" [3, 4]. Some work, albeit limited, has been carried out to investigate how supportive software can assist users in their information seeking activities. In this section, we take a closer look at these bespoke tools and how they affect the document triage process. We begin with an overview of information visualisation, which is employed in the individual tools presented.

2.8.2   Information Visualisation

One of the most challenging problems performing document triage from a results is the sheer amount of documents available. An information seeker is often inundated by more documents that can be possibly looked at. One way that researchers attempt to solve this problem is by using information visualisation. Information visualisation has not been restricted to the visual cues alone, but has evolved to include the interactions with the information [83]. Visualisations have, thus far, mostly been effective in a more structured or hierarchical form [127, 143, 152]. Research into query tools, utilising visualizations to a search a document corpus, has been conducted with positive results [87, 145]. This does not mean that visualisations are without problems [36], but the results are positive. Furthermore, advances in information retrieval algorithms (like the TREC conference [29, 156]), based on query terms, indirectly improve the tools that use the results themselves.

2.8.3   Visualisation Tools

2.8.3.1   ThemeScapes

Wise et al, implement a visualisation technique by employing spatial representations of large document sets [161]. Their aim was create a visualisation that "may then be visually browsed and analysed in ways that avoid language processing and that reduce the analysts' mental load". In their research, they used *Themescapes* (See Figure 2.10) ("abstract, three dimensional landscapes of information that are constructed from document corpora") and *Galaxies* ("displaying cluster

and document interrelatedness") "to present the notion of document similarity".
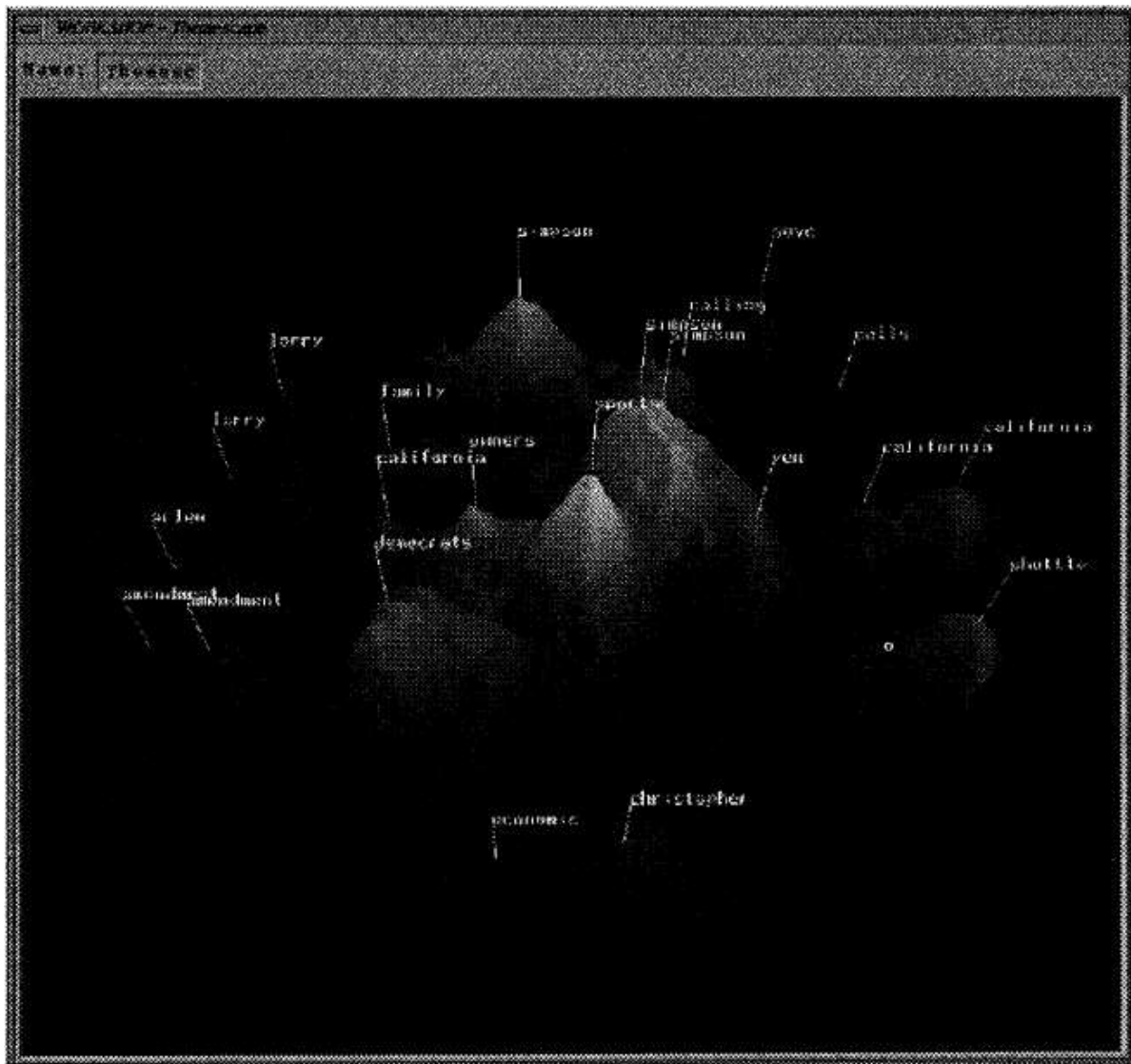
Figure 2.10: Themescape Tool

2.8.3.2 VKB

Another "interpretation of a collected body of materials" is presented by Marshall et al [104]. In this research, a spatial hypertext tool is presented which allows information seekers to interpret results from documents and identify the structure of the document set. This is made feasible by the creation of objects, composites and collections, and allowing relationships to be defined. Building upon this early work, Shipman et al, created the Visual Knowledge Builder (VKB, See

Figure 2.11) [136, 137, 139]. VKB "supports the incremental visual interpretation of information". This tool was thoroughly utilised for collaborative efforts on shared information space. Similarly, a prototype tool called SketchTrieve, was also created to assist information and document triage [67]. SketchTrieve, was based on a conceptual model which followed the pattern: select the services you need, connect them, press Run, and results will be displayed.
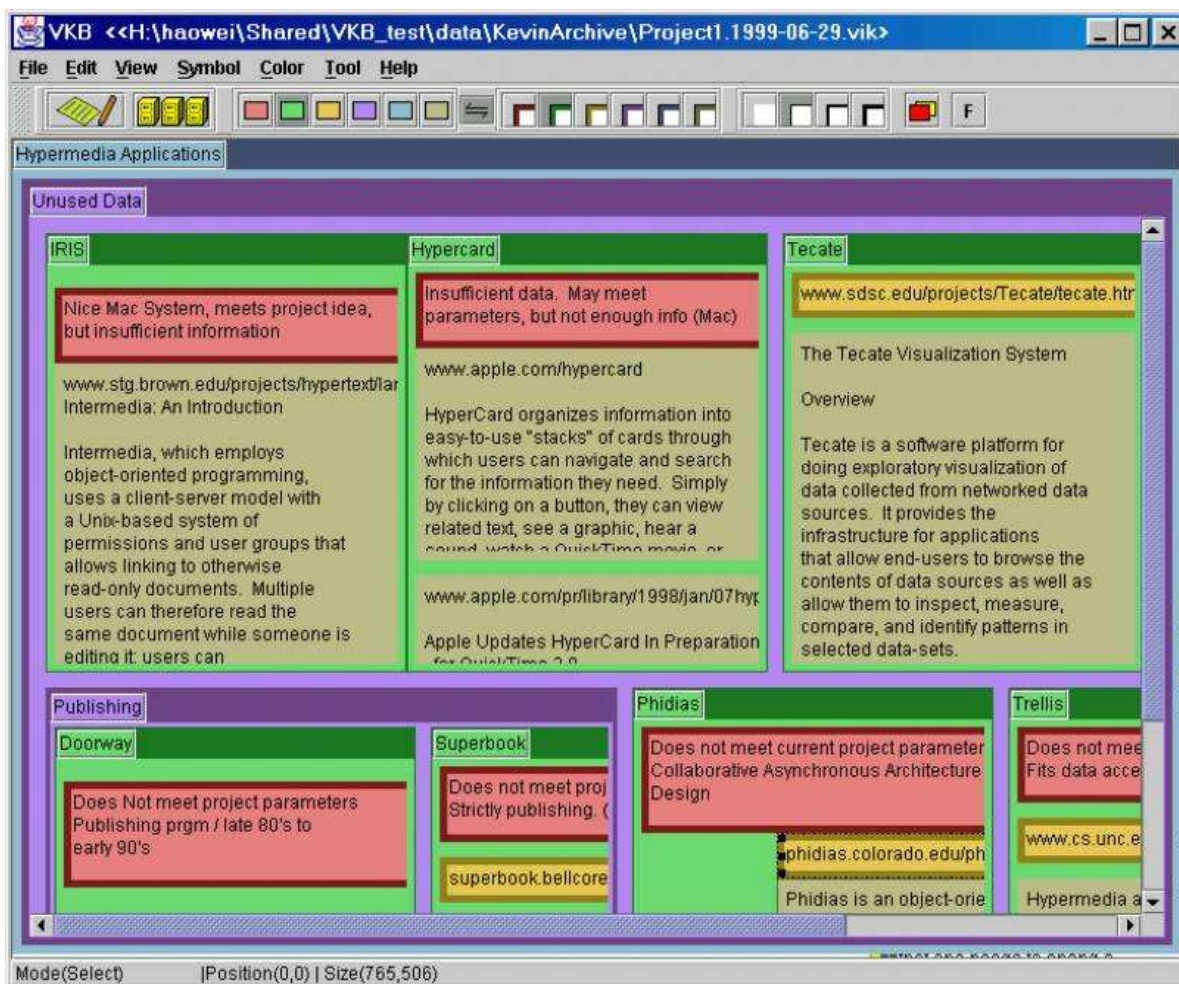
Figure 2.11: Visual Knowledge Builder Tool

2.8.3.3   TRIST

Another information visualisation tool, created specifically for information triage, is TRIST (The Rapid Information Scanning Tool) [79] (See Figure 2.12). TRIST is built on the analytical environment nSpace [162] and allows the "rapid scanning over thousands of search results in one display,

and includes multiple linked dimensions for result characterization and correlation". TRIST allows for the information seeker to compare queries and find documents that are more tailored to their need. By doing this, document triage is informed by information that would have otherwise have taken multiple steps to achieve, all within one environment.

Figure 2.12: TRIST: The Rapid Information Scanning Tool

Matching query terms to document content, like TRIST's attempt is important for information seekers. It helps them to relate their need to potentially relevant parts within a document. It is often hard however to locate the areas of the document which contain the query terms expressed by the user. A search engine will usually utilise the query terms in an information retrieval algorithm. Beyond that, there is usually no connection for the user, between the terms typed and the documents presented. Some users will use the Ctrl-F feature to find their terms within a document, but this is rarely the case [28]. It is evident that a more effective way to communicate the system's relevance decision to the information seeker is needed. One way is to match up the query terms to areas within the documents.

2.8.3.4   Using Task Bars

A common approach to supporting users' triage activities, is by enriched visual interfaces using scroll bars (or any bars representing the document length).

Two software tools, FindSkim and ProfileSkim, created visualisation in the task bars, to indicate the location of query terms in a document [61, 62, 63]. ProfileSkim, also added bar charts to allow the user to find heavily populated areas, query terms wise, within the document (See Figure 2.13). A similar basis was used by Donald Byrd, who used colour and term highlighting scrollbars [32] and Schwartz et al who used term distribution visualisations [135]. The argument for for making use of text structure when retrieving from full text documents, has also been investigated by Marti A Hearst, and a prototype "a visualization paradigm, called TileBars" is presented to aid the information seeker [65]. The same information and principle as Harper et al was implemented with some additions, such as snippets for reading the results before navigating towards the related area. This method was favourable with participants. Query term matchings have also been used in SmartFind, another hybrid Ctrl-F tool which uses Term Frequency x Inverse Document Frequency algorithms within a document to provide potentially significant document areas to the information seeker (See Figure 2.14) [96].

2.8.3.5   Opportunistic Search

Currently, query terms are the only means which an information seeker can utilise, to make a request to a search engine. Directed browsing strategies can be assisted by several methods explained above using these terms, or variations of these, formulated by the user. Opportunistic search however, is also a big part of the information seeking process. It requires the triage of information in a less structured way. This method is undertaken heavily in image searching, especially when the images are annotated [72]. As it is becoming evident that "keyword and hypertext cannot support all these new tasks well" and more opportunistic and exploratory systems are being researched [26]. One such software tool, uses Semantic fisheye views (SFEV's) to browser over collections with different metrics (See figure 2.15) [73, 74]. A similar approach was also implemented by Cockburn et al, this time, using space filling thumbnails with a zooming action to allow better space real estate [38].

One of the limitations of the above proposed prototype systems is that of screen real estate. The

Figure 2.13: ProfileSkim 2 Tool

Figure 2.14: SmartFind Tool

Figure 2.15: Semantic fisheye views
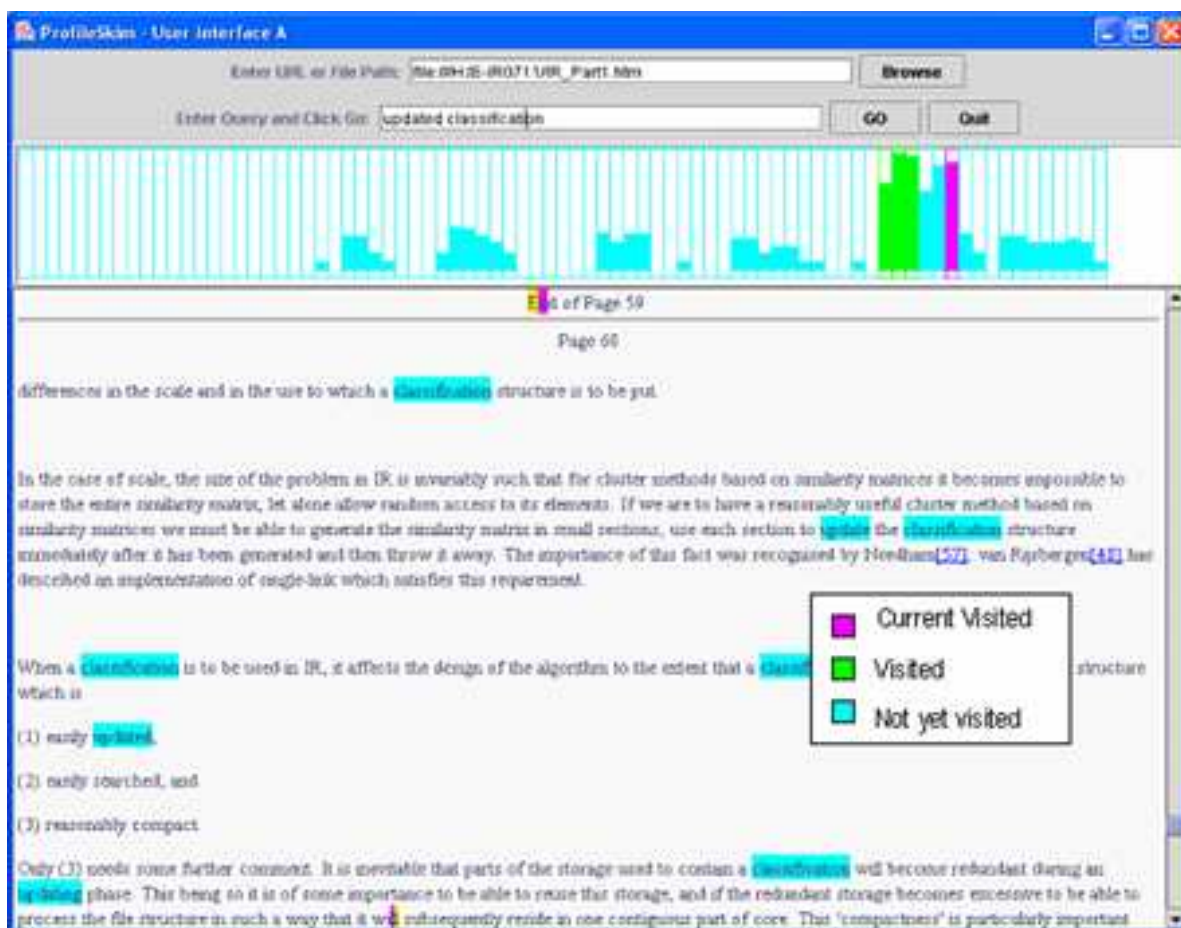
question asked by Bae et al was whether different display types, would have an effect on the way users perform document triage [14]. In their findings they report how there are more transitions using multiple displays rather than a single display. "Additionally, users evaluated documents more by reading their contents and less often relied solely on meta data. Users spent more time reading and interacting with documents that they valued". This finding corresponds with time reading correlates with assessing document value in the triage field also [35].

2.8.4 Summary

Most tools reported on for supporting document triage use a visualisation approach to present representations of the information. Interestingly, none of the tools we report on consider the individual document feature importance to the users. Furthermore, each tool focuses on extracting the amount of times that a user's search terms occur in the document rather than providing more rich representations such as document structure. While clearly they report in their findings that their methods improved triage performance measures, such as time or accuracy in locating relevant information, we argue that they can be improved even further. Using our research as a theoretical foundation, we encourage tools which give emphasis on the visual attention we report in our studies to investigate if these can increase the triage efficiency of information seekers.

2.9 Literature and Theoretical Foundation Summary

The document triage process is acknowledged to be part of the information seeking process. Beyond the digital library domain however, document triage is also influenced by other disciplines such as human computer interaction and psychology. In this chapter an overview of the related literature was presented. The aim was to equip the reader with a general understanding of the areas directly related to document triage and to give a solid foundation for the research being undertaken. The subsequent chapters build on this knowledge as more in depth studies are performed which investigate the gaps in the current research.

From the literature review we uncover the areas that have gaps in the research. As mentioned before, document triage is in general an area that has received little attention in the passed and therefore there is a large amount of elements thus far not reported on. There are three areas that we

could model and research. These are: behavioural, procedural and cognitive. From the literature the area most lacking in is behavioural and procedural. Thus far, there have been self reporting studies on the thoughts and actions of participants during the information seeking process. There have been next to no studies with quantitative observational data. We aim to extract observational data from visual studies and to model behavioural and procedural frameworks as a foundation for the document triage process.

We are also required to make a decision as to what type of documents to focus on during our studies. Among the obvious options are those of web pages, academic documents and digital books. Web pages have to a certain extent been explored, as presented in the literature. Academic documents (semi-structured documents) have also been widely reported on but only through self-reported data. The decision was therefore made to complement the existing data on semi-structured academic documents.

The medium used to interact with documents is also an element of our research that needed to be decided. Thus far, the majority of the research data on information seeking has been made on desktop screens. However, there has been a large surge towards other smaller displays in the market as well as research fields. Looking at the existing research therefore, we decided to focus our research towards document triage on desktop displays and extend this to recognise the validity of our findings on small screen devices.

reported on	mentioned but not reported
title	lists
abstract	diagrams
author	statistics
date	pictures
table of content	introduction
-	divisions
-	page numbers

One last area that we recognise gaps in, is the low level document elements and features that have not been reported on. Figure 2.9 gives us a table that summarises the elements that need specific attention, and that have not been visually and quantitatively reported on. These features,

in combination with the gaps explained in the previous paragraphs, constitute the basis for the choices we make when selecting material to investigate and gaps that we attempt to fill during our research.

The work described in this chapter testifies to the little research available specifically looking at the document triage process. We identify that document triage is a sub stage of the information seeking field, and directly recognise the stages of the information seeking process which document triage covers. Visual attention is influenced of course by the user's psychology and vice versa. We have not tried to provide reasons for the behaviour of our participants in terms of their cognitive functions but cover the literature required to guide the reader to similar related work. What we do provide is self reported pre and post study interview data to complement the findings and work alongside the related work on relevance judgements. We also introduced an area which thus far was not clearly related to document triage and has assisted in the a) analysis of document triage data and b) the evolution of new document triage tools; namely, data and information visualisation. Using all this research within the scope of interaction between digital library software and users we are able to understand all the information available.

Document triage is defined simply as the process by which information seekers decide the relevance of documents to their information need. Furthermore, most of the results have been self reported and qualitative in nature. This is the main gap in the knowledge that we have identified and investigate. Our first main question involves recording the visual attention of information seekers performing document triage on semi-structured documents using electronic means. What parts of a document participants reported using to make a relevance judgement has been previously documented. We need to empirically and quantitatively assess the extent to which the visual attention is only focused on these parts, and the patterns of visual behaviour undertaken during document triage. Once this has been established, we can conceptualise the findings in order to produce a framework and a foundation for comparison and further development. This model creation can tie together the related literature which seems fragmented, as document triage currently consists of different research areas and equip researchers with the knowledge to investigate document triage further.

Chapter 3

Empirically Exploring Document Triage Behaviour

3.1 Introduction

In our previous work [28], the differences in user behaviour between document triage on paper and electronic media were reported on. One aim of this thesis is to identify the visual features of a document which affect an information seeker's triage behaviour. In this chapter, an initial exploratory user study is presented, reporting on user navigation and behaviour in triaging electronic documents. **The focus is on how certain document features affect users' navigational choices and subjective relevance decisions.** The main aims of this chapter are to:

- Identify navigational patterns and behaviour of information seekers while performing document triage.

- Identify visual cues which affect information seekers navigational behaviours and document triage process.

- Compare the subjective feedback from information seekers against the objective results from their behaviour.

A laboratory-based study is conducted that observes users triaging documents using a custom made PDF viewer that systematically records users' interactions. An initial assessment of the

impact of common visual document features on both users' behaviour and their subjective relevance judgements is given. Using this method, common behaviours and influential document features to an information seeker's relevance decisions are identified. The deciphering of some cognitive issues of interaction with relation to document triage is also possible from the subjective feedback. This work forms a suitable starting point for our research, producing both data to cover the aims of the chapter as well as producing exploratory results to guide further research. The formation of a standardised framework can also aid in the conceptualising of the triage process and help us make predictions about how behaviour will be altered if we change elements in the process (see Chapter 8). Furthermore, system designers can be informed with guidelines for the development of effective user-centred document triage systems (See Appendix).

Firstly, a short description of the motivation for the research is provided, before describing the user study. The study is reported in three major sections: design, results, and discussion. The chapter concludes with a summary of the key findings.

3.2   Motivation

3.2.1   Automatic IR Vs. Manual Information Seeking

When discussing relevance decisions Cool et al identify a problem with traditional information retrieval definitions and criteria [39]. They state:

> "The problem appears to be that relevance, as most traditionally construed, deals with topical relationships between document and query, while the judgements that people make about the usefulness of a document to their information problems seem often to be based upon other criteria."

Information seekers as well as information providers seem to rely heavily upon the automatic process of information retrieval to facilitate their information needs. Information retrieval systems, such as search engines however, seldom address the subsequent challenge of manual user dependent relevance judgement. "...information seeking depends on the interactions among information seekers and other people and systems for representing information"[100]. This chapter identifies this gap

in our knowledge and begins to explore the manual interactions within the information seeking process and in particular that of document triage.

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 [28]. 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 [144]. Similarly, users choose only a few documents for further scrutiny, and appear to read little in detail beyond the first page [28]. This suggests that users are using only a limited amount of information to make relevance decisions. There is currently minimal research that identifies what that information is exactly.

3.2.2 Modelling Document Triage Behaviour

Information seekers have their own mental models or "personal information infrastructures" for systematic information seeking [100] in which they follow certain visual cues. There is a healthy diversity of information seeking models (e.g. [50, 91, 100]), 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 areas of information seeking and present very little data for a holistic representation of the document triage process. Some research on external factors that affect the document triage process is also reported on. [13, 14, 112].

3.2.3 Focusing on Visual Cues

Document triage can be viewed as a form of visual search. In human–computer interaction, the visibility principle asserts that the important aspects of a system must be visible [44, 78]. Applying this principle to documents, one can argue that users will be influenced by what is displayed of a document. The aim of this chapter's study was to investigate whether the factors suggested in the current literature (e.g. [39, 129]) did indeed influence how users move and view actual documents.

In the research presented in this thesis, we focus on identifying the visual features of documents

that influence user relevance judgements. 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 use that knowledge to improve support for document triage.

3.3   Study Design

In order to obtain the data we required, we chose to perform a laboratory-based observational experiment. These methodology is repeated throughout our research experiments in order to be able to compare the results [1]. Participants performed document triage on a closed corpus of documents, evaluating each for its relevance to a particular task. All documents were read using instrumented PDF reader software, which produced detailed continuous log data of the content displayed to the user. A pre-study and post-study questionnaire plus an interview collected further data for comparison with the participant's actual behaviour during the experiment. This general structure was chosen to permit the systematic assessment of the interaction between visible document content and visible document cues (e.g. headings, diagrams) and both user behaviour and relevance judgements.

In order to produce a clear understanding on the behaviour of users in terms of document triage activity, an observational empirical user study is undertaken to be able to analyse the data statistically. This aids in providing concrete evidence to support any conclusions made, based on the participants' common actions. Thoughts and qualitative feedback from the users were also recorded, in order to provide better understanding by the cognitive cues presented. Recent research [28] has shown an indication of the behaviour that users often engage in, by observational studies. This has raised questions such as 'why' and 'how often' these actions occur. The study investigates to what degree each action in the triage process is carried out, therefore having an affect in relevance decision making. Firstly, the implementation of the PDF reader software is reported on, before describing our participant selection and the study task design.

---

[1]Reasons for choosing this methodology can be found in Section 1.2.2

3.3.1 Apparatus

A custom made PDF reader was implemented using the PDFNet SDK tools provided by PDFTron.[2] The reader interface resembled that of Adobe Acrobat: e.g. main controls positioned in similar locations, and the document view display having similar defaults and behaviours. A thumbnail view to the left of the window was also available should the participants wish to use it but was not visible by default. The custom made reader allowed for the logging of detailed information on the participants' actions in an XML file in real time. It also facilitated for the investigator to pause the logging without interrupting the participant when his attention was not on the screen.

In order for researchers to acquire information from a participant reading or searching within-documents, the researcher would either need to record the whole screen and later analyse the results manually, or to time the user in real time; a process which is highly inefficient and difficult to undertake. It is a very time-staking and laborious for the researcher to carry out the experiment this way and so we chose against this methodology.

In order to unobtrusively record and extract this information, a prototype software was created, which emulates a standard PDF reader. This software, dubbed EmpTriage (See Figure 3.1), is a custom made PDF reader, which records document factors at regular intervals. By *document factors* we mean:

- File Viewed.

- Zoom Levels.

- Vertical Scroll Position.

- Horizontal Scroll Position.

The intervals at which the PDF software captures data is adjustable by the researcher. Any range can be used, from 0.001 seconds to 10 hour intervals. For real time recording, a recommended 24-30 recordings per second should be used. The main software component was written in the C# programming language (.NET Framework 2.5) using the PDFTron's PDFNet components API's[3]. This data is stored in XML(Extensible Markup Language) format. The reasons behind these are

---

[2]www.pdftron.com

[3]http://www.pdftron.com/pdfnet/index.html (2010).

Figure 3.1: EmpTriage Software: Automatically recording underlying screen contents in XML format for easy manipulation.

threefold. Firstly, XML is an open and widely used/accepted format which is well documented. Secondly, XML is easy to manipulate using any programming language. Thirdly, XML manipulation is less processor consuming than a database option and can be stored in plain text, thus reducing file size. A snippet example of an XML recording during an interval can be seen here:

```
<?xml version="1.0" encoding="UTF-8" ?>
- <Data>
- <file>
  <filepath>H:\Development\SDK\PDF\PDFNet\Samples\PDFView
  scroll\CS\bin\Debug\hci.pdf</filepath>
  <zoom>1.33949579831933</zoom>
  <Vertical>0</Vertical>
  <Horizontal>0</Horizontal>
  </file>
```

Upon the recording of a complete XML file, after the triage process is completed, the researcher can then manipulate the results easily. Microsoft Excel, is fully compatible with the output XML and the data can be imported there. Alternatively, the researcher can make use of a scripting language to extract specific data, and make calculations in real time without needing to export the data. One powerful feature of the software, is the ability to perform *instant replay on the documents viewed and participants actions.* By reversing the data of the XML file back into the reader values, a reproduction of the exact contents and view of the user can be performed. This can happen in real time, or at a sped up pace. The ability to also reproduce the exact contents of the document reader during a specific time frame in any participant's activity is also possible.

This software application is primarily targeted at researchers, since although it is fully functional, the features of the reader are not as advanced as other popular PDF readers. The applications however, can be applied to both reading and searching activities. This is especially helpful to analyse the reading stage of the document triage model, which requires a non-lab based and non-intrusive study exploration. Although the software is not open source and cannot be downloaded for free, any researcher wishing to use the software can make personal arrangements to receive it.

3.3.2 Participant Selection

Twenty participants were selected for the study. All participants had previous experience with using PDF reader software such as Adobe Acrobat Reader. They were also required to have only a limited knowledge of the topics covered in our study, to avoid effects from previous exposure to the study material. Nonetheless, they also needed sufficient expertise in the general area to be able to make credible relevance judgements. We therefore used volunteers studying at postgraduate level in a computer science discipline, avoiding the specific CHI sub-disciplines that appears in the study material. The participants' ages ranged from 21 to 38.

We chose a non-specialised group of participants. As mentioned above, we needed to make sure that the documents presented were not familiar to the participants as this would bias our experiments. We were very careful however in choosing participants that were knowledgeable enough in order to comfortably be able to understand the content of the documents. The documents were also specifically chosen so as not to be too specialised which would prevent the participants from performing a typical triage operation. We were therefore confident that the behaviour we were recording was suitable for us to report on and recognise that there may be slight variation if the information seeker is very specialised. We were interested in specifically addressing how people triage documents that are not known to them, which is the majority of the cases.

In the pre-study questionnaire, participants reported an average of 9.1 years of experience using document reader software (SD=3.04). The average number of documents participants stated they would typically triage per day was 3.7 documents (SD=4.44) while the amount of time taken per document was given as 9.15 minutes with a high standard deviation of 14.53.

3.3.3 Study Format

Ethical consideration when dealing with any users is vital. Professional bodies all have individual code of conduct and ethics procedures. Throughout our studies, we adhered to strict ethical checks and procedures to ensure that all the studies were performed properly with the proper ethical procedures. It is important to identify the ethical implications and requirements before the design of a study in order to adhere to the principles imposed. Although not exhaustive, Blandford et al have identified and summarised 3 factors for ethical consideration [23]. We use these as guidance

to inform our study designs. These are:

1. **Vulnerable participants.** These participants, like the word implies, can be more sensitive and exposed to physical or psychological pressures. Groups like children, disabled, and older people are considered vulnerable participants. Only one of the studies presented included vulnerable participants. Chapter 7 included minors reporting on their triage activities. Care was taken at all times for their legal guardian to be present during all communications. Furthermore, all the written research, by the participants as well as the researcher was screened and consented by both the minors as well as their legal guardians.

2. **Informed Consent.** While some groups may require special care when conducting studies, all participants should be treated fairly and with respect. It is common practise now to give a brief explanation and written description to the participants of the contents and background to the study. This briefing includes asking the participants to consent to the study by signing a form explaining what is asked of them. All our participants were required to fill in these forms. The individuals forms can be seen in the Appendix. Each participant also received a "bill of rights" form stating the right each participant is entitled to. This description includes rights, such as the right to be treated fairly and the right to end the experiment at any point. An example of the bill of rights can be seen in the Appendix. All participants were told before the study that the experiment was not to test their individual abilities but the effectiveness of the system or simply to get an insight into the processes they employ.

3. **Privacy, confidentiality and maintaining trust.** The privacy and confidentiality of the participants was stressed in that no names or identifying features (video or images) would be produced, unless clearly stated in the consent form. The information participants gave was anonymised and no names or personal information was kept after the interview was completed. If a quote was reported in any of the published or unpublished work, the quote would read *"Participant number x commented..."*. The raw data that was recorded from questionnaires or interviews would be kept for 1 year in physical form and for 5 years in electronic form, after which it would be disposed by destruction. Physical data would be held and viewed only by the related researchers whereas electronic data would be held by the author in two separate

hard disks for safety. Participants were informed that they could call back to be briefed on the results of the analysis 3 months after the study was completed.

Participants were given a pre-study questionnaire to complete. The first set of questions focused on establishing the participants' previous experience with document readers and document triage (frequency, duration, etc.). Rankings were elicited from each participant, regarding which parts of a document (e.g. title, abstract) they considered to be useful in assessing a document's relevance to their information need. Participants were then introduced to our custom made PDF reader and shown an example of it working by opening an example document and navigating randomly within it while using the zoom keys to indicate that the user can adjust the viewer to any size required. They were then given a training document to independently explore the use of the document reader in an open–ended familiarization task.

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 understandable. 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. We wanted to see if there was any effect in behaviour by giving participants less documents than a results list would typically give but were not able to uncover any differences. They were given 10 documents relating to the first information need in random order (the same as the first page of a results list would typically display) and 6 documents relating to the second information need, again in random order. 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 time frame 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. Special care was taken so as to not allow the participants to be able to achieve any combination on their own keyboards which would interfere with the data recording of our custom made PDF reader.

The entire computer display was recorded using screen capturing software (BB Flashback Express[4]). 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).

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.

2 documents which were abundant in features (for example contained more than 2 headings images per page on average) should rate high with the participants, while the other 2 documents considered were lacking in sufficient features (less than 1 heading or image per page) and should therefore perform poorly with the participants. These documents were specifically added as extremes, with regards the other documents, in order to explore any effects they would have to the participants' behaviour and ratings.

3.3.4 Summary

An empirical observational study was undertaken in order to gather empirical data on participants performing document triage. Users were presented with a custom made PDF reader and asked to triage some pre-selected documents and give a relevance ranking to each one, based on the information need. By recording actions and visible area characteristics in documents which affect relevance decisions and how users react to the stimulus of these properties can thus be identified. This data is coupled with a pre-study questionnaire and post-study interview to cross examine the participants' thoughts and justification of specific actions.

3.4 Results

It is important to remember that the data is describing the *visibility* of document content, not an interpretation of the user's point of gaze, as explained in Section 3.3.1. We, at this point, did not have access to an eye-tracker to continue to record more detailed information. Furthermore,

---

[4]http://www.bbsoftware.co.uk

the point of this study was to gain insight into the navigational patterns and a foundation of the visibility of the features rather than a detailed analysis.

The subjective importance that the participants gave to particular document features is presented first. These represent the perceived influence ratings that participants give to each individual document feature in terms of their value to their relevance decisions. The information can be used to compare the results with other study data and identify the items that users argue influence their decision making (e.g. [39]). The second part of the results describes the *navigational patterns* taken by participants during the document triage process. This gives the likelihood of users viewing different areas of a document. We then proceed to report on the details of the areas that were viewed by our participants. After we know the scrutiny that each document has had, we report on the participants' *relevance judgements*. The results are then discussed and compared in Section 3.5.

3.4.1   Subjective Importance of Document Features

Before focusing on the actual process of triage performed by the participants the subjective areas of feature importance is reviewed. It should also be stressed that the purpose of the study was not solely to evaluate how well the participants performed, but to gain insight into their actions and behaviour. Discrepancies between reported triage behaviour and actual behaviour through digital means are observed. It is for this reason that all the publications presented to the participants were carefully selected to be relevant to the scenario given to the users (See Section 3.3.3).

Figure 6.2 shows the subjective importance of each part of a document to the relevance decision making in the information triage process as reported by the participants. All scores are out of a maximum of 10. While the maximum score given by participants was the same (each feature was rated 10 by at least one participant), certain items were rated much higher than others. Even though there was a wide distribution of ratings for some items (e.g. body text), statistically significant differences did arise using Student's t–test: e.g. the body text versus title text reports $p<0.001$ (t=4.60,n=20). Comparing the title against the introduction yielded $p<0.05$ (t=2.25), and a similar result (t=2.35) when comparing the introduction and headings.

The findings suggest that a few features (e.g. title and heading text) would strongly influence
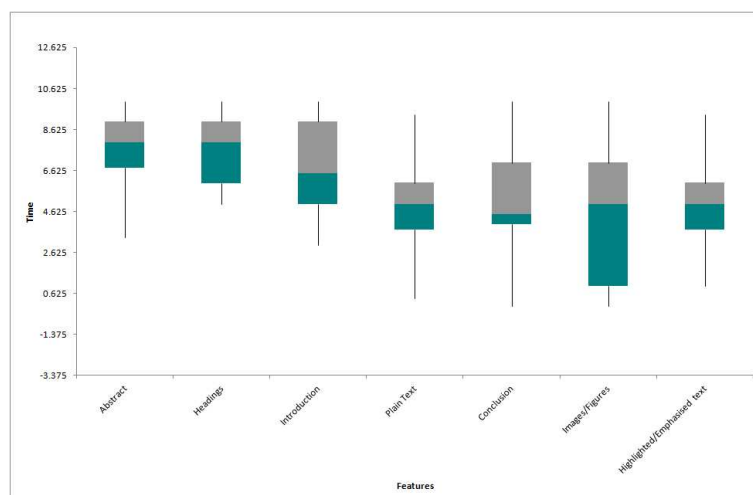
Figure 3.2: Subjective importance of document parts, Showing average, quartiles and minimum maximum.

most users, whilst many features (seen on the right of Figure 6.2) had a similar level of importance, and would possess only moderate influence. Naturally, such subjective data may not correlate well with users' actual behaviours, and these will be reported later in this section.

3.4.2 Document Navigation

From the viewing log data, a number of high-level navigation patterns were extracted. These recurred in the participants' viewing of documents. Despite being given a free choice of action, all participants viewed documents in the same order, proceeding through the list of documents in a linear fashion.

When looking at 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 order of frequency. Figure 3.3 presents caricatured time lines to illustrate the patterns visually.

The patterns shown are actual patterns from the raw data taken by the EmpTriage software. The page count may vary. They show a visual representation of the behaviour of what appeared to be the most common behaviours. In order to verify the behaviours 4 colleagues were consulted and asked to reproduce the patterns from the raw data and using a visual inspection technique. They verified that they would reproduce the same. One colleague, an expert in visualisation, confirmed
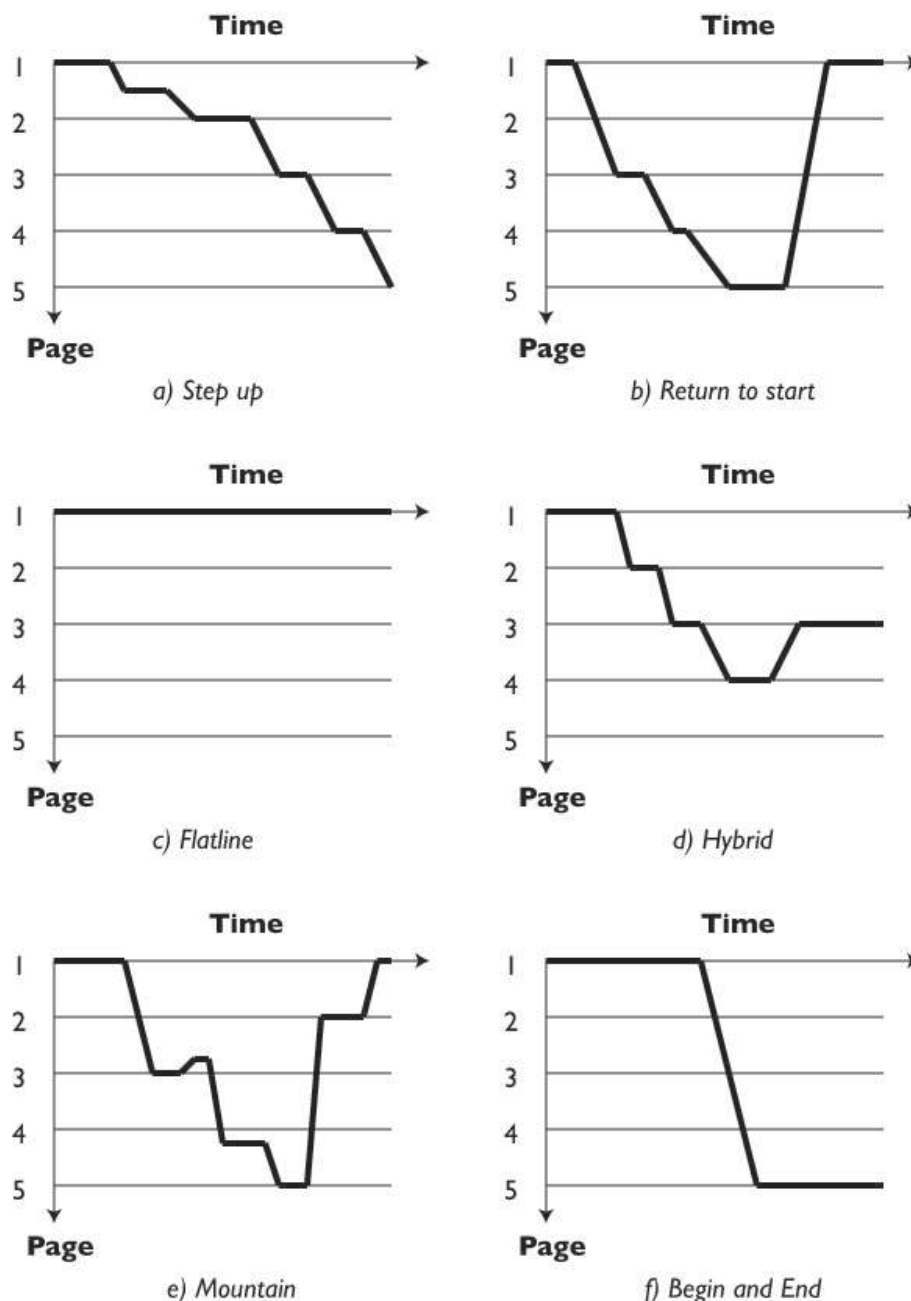
Figure 3.3: Navigational Behaviors

the patterns and findings. Further work can be found in [164]. In terms of mathematically proving the navigations, this is a difficult challenge in its own right.

Step Up Navigation a term we coined for the predominant viewing pattern, was reported on 229 document passes, and observed in all 20 participants' triage tasks. As seen on Figure 3.3a, the initial part of the document (usually the initial page) is viewed for a long period without any

scrolling. The remaining document receives is mostly skimmed, with specific areas being focused on when the user stops 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 3.4.3). 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 viewing is forward through the document, but with attention unequally spread over the content.

The second most common behaviour was the **flatline** pattern (see Figure 3.3c). Here users simply scrutinize the first page or visible part of the document. In contrast to the step–up pattern, there is no subsequent scrolling to later document content. This was performed on 73 documents and in the case of 18 individual 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 3.3e). In the 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 final behaviour observed is the **beginning and end** pattern. 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 is portrayed in Figure 3.3f and was observed on 33 occasions in all 20 unique 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(see figure 3.3d). We use the word hybrid to explain how a navigation behaviour can include clearly elements from two navigational behaviours or more, but do not represent one of the other models

in their entirety. Furthermore all the models for document triage apart from the flatline pattern have a common variation (see Figure 3.3b) 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 some 23 of these views were for less than one second.

3.4.3   Document Features & Navigation

We will now look at the impact of document content and visual features on the users' viewing behaviours. First, the characterisation of user behaviour on pages that lack any of the features being investigated are presented. We then report the viewing data for each of these features individually. Summarized figures are found in Figure 3.4.
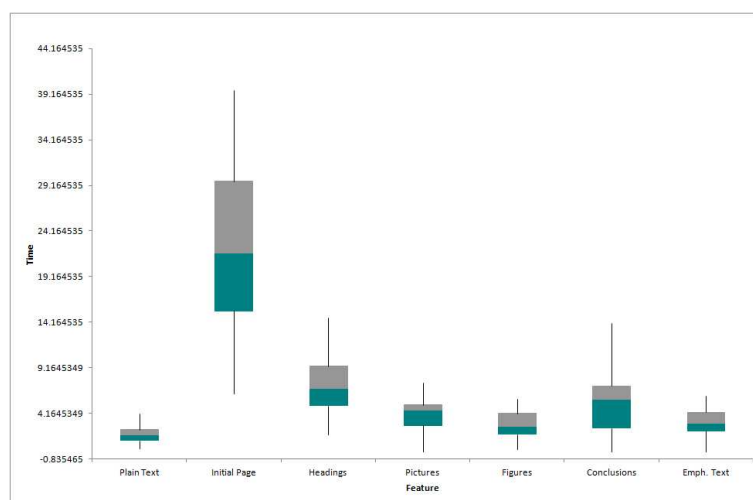
Figure 3.4: visible time per document feature (in seconds). Average, Quartiles, Minimum and Maximum

Document Length Greater document length does result in, on average, increased viewing times. The general view can be seen in Figure 3.5, 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). Therefore, the evaluation of 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 are reported on. Taking the entire viewing history of all documents, we see 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.
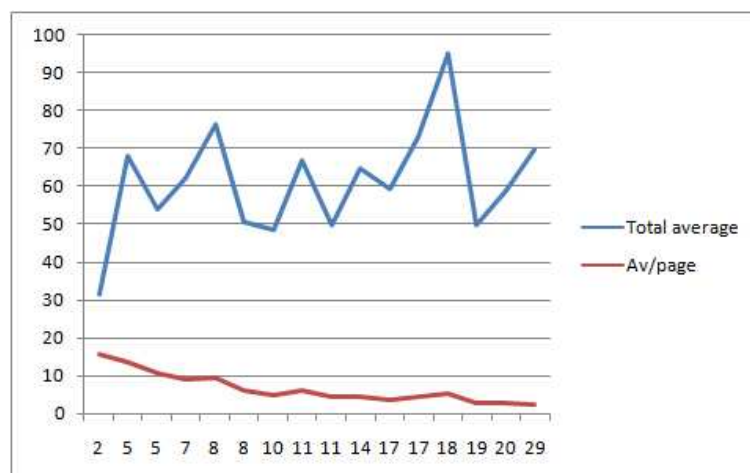
Figure 3.5: Timings per document

Featureless views Areas without any noticeable properties (i.e. plain text only - See Figure 3.6) were typically scrolled past by the users at a fast rate. This brief viewing time makes it unlikely that any reading or even superficial skimming of the page occurred that could contribute to their relevance judgement. The average viewing time for a page view that contained plain text only was over one second (1.38s). This figure is rather skewed by one particular page: the first page of the one featureless document. On average, the initial page was viewed for 23.5s. The remaining views are highly bimodal, with a large number of views appearing for less than 0.3s. In total, a plain page being moved past rapidly would be in view for less than 0.5s in total. In contrast, is a set of more prolonged page views with an average of 2.57s. The starkness of this is readily made clear when compared with the average time in view of all pages, which was 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 judgements [128]. These two parts of the document are most often located on the first page with a significant
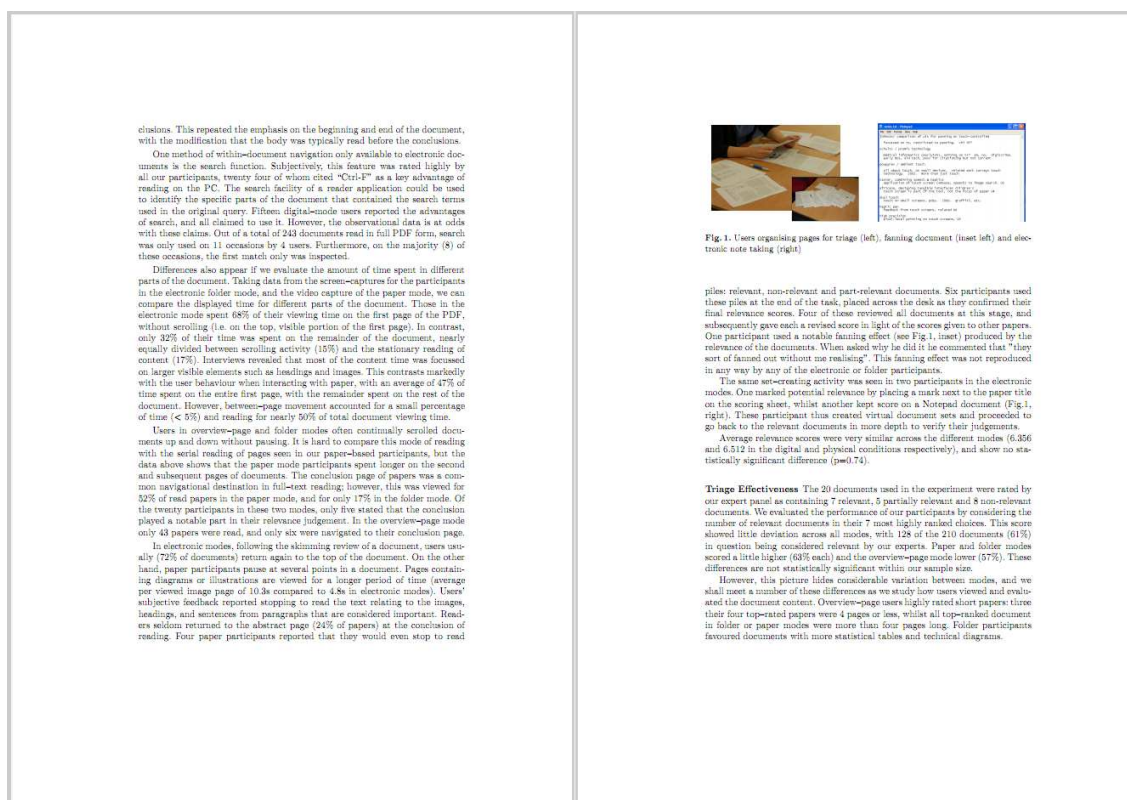
Figure 3.6: One the left: a featureless page. One the right: A page containing and image, emphasized text and/or heading.

proportion of the introduction. Across all documents, the viewing of the first page accounted for 48.2% (SD=15.24%, n=320) of all triage time. This immediately suggests the first page content plays a central role in the relevance decision making of users. The average time a participant spent on the initial page was 23.21 seconds. Again, this is a contrast with the average for all pages of 5.28s. As mentioned in Section 3.4.1, users relied solely on these three parts for making their relevance decision on 43 documents (see Section 3.4.2), 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 behavior on 11 documents each.

As noted in the navigation section above, the participants frequently returned to the beginning of a document. This occurred in the triage of 77 documents. In 23 instances the view of the last page was extremely brief ($< 1s$). 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 therefore reflect more than a cursory view, 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, the first page was discounted 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. Considering the more focused data of views, this correlation increases slightly to r=0.322 (df=4018,p<0.01). This suggests that headings have a modest impact on users' navigational 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). Applying Welsh's t-test to these sets 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 >= 0.3s of r=0.354. Thus, one can conclude that the presence of a heading does increase the period of time that a user will study the view or page, and also that it increases 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 result is between 0.1 and 0.3, which implies some (albeit limited) positive correlation between the area occupied by the headings and the and visual attention [5].

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 from such material, and this is further confirmed by applying Welsh's t–test on pages with figures versus those without, where a minimal difference in average viewing time (4.18s versus 4.38s) fails 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 but not headings. When the documents included emphasized text, these areas were visible on average 7.05% of the participants' time with a standard deviation of 3.56. On average participants

---

[5] More on this finding in Section 3.5

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 Welsh's t–test against all other pages results in $p<0.0001$ (t=4.36), confirming an effect, and Pearson's r is 0.432 [6]. No correlation or effect is discernible when testing for either views of $>= 0.3s$ (e.g. Pearsons=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 Welsh's t-test to determine an effect on the number of participants who viewed the page results in p=0.051 (onetailed,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 neighbouring ones). The observations of the study suggested, and the navigation patterns reported 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 fourteen documents that possessed on. This represents a rate of just over 60% being viewed. It should be noted that this number includes *all* conclusions that were displayed for any length of time, and not all were scrutinized in detail.

3.4.3.1   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

---

[6]N.B. this test would be considered inappropriate on a binary value, but its result is included for completeness

times and, in the case of headings particularly, an increasing likelihood of the page being displayed for more than a brief overview.

3.5 Discussion

In principle, the documents presented to the participants were all rated relevant to the tasks by subject experts. We decided to focus our attention on relevant documents. The reason for this is that we wanted to focus on where users will visually scan documents in general, but begin by exploring what visual path of users on relevant documents. We can then compare the relevance decisions and distinguish which documents, although being relevant, do not give enough information to the user to rate them appropriately. The reason is that ultimately we need to help users identify useful documents. Subsequent studies continue to address documents that vary in actual relevance. Nominally, the relevance ratings would be expected to be consistently high. However, **we found effects from visual document features on navigation and relevance, and that users' navigation follows some simple patterns.**

As reported in Chapter 2, there is a limited but sound corpus of literature dealing with the process of document triage. The findings agree with the current body of evidence. For example, the results concur with the reported importance of the main title and abstract.

3.5.1 Navigational Behaviour

There are four distinct types of navigational behaviour identified that occur during document triage. By far the most common is the *step up* pattern (see Section 3.4.2). 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 3.4.3). 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 3.4.2). These do not allow the information seeker to interact with the entire content of the document and may result in the information seeker making poor relevance decisions.

3.5.2   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. 3.4.3). While content text is important, the evidence from other visual cues leads the reader's attention to a particular point. Participant 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 [95] has suggested that electronic reading is used less for in–depth reading. The data from the 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 [32] and Harper [63] have both endeavoured 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 relevance of 8 out of 10. The Abstract and the Introduction ranked 3rd and 4th 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 observed 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%. Therefore, concrete data for viewing times of visual data can be added to the findings of Saracevic, Cool, and others [28, 75, 128].

Subjective feedback reported that obscure and misleading titles are a negative influence in

ranking as well as time consuming since the participants noted having to look further to distinguish the relevance of a document. Similarly, the abstracts of some documents were criticized for being imprecise or too large. Participant 6 commented that he would "prefer a short abstract which doesn't over-analyse the paper". Participant 13 enforced both the abstract and main title views by saying "I love a short simple abstract which gets to the point, and a good main title". Due to the format of the study, impact of these different factors cannot be discriminated in detail.

Headings have long been held as a significant factor in document triage[39]. The participants rated headings as the second highest visual factor when making a relevance decision (see Section 3.4.1). Participants repeatedly observed during interview that "headings are really important". However, headings alone may not be 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: e.g participant 15, commented "I don't like this, there are no headings or anything". Participant 6 said: "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. This phenomenon has been observed only once before [28]. Participant 1 explained that "pictures give you a better idea and makes a paper more appealing to someone". Obscure diagrams attracted considerable criticism. Participant 8 noted that reading some of the figures was simply "too much effort". Figures did not produce a significant impact on user behaviour.

Emphasized Text Emphasized text received a moderate rating from the participants during the pre-study questionnaire. This shows awareness on behalf of the participants for set targets for likely information gain. The acquisition of this information can result in accurate and timely relevance judgements. Time spent on parts of the documents containing emphasized text was low, however, proportional to the amount of text this document feature occupied the amount of time is quite substantial allowing for a focus point to be established and analysed on this text(see Section 3.4.3). Documents containing emphasized text as one of the dominant features were rated high by the participants, suggesting perhaps a tendency to read more in depth when the user is presented with a short amount of text. This behaviour is also noted in [96] where participants were comfortable with focused text results rather than larger chunks of text to read through for

document triage and relevance decision making. Participant 5 noted that emphasized text will capture attention. He stressed this by bringing the example of how "section headings should be emphasized in bold" especially.

Conclusions Section The conclusions section had a tangible impact on users' navigation patterns. Despite falling late in the document, conclusions received significantly more viewing time than nearby material. Nonetheless, 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. This is rather at odds with the pre–study questionnaire ratings. The observed behavior is consistent with reading of conclusions being informative and summative, rather than a key point in a user's triage decision.

Keywords is an often neglected part of a document that is often only given attention for categorization or for automatic information retrieval. 4 participants shared the view reported by participant 2 in that "Keywords are a useful part of the document" and aids in gaining insight into the context of the document. As the keywords are located on the initial page they are categorized with the figures produced for the main title, abstract and introduction.

A **within document searching find feature** (*Ctrl-F*) was used by 3 of the participants, all of which gave no worthwhile result. One participant, upon opening the dialogue window of *Ctrl-F*, said: "I don't know what terms to use" and proceeded to close the dialogue window. This shows that although the 'find' tool does not seem to be a useful tool in information triage it is often sought after for assistance. Tools to assist a user which advance this basic search feature have been introduced in [62, 63, 96].

3.5.2.1   Summary

Liu [95] characterized digital reading as hasty and incomplete. It seems that this characterization 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 berry picking behaviour [17] 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 "honey trap" 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, and hence further work will be needed to finally prove or disprove this hypothesis.

3.6 Conclusion

3.6.1 Overview

Through a controlled laboratory–based observational study, empirical data on visual and navigational patterns was collected. Our initial aim was achieved in that **navigation patterns of users when viewing documents for triage were produced**. Furthermore, by combining self–reported significances of visual document features with the detailed navigational logs provided by the instrumented PDF reader, **it was feasible to produce a representation of document triage as a visual search task** which was our second aim. A number of minor **differences between subjective ratings and actual behaviour were uncovered**, thus fulfilling our third aim. The study data confirms the importance of known visual cues, whilst illuminating some distinctions between them.

3.6.2 Summarisation and Visual Guidance

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. As Marchionini states:

> ...highly interactive electronic information systems are causing incremental changes in how we seek, acquire and use information. The high volume and diverse forms of information demand better tools, which in turn change our behaviours, expectations and attitudes.[100]

Several overview summarisation techniques emphasising the important and beneficial features to users may be implemented on a larger scale, when searching through a corpus of documents rather than within-documents, in order to improve the efficiency of the document triage process.

Headings were considered of high importance (15%) by the participants in the pre-study question-naire. Headings are reported as an incentive for information seekers to proceed in further scrutiny of the document rather than making a relevance decision simply based on them.

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 may vary between the media concerned and the information needs and skills of the users (see Chapters 5 and 7). 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"[160] (see Chapter 8).

Finally, the need for a more detailed analysis of visual attention is identified. The strong evidence that an information seeker is affected by visual cues invites further detailed research into their influence. In this study, we acknowledge a limitation which prevents us from extracting as much detail as we would like on the visual features within a document. Using the bespoke software tool, we were able to identify zoom levels and areas visible on the information seekers' screen. Although this allowed us to infer the likely relevance of each of the features, a more detailed scrutiny of the users' eye-gaze is needed to confirm the findings. This will allow us to verify our results further and also to uncover the details as to how, and for how long, the individual features are viewed. A study tailored to facilitate these questions is presented in the next chapter.

Eye-Tracking Document Triage

4.1 Introduction

In the previous chapter a study was performed that modelled navigational patterns of users while performing document triage, and the influence of document elements to these patterns [97]. The focus on that study encompassed how information seekers are affected by the elements on a document when performing document triage. Using bespoke document reader software, data was recorded by capturing the location of the participants' window views on the documents as well as the zoom levels. It was empirically informative in quantifying the time spent by the participants on certain areas of a document. We were able to report which features influence an information seeker's visual attention. This study unveiled the complexities of the document triage process. Although we had evidence of which area was being viewed, and which features were found within that area, we were unable to report on which individual features were viewed specifically. For example, when more than one heading was displayed in the participants' visual space, it was impossible to distinguish which one (if any) received the information seeker's focus. Was there any subsequent reading after a relevant heading had been identified? We know that the initial page is an important part of the triage process. For example, it was reported in the previous literature that the abstract and main title (which are found on the initial page) are important to relevance decisions [128]. We are not able to find any details on other parts of the initial page such as the introduction, the author and key words, as to their effect on the triage process. We also know from our previous chapter that

the conclusion section receives a high degree of visual attention. Which part of the conclusion is scrutinised is left unanswered. Plain text, which constitutes the majority of an academic document is reported as being skimmed at a fast rate. Our 'Step navigational pattern' shows information seekers do pause to focus their attention on document sections which include plain text and other document features. We are unaware if any of the plain text is being read, or if only the other document features are scrutinised.

In this chapter, we report more detailed data on individual features of a document regarding the users' visual attention. Like our previous study, participants were given semi structured documents and asked to rate the document relevance based on an information need. Comprehensive gaze data was recorded using an eye-tracker. The reason why we did not use an eye-tracker in the previous chapter was simply that we did not have access to one. Seeing the results from our study in the previous chapter however, we identified that an eye-tracker was needed for subsequent studies.

By identifying the exact visual behaviour from users, we can contribute this information to our initial conceptual model (See Chapter 8). Although visual attention on a feature does not evidence a positive or negative contribution to a relevance decision, it does provide us with the knowledge that the feature has been taken into consideration. We are then able to tailor studies specific to pinpointing which features produce bottlenecks and cause misinformation which can lead to false relevance decisions. In turn, this information can be utilised by designers and developers to inform supportive triage software. It is our intention to promote future research into software which encourages a more human involvement and interaction in the relevance decision process rather than mostly relying on automatic information retrieval results.

4.2   Related Work

Document triage is classified as a form of visual search. A human computer interaction visibility principle states that important elements of a system must be visible [78, 112]. We can also infer from this that information seekers will be influenced by what is displayed to them. Some work has studied the natural process of looking for information without external influence [13] whereby users' activities and reading was taken into account. Marshall et al studies the triage of information of large sets of documents with time constraints [103], while Saracevic allows for the inference of

feature influence to relevance decisions by gradual presentation of different document features to the users [129]. Bae et al looks at the effects of different screen displays on the triage process [14]. Cool et al observe specific characteristics of documents such as format, topic and contents regarding their influence to relevance judgements [39]. Wacholder et al [157], as well as our previous work [28], also suggest differences between paper and electronic information triage. Search time is primarily determined by two factors: the number of features which contain the psychological subject of the target icon, and the depth to which the features and their associated text and images must be evaluated before they can be rejected or accepted as the target [106].

We borrow the findings from related work to guide our study design and assist in the sense-making of our data. Spink et al describe a phenomenon of selectivity which is how users will mainly focus on the first two pages of a results list [144]. If we apply the same principles to the research presented in this thesis, we can identify common behaviours. For example, when performing document triage within a document, the information seeker will focus heavily on the initial page and drop his attention on subsequent pages [28]. Both triage processes described in [28] and [39] look at a differently structured set of documents. By reading the self reported data from participants, from our work and from related work, we hypothesised that users make relevance decisions early on. The participants were asked to explain their triage behaviour and replied that triage beyond the initial page is usually done in order to verify their initial decision, rather than rather than to formulate it.

Previous research by Saracevic discusses elements such as titles and abstracts of documents that affect a user's relevance decisions and document triage process [129]. Although insightful and a good foundation, this research is limited to only few of the properties of the documents. We extended Saracevic's work to focus upon the scrutiny of more document elements and to compare between paper and electronic media [28]. We had mainly focused however, on differences between the two mediums, while only touching upon the document elements and mainly relying upon the participant's qualitative feedback. In the previous chapter, a bespoke software tool was created to record the exact times and visible document area. We were thus able to empirically report on document features which visually affect the navigational patterns of an information seeker [97]. It was made evident however, that more detail was required to investigate the exact focus within the

visible area of information seekers.

Looking into document elements of academic documents was deemed a suitable starting point since the authors and majority of readers are familiar with this format. Furthermore, following the same user study format as our previous study, the results are comparable. We wish to remind the reader, from our own work and from related work, the basic findings on document features. We have so far identified the importance of the initial page. Saracevic has also reported on studies which attest to the importance of the above mentioned features [128]. The exact visual attention and a comparison of the importance of each however, is not reported on previously. Headings were subjectively considered of high importance according to pre-study questionnaire responses by participants. However, statistically these were at odds with the actual behaviour of the participants [97]. Headings played a modest role in attracting the attention of the information seeker. What a user does visually after he or she locates a relevant heading however, is not known. In this chapter, we introduce the visual importance of the first few lines after a heading. By inspecting the gaze opacities, we found that text within the first few (5 lines) of the plain text after a heading would be scrutinised by more than 75% of the participants, after which there was a sharp drop in visual attention. Plain text carries a very low value in terms of a user's attention [97]. We do not know however if plain text which occurs around recognised areas of interest, such as headings is an exception to this. According to our previous studies, images are less informative than table figures. Emphasised text played a marginal role in user attention. Questions arise as to whether the actual table/figure/emphasised text is scrutinised or if the users then continues to read the surrounding text. The conclusion section receives substantial visual attention. Although general visual attention deteriorates progressively with page number the conclusion section produces a peak in the viewing time. Which parts of the conclusion are looked at has not been reported on. In general, research lacked detail in terms of the exact gaze and focus of the participants within a viewable area. If two features were inside the viewable area then it was impossible to identify which one of the elements the participant was looking at. This lack of detail needed to be addressed in order for an informative model to be produced. In this chapter the granularity of our data allows us to describe findings which cover all the above gaps.

4.3 Study Design

4.3.1 Introduction

In order to observe the impact of individual features on the navigation and attention of the participants, a controlled lab study similar to the study in the previous chapter was performed. We used an eye-tracker to record participants' visual attention. This methodological approach restricts the participants' freedom of head movement and thus reduces the time which a user can comfortably sit in front of the screen. For this reason, although the type of documents triaged remained the same, we reduced the number of documents presented to the participants. Firstly, the participant selection process is explained, followed by the pre-study questionnaire, main study specifications and post-study interview.

4.3.2 Participant Selection

Twenty participants, 7 female and 13 male, were chosen for the study. All participants were of postgraduate level or above in a technology related discipline (due to the light information technology content of the documents). Their ages ranged from 21 to 50 years of age. Our participants were required to be familiar with academic documents and have a working knowledge of human-computer interaction (the theme of the documents used). Special care was taken however, in order that the participants were not directly familiar with the specific documents presented to them in the study. All the participants were given £10 as a thank you for their participation.

4.3.3 Pre Study Description

Participants were briefed with an explanation of the term 'document triage'. A pre-study questionnaire was given to the participants containing general questions such as: age, average documents triaged per day and time spent on document triage daily. We also asked the participants to give a subjective rating of the different document features regarding their importance to the document triage process. Participants were also asked to choose their medium of choice for document triage (electronic or physical).

4.3.4   Main Study Description

The main study included the participants performing document triage on a corpus of 6 documents and giving a relevance rating on each, based on an information need. The information need given to the participants was in the form of a scenario in which they would have to pretend to be a character, in this case a teacher, looking for material involving specific tablet technology which would assist in teaching in a classroom environment. The documents chosen varied in length, relevance and the number of features they contained. The summary of the properties of the six documents can be seen in Figure 4.1. They were presented in a random order and the participants were instructed to report if they were familiar with any of them.

The four categories the participants could place the document in were: Relevant, Somewhat (Mostly) Relevant, Somewhat (Mostly) Irrelevant and Not Relevant. Two documents were relevant, two not relevant and two were somewhat relevant. The document relevance was determined by three information seeking experts, with regards to the amount of information they contained to the information need.

Doc. No.	Relevance	Headings	Doc. Length	Expert Average Rating
1	High	14	4 Pages	8
2	High	16	8 Pages	9
3	None	7	10 Pages	2
4	Somewhat	16	34 Pages	5
5	None	5	2 Pages	1
6	Somewhat	18	8 Pages	6

Table 4.1: Summary of Document Properties

Participants were given an open-ended time frame in order to complete the task but were reminded that this was not an in depth reading exercise. The documents, in Portable Document Format (PDF), were presented one at a time on a 19 inch screen. The participants' eye gaze was tracked using the non-intrusive Tobii x50 eye-tracker. The Tobii x50 allows for 0.5-0.7 degree accuracy, a spatial resolution of 0.35 degrees, a drift of < 1 degree and a frame rate of 50 Hz. The participants' freedom of head movement was 30 x 15 x 20 cm at 60 cm from tracker. For this reason, the amount of documents used for triage was limited to six, thus not causing any discomfort to the participants. The computer used was a Dell OptiPlex GX620 3GHz with ATI RADEON X600 256Mb dual head graphics card, 1Gb RAM and running Windows XP SP2. A dictionary was

supplied for the clarification of unknown words and questions could be asked at any point in the study to further clarify their goal. Guidance regarding the documents themselves however was not given.

4.3.5 Post Study Interview and Summary

Following the main study a semi-structured interview was performed, in order to gain qualitative feedback from the participants. Questions that were covered included 'how close participants were to a decision after the initial page?' and 'when was an initial mental decision made and when was a final decision clear?'. Participants were asked to provide a description of their experiences doing the tasks after each or all of the tasks were completed, a method called Retrospective Think Aloud (RTA) [70]. This way, we did not interrupt or influence their interactions during their study. Some participants, could not remember all their interactions clearly. For these participants, we replayed their behaviour from the experiment to them as part of the post-study interview. Using this playback as a prompt, we then elicited further information on the reasons for their behaviour and to better understand their decision-making processes. [148].

4.4 Results

4.4.1 Introduction

In this section we report the data recorded from the study. For a discussion of these results, and a comparison with our previous results, see Section 4.5. We begin by looking at the subjective feedback from the participants obtained before the study and then continue onwards to look at the empirical data to report the impact they have had on the users in terms of actual attention and navigation times. We then conclude by looking at the document ratings given by the participants.

Gaze Opacities. Throughout the section images of gaze opacities are presented. These are extracts from eye gaze charts which give a representation of actual user visual attention. The figures presented here synthesize a view of all 20 participants on a single example document feature. They display what participants were focused on when looking at a feature: Darker areas represent less attention time, while brighter, clearer areas, represent more attention. These images summarise

the main behaviour of information seekers' interaction during triage.

4.4.2   Subjective Importance

Table 5.3.1 shows the subjective importance of document features according to the participants in their pre-study questionnaire. The scores are out of a maximum of ten (ten being extremely important and zero being not important). The *Main Title* and *Abstract* are considered by the participants as the two most important features subjectively, with average scores of 7.6 and 7.1 respectively, followed by the *Headings* (6.85), *Conclusions section* (6.7) and the *Introduction* (6.6). Images and tables rate low in subjective importance according to the participants (averaging a score of 5). The least useful feature appeared to be the references with 3.7. It is interesting to note that both the references and the figures are rated by our participants on average as less relevant than the plain text (5.3).

When asked whether participants preferred to triage on paper or using electronic media, the majority of participants (15 out of 20) preferred electronic triage while one was undecided. All participants had over one year experience in using document readers and claimed they would triage an average of 12 documents per day, spending an average of just under 6 minutes on each document.

	Title	Abstract	Headings	Intro	Plaintext	Conc.	Refs.	Figs.	Emph.
Minimum	3.00	2.00	2.00	3.00	1.00	2.00	1.00	2.00	3.00
Average	7.60	7.10	6.85	6.60	5.30	6.70	3.70	5.00	6.05
Maximum	10.00	10.00	10.00	10.00	10.00	10.00	8.00	9.00	8.00

Table 4.2: Subjective Importance of Document Features

4.4.3   Document Features and Attention

In this section, the impact of the visual features of the documents to the participants' attention during document triage is reported. We are able to identify the exact visual focus of users in real time to at least 0.5cm accuracy. We therefore record the areas of interest by the fixation time and also the area they occupy. We report both these measures and also present the visual processing ratio to give us an understanding of which areas take up the most attention. We do not report on importance in the traditional sense, meaning that something is important in the decision making process. Instead, we give a visual importance to the features in terms of their fixation spatial density

as reported by Cowen et al [41]. We measure the area features occupy and the amount of attention they receive and give a ratio between the two to account for a density of the fixations. Cowen et al report "how fixations concentrated in a small area indicate focussed and efficient searching whereas evenly spread fixations reflect widespread and inefficient search". Furthermore, it is argued that "a longer fixation duration indicates difficulty in extracting information, or it means that the object is more engaging in some way" [80]. Since we want to report both features which are likely to assist the user or not, fixation duration is an important measure to us.

We begin by reporting on participants visual scrutiny of the initial document page. We then continue to give findings on users' visual focus on individual features such as headings and images. It is worth noting that although we anticipate some of our findings to repeat the data reported in the previous chapter, we will now be able to validate them empirically also. We are using an inductive method to verify the validity of our previous results, but also present them with a finer granularity and with data previously unreported.

4.4.3.1 Features with Relation to the Initial Page

	Title	Abstract	Abstract (*)	Keywords	Authors	Intro.	Intro. (*)
AAO	14.05	12.57	3.46	5.48	7.20	52.05	3.70
ATV	12.12	43.18	19.3	4.17	5.46	26.91	7.07
ATV/AAO	0.86	3.44	5.58	0.76	0.76	0.52	1.91

Table 4.3: *** denotes the first 5 lines after the heading of the feature described.**) Initial Page Features: Average Area Occupied (AAO) % versus Average Time Viewed (ATV) % (see Section 8.5.2)

Initial Page It was reported in our previous study, and in related literature, that the initial page is considered highly important to the document triage process and is closely scrutinised by information seekers [28, 97] (also see Chapter 3). Other researchers have noted that abstracts and the main title play an important role in the relevance judgements of information seekers [129]. The structure of the initial page of an academic document is highly standardized and almost always contains certain elements: e.g. the main title, keywords, authors, abstract and all, or part of, the introduction. The initial page also contains other content which may also appears on other parts of the document such as diagrams and headings.

The results strongly correlate with the data from our previous study (Chapter 3). The eye-

tracking data reports 48% of the total triage time was spent scrutinising the first page, an average of 44.90 seconds. In our previous study, the initial page was on screen for 48.4% of the triage time.

Table 4.3 summarizes the data concerning the participants visual focus on the initial page. The most viewed feature within the first page, and therefore the most influential, is the *Abstract*. Although the abstract accounts for an average of 12.57% of the area on the initial page, participants spent an average of 43.18% of their time on their first page viewing the abstract. The share of document area is much less than the share of viewing time, suggesting a high impact of the abstract to the attention of the information seeker.

Until now, the abstract was reported as being an important feature in the document triage process as a whole. Our data however indicates a bias towards the first 5 lines of the abstract. These 5 lines were contained in 3.46% of the initial page and received 19.3% of the participants' focus.

The second most viewed part within the initial page was the *Introduction*. Although the introduction is not always entirely located on the first page, it still causes an impact on user attention. On average, the introduction occupied 35% of the initial page. The focus devoted by the participants however, was 26.91% of their triage time on the initial page. The visual processing ratio is therefore 0.52. This is again unsurprising, as we are simply confirming the subjective feedback of the participants. We therefore, report on findings in finer detail which attest to which part of the introduction has the highest visual impact on information seekers.

The first five lines of the introduction, which were always located on the first page, took an average of 3.7% of the space on the first page. The participants spent an average of 7.07% of their time on the first page focusing on these five lines, making the first five lines after the introduction the third most scrutinised feature on the initial page with an importance ratio of 1.91.

The *Main title*, although reported subjectively as the most important feature on a document by our participants, received limited attention. The main title ranked fourth in visual importance on the initial page with a ratio of 0.86. Subjective feedback from the participants agrees in that "the title is often misleading and does tell you what the document may actually be about".

Authors and Keywords. Previous studies, including our previous chapter, did not report on the author section or keywords section with regards to document triage. *Author* space and *Keywords*

also received limited, yet definite visual focus within the context the initial page. Although less important than the main title with a visual processing ratio of 0.76, they are still ranked higher than the introduction.

4.4.3.2 Initial Page Features with Relation to the Entire Document

Within the first page, the average time taken to view the keywords and author section was less than the area of the feature. However, when looking at the features in the scope of the whole document, we can see that there is a increase in terms of attention compared to the average area space. Although 0.91% of the total document space is occupied by these features they receive an average of 2.3% of the participants' attention time. The facet, within the whole document, has more than a modest share of the user's attention with a ratio of 2.53.

	Average Area Occupied(%)	Average Time Viewed		Time/Area (Ratio)	Importance Ranking
Main Title	2.84	5.53s	5.97%	2.10	8
Abstract	2.04	19.17s	20.44%	10.02	2
Keywords/Authors	0.91	2.19s	2.30%	2.53	6
Introduction	5.35	12.09s	12.96%	2.42	7
Headings	4.02	7.60s	7.69%	1.91	9
Emph. Text	1.39	1.61s	1.47%	1.06	11
Images	4.42	1.87s	1.95%	0.44	13
Diagrams/Tables	6.78	3.08s	3.54%	0.52	12
Conclusions	2.29	8.07s	8.42%	3.68	5
References	4.24	1.07s	1.21%	0.29	14
5 Lines after Abstract	0.54	8.56s	9.18%	17	1
5 Lines of Intro	0.70	3.20s	3.44%	4.91	4
5 Lines of Headings	4.15	7.28s	7.64%	1.84	10
5 Lines of Conclusion	0.50	3.96s	4.22%	8.44	3

Table 4.4: Whole Document – Feature viewing times and surface area (see Section 4.4.3.3)

Main Title Subjectively, participants rated the main title as the most important feature with regards to document triage. The importance of the main title within subjective terms is common among information seekers and was evident in the previous study as well as the author's previous work [28, 97]. This claim is at odds with the visual data we extracted. On average, the main title occupies an area of 2.84% of a document while participants spent 5.97% of their viewing time looking at the main title. If we consider the title within the whole document, it is rated as a

visually important feature. If, however, we consider the main title within the domain of the initial page only, then visual attention drops (14.05% area and 12.12% viewing time). Figure 4.1 shows a representative gaze opacity of the main title. There is an evident biased view to the left, indicating attention is highest on the beginning of the title.

Figure 4.1: Gaze Opacity of Main Title (see Section 4.4.1): Attention is given to the main title although it is completely scrutinised and in depth.

Abstract Both the subjective feedback and participant attention within the initial page indicated the importance of the abstract to the document triage process. This view is further supported by considering the abstract within the context of the whole document. Abstracts took 2.04% of the total document space, and took 20.44% of visual attention time. The time given to this fragment of the document is therefore, nearly 10 times its share of the document space. This in turn suggests a very high level of attention on this part of the text.

The first five lines of the abstract also receive the highest scrutiny, ratio wise, in terms of attention time. They occupied 0.54% of the document space, but participants spent 9.18% of their time looking at this specific area. Figure 4.2 shows a representative gaze opacity of the abstract. We note that the first five lines are heavily scrutinised [1]. Due to the eye-tracker limitations we are unable to comment on whether the exact eye-focus is specifically on the right or left. We do however observe two things: a) The attention of the seeker starts on the left. This creates a longer period on the left of the paragraph until the eye-gaze focuses on the initial word of the paragraph and b) The participants' self-reported feedback was that they read the entire line of the test. Although the eye is focused on the centre of the paragraph, peripheral vision can read / locate text surrounding the exact point of attention.

Introduction Section The introduction section receives a high rating of importance from the subjective feedback in the pre-study questionnaire (see Section 4.4.2). This is also reported by previous work by the author and in the previous study [97] (see also Chapter 3). The introduction occupies an area of 5.35% of the document and is in visual focus for 12.96% of the participants' time. Within the initial page, the introduction does not have a large impact on the users' visual attention

---

[1]Due to an inset problem the focus was off by about 1cm to the left
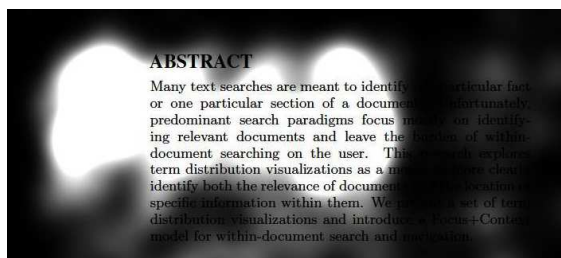
Figure 4.2: Gaze Opacity of Abstract: Very clear in depth focus especially on the initial few lines.

relative to other features like the abstract. It does however, rate higher when compared to features occurring throughout the document such as emphasised text. Figure 4.3 shows a representative gaze opacity of the introduction. Looking at this images, although the heading of the section is brighter than the surroundings and there are small areas of interest, the majority of the main text is not very bright and does not appear to receive an in depth scrutiny. When compared to the abstract we see the visual impact of the introduction is substantially lower.
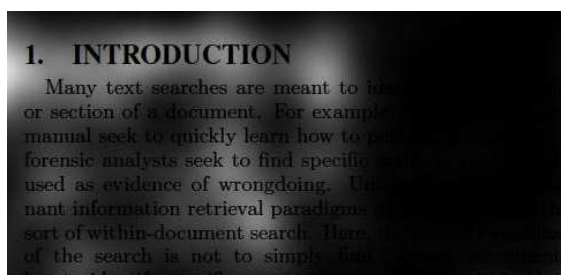
Figure 4.3: Gaze Opacity of Introduction: Receives limited attention, most likely on the first few lines.

Using the Word Count Measurement. Earlier in the introduction to the results section, we explain the reasons for choosing area and time as opposed to taking precise word count measures. We do not wish to report on how much information the seeker extracts, but instead, where the visual attention is; hence using the eye-tracker methodological approach. There are alternative measurements that we could also report on. One such measurement involves the word count per area viewed. This can also be viewed as an important measure in terms of the visual attention of the participants, and provides insight into how much information could have been accumulated by the users. Being aware of the importance of the role of the initial page to the triage process, we consider it an important section for further measures to be taken. In this section, we report on the main findings for the main title, abstract, introduction and initial page results with regards to

the word count and time on the initial page. We therefore can compare our results between two measures and comment on similarities and differences.

Table 4.5 shows the average word count compared to the time taken for the features to be triaged.

	Main Title	Abstract	Abstract(*)	Intro	Intro (*)
AWC	9.67	125.50	49.00	301.00	50.50
ATV(%)	12.12	43.18	19.3	26.91	7.07
ATV/AWC	1.25	0.34	0.39	0.09	0.14

Table 4.5: Initial Page Features: Average Percentage Time to View versus Average Word Count (*** denotes the first 5 lines after the heading of the feature described.**) All values are rounded to 2 d.p

Our word count results correlate well our area/time results with one exception. The abstract and particularly the first few lines after the abstract, receive a larger share of attention per word than the introduction. Similarly, the 5 lines after the introduction and the abstract, receive a larger share of the attention than the entirety of the introduction and abstract respectively. The starkest difference is the main title. Clearly, the title will be suffering artificially from the area measure because of its larger font size. In terms of word count and time ratio, it is receiving about four times the attention the abstract is receiving as a whole. We can understand that the words of the main title, have some importance but, nonetheless, they do not seem to have the same cumulative impact as the abstract.

The main title is the largest feature font-wise and appears first on the initial page, thus attracting attention. It would probably be taken as the psychological subject of the propositional representation, in our case, the document. This is a phenomenon known as featural salience [106]. If this was the only document feature, users would not have to learn about the predicate features such as the introduction and the abstract. Further study to distinguish between visual attention and cognitive understanding and interpretation would complement the findings.

4.4.3.3 Further Features Within the Whole Document

The initial page, although scrutinised heavily, received just under half of the participants' visual attention. We now look at document features on subsequent pages. The findings are summarised in Table 4.4. A χ^2 test reveals a significance of p = 5.59 x 10^{-18} (χ^2 = 210.15, df = 11) of document

features proportionally between the area occupied and the time viewed. These do *not* include the items relating to the first 5 lines.

Headings Headings are another feature considered by participants as important to the triage process. Headings in this context do not include the headings in the features included in the previous subsection; such as the introduction. Furthermore, data was captured only from headings and subheadings.They are given a subjective rating of 6.85/10 (see Section 4.4.2). Headings occupied an average area of 4.02% of the document space, but participants gave them a greater share of attention with 7.69% of their time. The first five lines after a heading occupied 4.15% of the document and viewed 7.64% of the time. Figure 4.4 shows a representative gaze opacity of a typical heading. We can see a brighter view of the first few lines in contrast to the surrounding area.
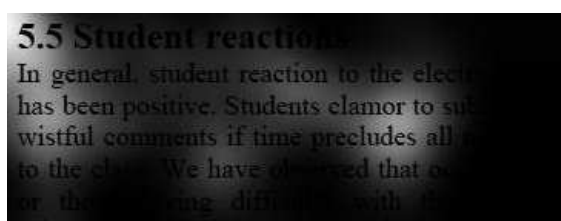
Figure 4.4: Gaze Opacity of a Heading (see Section 4.4.1): A good impact on user attention. If further scrutiny occurs, it occurs on the first few lines.

Emphasized Text Emphasised text showed very little impact on participant attention, when compared to the average occupied area. It is interesting to note that emphasised text was considered quite high by the participants in terms of triage on their subjective scores(6.05/10) see Section 4.4.2. Our previous study has reported emphasised text present in an information seeker's triage process. We can now verify its relatively small impact on the visual attention and focus of users.

Images Images[2] were rated of average importance in document triage (see Section 4.4.2). Images receive little attention from the information seeker. They had an area that occupied 4.42% of the document space, yet received only 1.95% of the participants' average attention time.

Diagrams and Tables Like images, diagrams and tables, although rated with average importance in participants' subjective ratings, seem to have a minimal impact on user attention. Again, we see the same effect when comparing average time viewed to average area occupied as that of

---

[2]Images are pictures and drawings but does *not* include tables and diagrams which require further cognitive load for understanding.

images. When present, an average area of 6.78% was occupied by diagrams and tables, while only almost half that time (3.54%) was spent triaging them. Figure 4.5 shows an example gaze opacity of figure attention within a document. The large dark areas and lack of bright areas give us an indication of how users do not scrutinise images heavily.

Figure 4.5: Gaze Opacity of Diagrams/Tables/Images: Although some user focus is shown, there is very little scrutiny.

Conclusions Section 'Conclusions' were rated the third highest feature, in terms of subjective ratings (see Section 4.4.2). Participants rated conclusions higher than the introduction. The conclusions section occupied an average of 2.29% of the document surface. They spent an average of 8.42% of their time triaging the conclusion a ratio of almost 4. By further dissecting the conclusion section and investigating its individual sections we can report on the most important visual areas.

By focusing on the first five lines of the conclusion, which occupied on average an area of 0.5% of the document we can see the impact these lines had to the information seeker. Participant attention time averaged of 4.22% of their total time spent on this section alone. Figure 4.6 shows a representative gaze opacity of the conclusion.
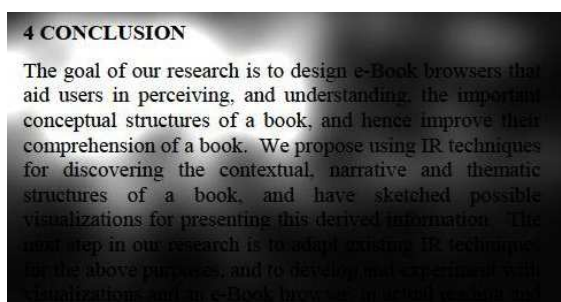
Figure 4.6: Gaze Opacity of Conclusions Section: Clearer than the introduction, the conclusion receives a good amount of attention, especially on the first few lines.

References To date, the visual effect of references during document triage has not been re-

ported on. References, had the lowest subjective ratings (See Section 4.4.2). They covered an average area of 4.24% of the documents while attention time was much lower at 1.21%.

4.4.3.4 Summary

In this sub-section we have seen how information seekers focus their attention on the initial page of the document. Not all the features contained within the initial page are heavily scrutinized (see Section 8.5.2). After the initial page, users look at certain features like headings and the conclusion (see Section 4.4.3.3). Although participants focus attention on plain text it is predominantly on specific areas such as the few lines after headings.

4.4.4 Participants' Relevance Ratings

The participants were asked to give a relevance rating for each document based on their information need. Table 6.2 shows the relevance judgement ratings by the participants during their document triage. The ratings were one of vr (very relevant), sr (mostly relevant), si (mostly irrelevant) and vi (very irrelevant).

Document	1	2	3	4	5	6
Very Relevant	12	2	2	1	7	11
Mostly Relevant	8	12	4	4	5	7
Mostly Irrelevant	0	6	8	5	5	1
Very Irrelevant	0	0	6	10	3	1

Table 4.6: Participant Document Relevance Ratings.

The primary reason for recording participants' document ratings was not to compare how well they were able to rank the relevance of the documents. Our previous study shows a bias on features of a document, such as length, which can play a deciding factor in relevance decisions during document triage [97] (also see Chapter 3). It is also reported from participants' qualitative feedback that more specific document features, such as clear abstracts, clear titles and well organised structures, can play a more dominant role in influencing the decisions of an information seeker during document triage. Cool et al reports document format as an affecting factor in relevance decisions [39]. In order to extract some further data and justify future scrutiny, we selected two documents in each relevance category. Our selection was based on the clarity of the main title

abstract and headings of the documents to the information need as verified by our experts.

In the case of the two *relevant* documents (Document 1 and Document 2) we can see a favourable vote towards Document 1. A χ^2 test reveals a statistical significance of p = 0.000938 (χ^2 = 13.94, df = 3) between the ratings for these two documents. Document 1 was specifically chosen for its clear title, and abstract, as well as its high relevance to the information need. Conversely, document 2, although highly relevant to the information need, did not have a clear title and abstract.

The majority of participants gave document 6 higher ratings than document 4. A χ^2 test reveals a statistical significance of p = 0.000251 (χ^2 = 52.31, df = 3) between the ratings of the two documents. Although both the documents were ranked with average relevance to the information need, the shorter document receives favourable scores. We cannot claim that this is the only reason for the score ratings, we can justify further study.

Comparing the two documents (3 and 5) in the *not-relevant* category, we again see participants being more willing to rate document 5, which contains relevant headings, higher than document 3. A statistical difference in a χ^2 test however, does not show a statistical significance between the two document ratings (p = 0.2051 χ^2 = 8.08, df = 3).

Our rating findings, produce mixed results. We are not able to conclusively derive any principles from the data we have recorded, but we do have indications on which areas, such as clear headings and length, future studies can benefit from.

4.5   Discussion

The results in Section 4.4 are now considered by supplementing the quantitative data with qualitative data from the post-study interview and *think aloud* comments made by the participants during the main triage study. We begin by looking at the initial page and proceed to individual features. Finally, we comment on the remaining elements which do not attract much of an information seeker's focus. These comments are summarised by visually presenting the main observations in Figure 4.7. Figure 4.7 is an extract of actual eye-tracking data, provided in a gaze opacity (see section 4.4.3), which represents an information seeker's attention during a typical triage of a document.

Initial Page It was expected that the initial page will have a high impact on visual attention
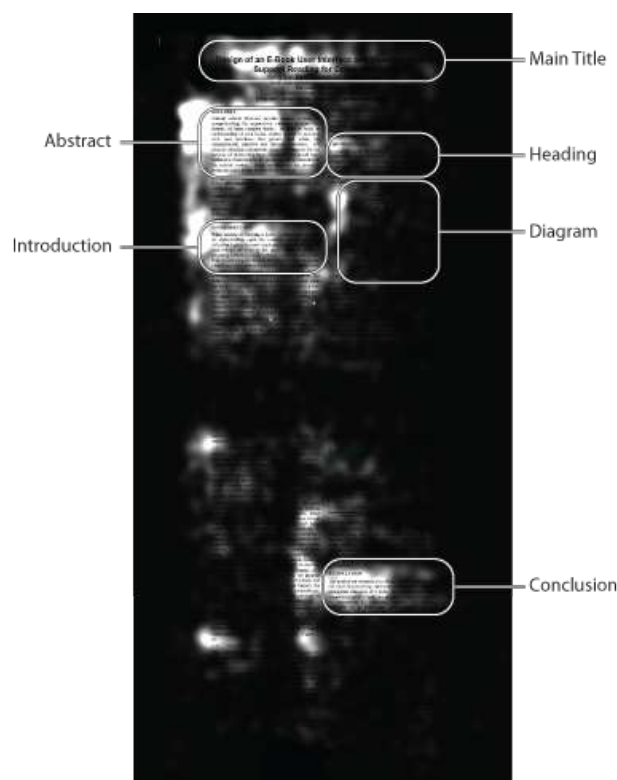
Figure 4.7: Gaze Opacity of a Typical Document During Document Triage (See Sec. 4.5)

[28, 97] (also see Chapter 3). Participants were, retrospectively, asked to describe how close to making a decision they were upon completing scrutiny of the initial page. All the participants reported either making up their minds while on the initial page or being very close to a decision. As 50% of our participants reported: further reading to the initial page was done for "further reassurance". As one participant put it, "all the important things are on the first page, so I quickly sort of know what I think of this document". A representative gaze opacity of the initial page can be seen on the top half of Figure 4.7.

Abstract. The abstract contains what participants described as the "core of the paper" or the "summary". One participant, that directly started triaging documents by reading the abstracts, commented that "if the abstract is right (representative of the document), then I don't need to look at the rest (of the document)". Although the whole of the abstract attracts a user's visual attention, the first few lines under the heading are heavily focused on by information seekers (see Section 4.4.3). When questioned after the main study, 85% of the participants agreed that the abstract alone was the most influential part of their decision making process. A representative *gaze*

opacity of the abstract can be seen on part B of Figure 4.7.

Main Title When an information seeker first looks at a document, the first feature they are most likely to focus on, according to our participants' first fixations, is the main title. The participants reported that titles may imply what the paper is about, but can often be obscure and, sometimes, even misleading. For example, one stated: "unless the heading is amazingly accurate to my need or has absolutely nothing to do with the topic then" they would read further. In Figure 4.7 we can see that the title receives a substantial amount of attention for a relatively small volume of text (albeit in a large typeface). Per word, the main title receives a large amount of attention, but overall it does not seem that it is simply the dominant factor in the user's decision. Clearly however, it plays a very important role.

Introduction We are again, able to report on the majority of the visual attention received by the first few lines of the introduction. Though it receives more total attention time than the conclusion, this relationship is reversed when the size of each part is considered (see Section 4.4.3). In the post-study interview, one of the participants suggested that the introduction is expected to "say what the paper is about, but it seems that people like big introductions rather than getting to the point". In Figure 4.7 we can see an example of the cumulative attention the introduction receives. Participants' main feedback on these lines, is that they are generic in nature and do not provide the seeker with the specific information needed to formulate their relevance decision.

Conclusions Section As reported in the previous study and in the previous sections in this chapter, the conclusion section plays a larger role, visually, than the introduction to the document triage process. Our study in the previous chapter suggested that although overall attention of users dropped on subsequent pages after the initial page, there was a 'spike' of interest on the page containing the conclusion [97]. We can now prove that the information seekers' gaze focuses upon the conclusion section. Once again, the first few lines receive substantially more attention that the remaining conclusion. Participants, reported that the conclusion acted as a confirmation of their decision rather than new information. One participant commented: "actually I'll read the conclusion first. Should do that before giving the final decision".

Plain Text and Main Document Body. Beyond the initial page users demonstrate inconsistent attention to different headings. No individual user has give equal attention to all headings

and many headings have received little or even no visual focus. Furthermore, these patterns are not consistent between users. This suggests that the selection of which headings are looked at is quasi random and might be crudely characterised as a 'pecking approach'. Like a bird pecking for seeds there is some intention but there is also arbitrarily some randomness as to the seeds that are selected. So is the randomness of the headings selected by the information seekers while performing document triage. Beyond the headings, the plain text as a whole does not attract the focus of the information seeker.

Images, Tables and Diagrams These are considered by information seekers to play an important role in their document triage process, according to their subjective ratings (see Section 4.4.2). However, looking at the focus of the users, we can determine that this is not the case in general. Although some points of focus were registered by the eye-tracker, participants did not examine any of these features in depth (as seen by the gaze opacity in Figure 4.7 part E and described in Section 4.4.3). When questioned about this discrepancy, participants commented on the lines of "although a picture can show something specific it's hard to get the gist of the whole document". Furthermore, diagrams and tables may require more cognitive processing and time than an information seeker is willing to give during the fast paced process of document triage. Conversely, an image can require only a small amount of time in order to convey the required message to the information seeker.

The remaining features (plain text, references and emphasised text) receive no particular attention during document triage, as is demonstrated by the eye gaze plot in Figure 4.7 and the viewing data in Table 4.4. Overall presentation - a factor not readily quantified, but at least associated with regular use of headings and a balance use of features across the document - seems to impact on users' subjective relevance ratings. Specifically, more 'symmetrical' and 'regular' documents have higher ratings, and those devoid of common cues such as titles and headings are negatively assessed (See Section 4.4.4).

4.6 Conclusions and Summary

In this chapter we described an in-depth study of information seekers performing document triage. Using an eye-tracker, we traced users' visual attention with a new level of precision.

As expected, the initial page receives a large amount of the seeker's attention. We looked at the first page from two perspectives; namely, word count and area occupied. We report on the relative strength of the abstract and introduction, and even more specifically, the few lines after their headings. We also report on how the main title has, in word count terms, a high statistical number but not in terms of area. This makes areas with big fonts looking better and areas but also elevates the importance of lots of text in a small area being viewed for a large amount of time. Overall, the main title, abstract and first few line of the introduction receives the majority of the visual attention within the initial page. The remainder of the document is skimmed in a lightweight fashion with participants focusing mainly on the headings, conclusions and the first few lines after a 'relevant' heading. When a potentially relevant area is identified, further localised reading takes place.

There were two main aims to this study. The first was to verify the findings of the previous study using precise eye-gaze data. This study verified the main findings of the previous one. The relative importance of major features such as the initial page and conclusions were confirmed. Features which were reported to have a high visual impact in the previous study, such as the initial page, headings and conclusion section were focused upon by our participants. Conversely, pages with no specific features were skimmed without any reading behaviour.

Our second aim was to analyse the detailed visual focus of users. Two major new discoveries were the detailed relative importance of features within the first page, and the very limited reading following a heading on any page - typically no more than five lines. We can now begin to build on previous work, to tie together our findings and contribute both to our triage model and to the general information seeking research work.

We have previously reported on the local selectivity rule [144] in which "users continue to have low tolerance for wading through large retrievals". In the papers' study, participants were not likely to view more than two results pages per results list. Our participants replicate this "low and declining level of user interactivity". Once the initial page is scrutinised, the information seeker will either move on to the next document (equivalent to moving on to another search term) or continue to triage but with reduced visual attention. We are able to also report that there is a higher probability for users performing triage to go through all the document pages, than for users

to go beyond the initial two pages of a results list.

By continuing to triage a document beyond the initial page, a user is likely to skim without reading until he reaches areas of interest where he will focus his attention for further scrutiny. These main areas of interest are likely to be a heading and, the few lines after it. This 'pecking behaviour', as we have previously characterised it, represents user behaviour as moving between skimming the document and focusing on some of the further document features and plain text surrounding them.

By identifying these behaviours, we begin to contribute towards building our document triage model (see Chapter 8) grounded on empirical statistical evidence. We have, through our first two studies, identified areas which can add value with further research, such as the effect of document length on relevance decisions. Although these gaps need to be investigated further at a more detailed level, we are focusing on the within-document triage behaviour as a whole in our work. In the subsequent chapters, more constraints and interfaces are examined to begin the formulation of a general model for the triage process.

Chapter 5

Small Screen Triage

5.1   Introduction

Our research aims to produce a basic document triage model to conceptualise the visual behaviour of information seekers. This can assist us in informing the design of assistive software for information seeking. Though there are several variables that we could investigate and change (such as topic, document familiarity and structure to name a few), we are investigating the basic document triage activity; namely, how a user makes decisions on realistic unmodified content and so do not alter the documents in any way. One of the main variables which has not yet been reported on, is that of small screen effects on the document triage process. Investigating small screen influence is also timely as these devices are becoming increasingly main stream.

In the previous chapter, we reported on the exact visual attention of users during their triage activities on a desktop setup. Research using other screen displays and its effect on document triage has already been reported on by Bae et al [13]. The main focus of the study was to increase the screen real estate in order to reduce window transitions. The results showed notable differences in the behaviour of the participants between different screen sizes. We already know that web browsing and information seeking varies between desktop screens and small screen displays [77]. Furthermore, we know that display size effects other activities, including critically, reading [125], something dominant in the triage process. Interestingly, in their research Richardson et al, state that some participants quoted that "visual scanning of text was probably easier with the small

window as it required less ocular movement and it was easier (quicker) to come to a decision as to whether a given page contained the target information." We hypothesise that there will be notable differences in smaller screen size displays for academic document triage also and therefore deem it as an essential factor to investigate within our work.

We no longer associate small screen devices solely with mobile phones and only as a means of communication. An area which is rapidly gaining popularity, with regards to small screen devices, is that of portable document readers. Amazon's *Kindle* [1] and *Sony Reader* [2] are examples of specialised small screen document readers which are commercially popular. The evolution of mobile phones and PDAs (Personal Digital Assistants) have driven the creation of bespoke reader tools for mobile devices such as the *Aldiko* [3] for Google's Android OS, a built-in PDF reader for the I-Phone and *Mobipocket reader* [4] for Palm OS. Furthermore, Adobe, a key stakeholder in the electronic document field, have adapted their popular Acrobat reader software for use on small screen devices.

As these small screen readers continue to become popular, users are faced with challenges in information interaction for document triage; namely, how to use the available space to make sense of the relevance of documents to their information need. In this chapter, an exploration of the behaviours and requirements of information seekers when performing document triage on small screen displays are investigated. There are two main objectives to the study presented. These are:

1. To verify which of the previous behaviours are replicated on small screen triage. Small screen can have the effect that it is "easy to forget which section they were in or when one section had ended and a new one begun" [125]. Although this is an important factor when reading documents, we are not sure if our users will make a distinction between different sections when it comes to locating information simply for making a relevance judgement on the document. We know thus far, that the introduction, the abstract and the conclusions sections are the main visual attractions for triage. Whether this visual stereotype will be altered in a small screen triage environment is not yet known.

---

[1]http://www.amazon.co.uk/ (2010)
[2]http://www.sony.co.uk/hub/reader-ebook (2010)
[3]http://www.aldiko.com/ (2010)
[4]http://www.mobipocket.com/en/HomePage/default.asp?Language=EN (2010)

2. To identify variations to document triage behaviour when using a small screen real estate. We are aware of how reading on small screen displays creates distinct behaviours for reading. One example is that of "the need for readers to page through twice as many small screens to access the same amount of information"[125]. Since document triage is a form of reading (skimming), we are interested in the variations that will take place when the users search for visual clues to relevant information to their satisficing criteria.

The study is designed to be in-line with our previous investigations. Consistency in the experimental design is key to the reliability of the reported data. In order to be able to compare the results, and also to report on the effects of a small screen variable, we alter as few remaining variables as possible. Structured documents are used as in the previous studies. Participants are asked to triage a set of documents according to an information need. Their actions are recorded using specialised mobile testing hardware, such as an overhead camera recording and streaming the user's actions in real time on a lightly coloured background. Empirical data for a quantitative analysis as well as qualitative feedback from questionnaires and interviews are then extracted.

5.2 User Study Design

Pre-study. 10 participants (6 male, 4 female) were recruited for the study. All participants were required to be of post-graduate level or full academic staff, and experienced in reading academic publications. Due to the nature of the study, participants were not required to be experts in a specialised field, but were required to have basic fundamental knowledge of Information Technology. Participant were handed a questionnaire containing general questions such as age, experience in using small screen devices and document readers, documents triaged on average per day and time spent on the process per day. Further questions asked the participants to rate document elements with regards to their own subjective importance to the usefulness in the triage process.

Main Study. Participants were given a single information need and asking them to triage six documents, randomly presented, on a small screen. The short study time and task was enforced to not inconvenience the participants due to the study layout. They were asked to rate the documents between 0 and 10 (10 being extremely relevant and 0 being not relevant) of how relevant each

document was to their given information need. In order to achieve this task the participants were given a Dell Axim X51 palmtop with Adobe Acrobat Reader installed. The documents were stored on a local folder in PDF (Portable Document Format) form and were accessible by pressing a single on screen button. The reason for choosing Adobe Acrobat for the study lies in user familiarity with the software. Participants were given an open-ended time frame in which to complete this task. Users' actions and screen content were recorded in real time by using a specialised overhead recorder streaming the data to BB Flashback Express Screen Recorder [5], which records the videos (See Figure 5.1). Think-aloud comments were recorded using a separate microphone and synchronized on the video post-study.

Figure 5.1: Participant Performing Document Triage (Researcher's view on the top left).

Post Study. Following the completion of the main study a second questionnaire was given. Our participants were again asked to rate the individual document features subjectively, based on their importance to the triage process. By comparing the ratings to the ones given before the main study, we tested for a change in perceived importance of the document features. Furthermore, the use of a small screen may alter the initial feature relevance to the participants in this specific situation. This test is merely exploratory and can can give us reason to investigating further. A semi-structured

---

[5]http://www.bbsoftware.co.uk/bbflashback.aspx (2010)

interview followed, in which the participants were asked to comment on the difficulties in finding relevant information, preliminary and final decisions on documents and to explain their behaviour with the aid of a video playback. Participants were given £10 as compensation for their time.

5.3 Results and Discussion

In our previous two studies, we reported on the behaviour and visual attention of users performing document triage on a desktop computer. The general patterns we identified in the two studies were consistent with each other. We now have acquired further data to compare information seekers triage behaviour between small screen devices and desktop PC's. In this section the results from the qualitative and quantitative findings extracted from the study are presented. We focus on the findings which are directly relevant to our research and can be compared to other chapters. Although major findings are highlighted here, a direct comparison to our results from other studies is found in Chapter 9. Firstly, we report on the subjective importance of document elements to the information seekers, before and after the main study. The navigational patterns and areas which participants focused their attention on are then reported on, followed by the document ratings. Barriers to interactive small screen document triage are also identified by qualitative feedback.

5.3.1 Subjective Ratings

The participants were asked to rate document elements as to their perceived importance to the document triage process on small screen devices. A summary of the results can be seen in Figure 5.2.

The highest average ranked element pre-study, was the *Main Title* closely followed by the *Abstract*. The lowest ranking elements included *Plain Text* and *References*. Upon completion of the study, these four elements, unsurprisingly, retained their rating order. Overall we see a sharp fall in average ratings between pre and post study interview ratings. A Wilcoxon's test for non-parametric data sets on the overall ratings between the two sets finds a significance of P<0.0001 (n=90). Upon focusing on the elements individually, this hypothesis stands true, again using a Wilcoxon's test, for the *Main Title* (P<0.0273, n=10), the *Introduction* (P<0.039, n=10) and the *Conclusions Section*. The other elements do not produce significance using a Wilcoxon's test to

	Main Title	Abstract	Headings	Introduction	Plain Text	Conclusions	References	Figures	Emphasized Text
■Pre	8.9	8.6	7	7.2	3.8	7	3.7	6.4	6
■Post	7.5	7	6.8	4.5	3.2	5.4	1.5	5.8	5.2
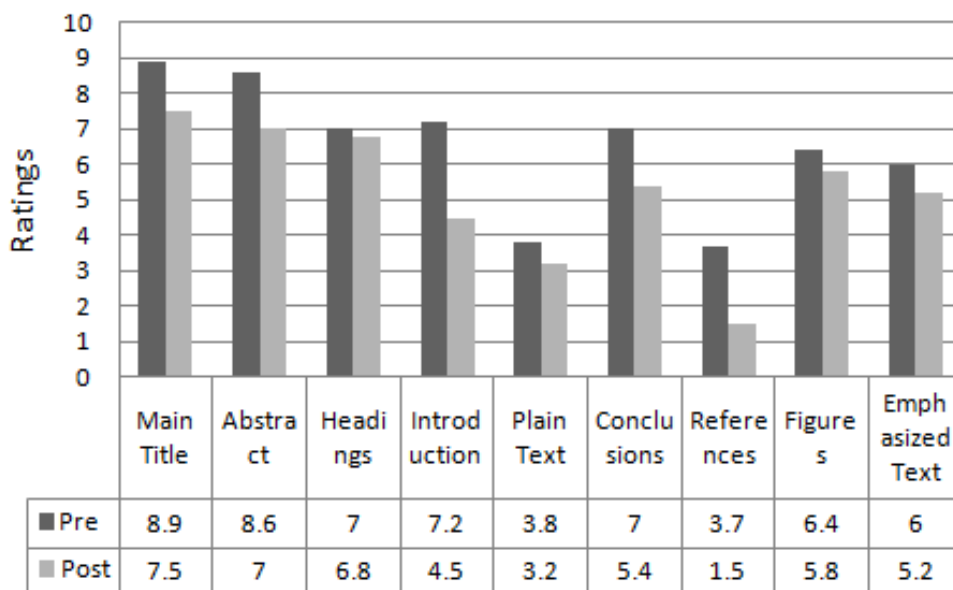
Figure 5.2: Pre and Post Study Ave. Subjective Ratings of Features

the 5% level. The overall understanding from this data and participants' feedback is that the general confidence levels drop in that the participants do not justify one specific feature as able to completely cover the information need.

5.3.2 Navigational Patterns & Areas of Interest

Navigational Patterns. When information seekers triage documents, they follow a set of scrolling patterns [97] (See Chapter 3). There are four main navigational patterns, which can be combined, to produce hybrids to model navigational behaviour. When performing document triage on small screen devices, participants replicated the patterns. 60 instances of document triage were performed (1 per document scan). Of these, 49 were performed using *Step* navigation (all the participants were reported using this behaviour at least once), 9 using *Flatline* navigation and 2 using *Begin and End* navigation. On 5 of the instances the participants returned to the beginning of the document before ending their triage. When comparing these results to our previous studies the predominant of the *step* navigation pattern [97] (Chapter 3) is recognised.

Areas of Interest. Not all parts of a document have the same value to the document triage process. The hypothesis is that this statement will hold for document triage using small screen devices. Figure 5.2 shows a comparison between the area occupied by document features and the

actual attention time given to them. Our two previous studies in Chapters 3 and 4 identify that the *initial page* plays a vital role in the visual document triage process. Data in both studies reported close to 50% of the participants viewing time devoted to this area. When performing triage on small screens, users rely on the initial page even more, scrutinising it for an average of 65% of the total viewing time. A t-test of these differences was significant at the 5% level (p=0.000193). **Plain text**, which does not include any special characteristic, accounts for 71.78% of the area on the documents. However, the attention received in this study was 46.67%. The attention time mainly includes scrolling, which makes any sense making textual extraction impossible. The same findings are replicated on larger screen display triage, were limited meaningful attention is given to plain text [97]. The most attention, when comparing area occupied to attention time, focuses on the **abstract**. Although the abstract covers an average area of 1.52%, 21.82% of the time is spent triaging the document though the abstract, giving us a ratio of 14.36. The abstract is also the most visually processed feature in our previous eye tracking findings, making these results consistent between the two screen real estates. The participants scrutinised the **Main Title** over 7 times its corresponding area percentage. The **introduction**, which rates high on pre study subjective ratings, gives a ratio of 3.4 to its corresponding area. Overall, the area of the introduction which is located on the initial page receives the majority of the attention, while the introduction part on the subsequent page is not scrutinised. All the participants performed this behaviour on 95% of the triage instances. The main title, abstract and introduction are often located on the initial page of a structured document. The document is presented to the information seeker with the initial page displayed first on the screen. This is therefore a factor that may affect the bias towards the initial page viewing times.

Headings occupy 2.9% of the document area. Participants spent more than twice that time however, triaging with the headings in focus with 6.21% of their time. Headings were rated the fourth most important element in the triage process before undergoing the study and third most important after the study, thus increasing its subjective importance. Beyond the initial page, excluding the headings, only the **conclusions section** receives a favourable ratio regarding the time spent by the participants and the area occupied is. With 0.95% total area and 4.24% of the users' attention time the conclusion has a positive 4.46 ratio. Elements that are not looked at

	Main Title	Abstract	Intro	fig	head	emph	conc	ref
■ Attention	8.78	21.82	7.31	10.6	6.21	0.59	4.24	0
■ Area	1.23	1.52	2.15	13.46	2.9	0.64	0.95	5.38
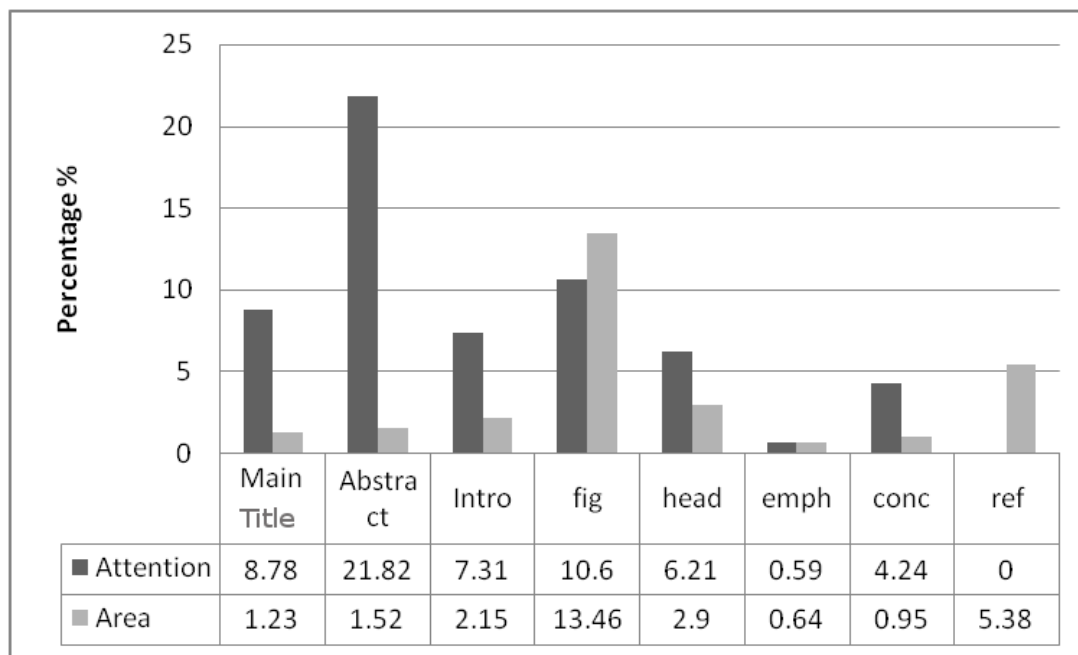
Figure 5.3: Ave. Attention Distribution (%) Compared to Ave. Element Group Area (%).

considerably, other than the plain text (discussed above) are **figures** (ratio: 0.79), **emphasized text** (ratio: 0.92) and *references* (ratio: 0.01).

Generally, the visual attention of users is consistent between the mediums. The initial page is, as expected, the most important area of the document and the most viewed in all cases. Interestingly, our previous study (Chapter 4) reported the main title as the 5th most important feature for triage with an visual processing ratio of 2.10. Participants in our small screen triage experiment, produced an importance ratio of 7.34 ranking the main title as the 2nd most visually important feature in the document. We hypothesise, that due to the physical constraints of a small screen display, information seekers rely more on the initial page which is easily accessible in order to make their relevance decisions.

5.3.3   Document Ratings

Participants gave average ratings to all documents of 6.3/10 or above with the exception of one document which received an average rating of 5/10 (S.D. = 2.8). When questioned, 9 of our participants reported that they found it difficult to look for information on a small screen on "such

a relatively large document" and had to "rely on the abstract and title to cover all the document contents". The document was in fact the journal article, reported in our previous studies [97] (Chapter 3. In our previous work we have reported on the effects of length to the decision making process of individuals; namely that document length can play a negative effect on one's rating. This effect was later tested in our eye-tracking experiment with regards to the effect of having a large number of features in a document. It was reported by the participants' post study interviews that having clear document features can help them to overcome the bias effects to lengthy documents. Our previous studies also concur with the ratings given here. Information seekers, in our initial study gave an average rating of just under 6/10, while participants in our eye-tracking study gave an average rating of 5.4/10. A Mann Whitney test reveals no statistical significance between the three groups in terms of relevance ratings. We can report therefore, there is no drop on relevance decision efficiency, based on the screen size of the medium used.

5.3.4   Barriers to Small Screen Document Interaction

Following the main study, an interview took place where participants were questioned using a semi-structured interview to identify further qualitative details regarding their behaviour, thoughts and feelings during their triage activity. A question asked was when during the process did participants form initial relevance ratings for the document? Upon answering the question they were then asked at which point was that decision finalised. This two-part question reduced any potential priming bias of whether two decisions were made. Participants commented that the decision was made after the main title and abstract were scrutinised and their decision remained the same until the end. All the participants who did not follow a flatline pattern (see Section 5.3.2) commented that they simply verifying their initial relevance decision by scrutinising the document further.

Participants were then asked to comment on barriers to interaction; namely, what kept or delayed the participants from making their triage decisions. The most common factor, reported by all the participants, was that of scrolling horizontally (left and right). All the participants' adjusted their zoom level out, to the point where the most text could legibly fit on the screen. This created problems when reading, even when the document was double column formatted. As one of the participants commented: "scrolling left and right on the title may be OK but having to

do it for every line is ridiculous and frustrating". These comments are in line with earlier design principles which state that systems need to "reduce the amount of vertical scrolling by simplifying the text you wish to display" [27]. The research reported this heuristic with regards to mobile website design. We can now positively verify the validity of this rule with structured documents and triage.

A further barrier that surfaced from the majority of participants' was that of locating document elements such as the conclusion section. As participant 3 commented, it "is a problem to easily locate relevant document parts". Interestingly, none of the participants searched for the conclusion using the built in search feature, although they were informed that it existed and was available to them, therefore portraying similar behaviour to that mentioned in [96]. Only 6 of the 10 participants reached the conclusions section and only 5 remained for more than 5 seconds (thus being able to read the text in depth rather than skimming).

5.4   Summary and Conclusions

In this chapter a study reporting on document triage using small screen displays with structured documents was presented. By repeating the experiment from our previous studies and altering as few variables as possible, we are able to validate already existing behaviours with a small screen real estate. The majority of the behavioural patterns recorded, encouragingly coincide with our previous findings.

One of the things we did differently in this study, was to ask the participants to rate features before and after their triage activity. The results showed a fall in the ratings generally after the triage occurred. Although we were not able to analyse the data before asking the participants to comment on why this was the case, the general consensus from the participants' post-study comments was that no one feature is concise enough to allow an information seeker to make a relevance decision. We have reported on the models of navigational patterns on small screen devices. The majority of small screen triage instances follow a step exploration pattern [97] (also see Chapter 3). The inclination of the information seekers to follow a step navigational pattern is recorded to be much stronger on the small screen than on a desktop activity. Furthermore, we report the elevated scrutiny of the initial page, from a 50% of the time average on a desktop

to 65% on a small screen device. Our results verify the perceived visual processing dominance of the abstract as this area receives the most attention from the participants with regards to the time and area occupied. Beyond the initial page features, we can report on the headings and the conclusion section as having the only notable impact on visual processing. Again, these features receive elevated attention when triaged through a small screen display rather than a desktop screen. Interaction barriers which were valid for websites on a small screen device are also recorded to have the same effect on semi-structured documents during triage. Two examples are limited use of the search functionality and the negative attitude to horizontal scrolling for longer blocks of text.

These results give us vital information on the behaviours which agree with a larger screen real estate and ones which vary on a small screen device. This covers the two main environments that users are likely to perform triage in. We therefore now have a foundation to begin formulating our model, based on the behaviours of the information seekers, such as their navigations, and their visual attention. The process was kept as close to a representative triage activity in that our documents were not altered in any way. In our studies we report on parts of a document that seem to be more visually attractive to our users. We have thus far distinguished features that are more visually scrutinised than others, such as the initial page, conclusions sections, headings and the first few lines under relevant sections. In a realistic triage activity, satisficing is reported to be the tendency to select the first option that meets a given need or select the option that seems to address most needs rather than the optimal solution (see Chapter 2). We wish to negate any effects we have when users take part in our lab-based studies of trying to find the optimal solution and to confirm which parts are visually scrutinised due to the user's perceived importance to the task at hand. Furthermore, limited time can also often be a factor to an information seeker when performing document triage. In the next chapter, we employ time constraints to participants to test for this.

Chapter 6

Applying Time Constraints to Document Triage

6.1 Introduction and Related Work

Our previous chapters report on the direct visual impact of document features to information seekers. This corresponded to and complemented earlier work which began looking at how document features influence relevance decisions [129]. Relevance decisions were reported, by our questionnaires (in Chapters 4 and 5) and qualitative feedback from the previous studies, to be taken in the close to the beginning of each document's triage; generally before the seeker has left the initial page. The users reported continuing the seeking process generally to confirm their initial decision rather than to formulate it. In limiting the time given to the participants and comparing the relevance ratings, we can report whether or not this is true. If the document relevance scores are similar between the groups that perform triage with no time limitation and those that do, the decisions are likely to be those taken at the beginning of the triage process.

The literature identifies "limited support for the idea that, when there was insufficient time to read all the text available, readers could effectively focus on the most important parts of the text" [47, 105]. This chapter will be looking at the possible impact of time constraints to information seekers performing document triage. We investigate visual attention relating to the document features, when the user needs to sacrifice non-important parts of the document due to time. We

are interested in giving the participants a time constraint in order to observe what the variation to their visual behaviour and ratings are when they knowingly have a time variable to address.

Most document triage activities are, by definition, a fast-paced part of the information seeking process [28]. Information seekers will triage documents by skimming their content for relevant information. In-depth reading is minimal and occurs only after a document is perceived to be relevant to the user's information need. During this fast scanning activity, users have a low tolerance for searching through information [144]. We also uncovered evidence which testify to information seekers being biased against long documents [28, 97]. We hypothesise that all these factors contribute to users making incorrect judgements on document relevance.

There is evidence to suggest that skimming (such as the behaviour during document triage) is more beneficial in terms of information assimilation than reading [48]. We are aware that users glance at a page before settling to read the pieces of potential interest [48], a behaviour that is based on visual attention. The document triage process, as we have seen in all our previous studies, is a fast-paced skimming exercise aimed at identifying key pieces of text which are used to formulate the document relevance to an information need. The importance of the skimming and visual attention therefore is heightened when there is a limited time frame and skimming to identify potentially relevant features needs to be undertaken on a larger scale. What guides the user the most is the visual cues, making our work of paramount importance in terms of modelling behaviours.

6.2   Study Description

The methodology and study design follows closely with our previous studies, in order for us to accurately compare the data and enrich previous findings. We required the use of an eye-tracker to record the exact visual gaze of the participants during their limited screen viewing time. Think aloud comments and interruptions during any of the tasks would produce false information from the participants due to the narrow time margin available. We therefore required no interaction from the participants with the investigator during their triage activity. We begin by explaining our participant selection criteria and give a description of the study design, highlighting the differences from our previous methods.

Participant Selection.   20 participants were chosen, 7 female and 13 male, for the study.

The first 2 participants were chosen as pilot study subjects, while the remaining 18 participated in the main experiment. All participants were at postgraduate level or above in an HCI (Human Computer Interaction) related discipline. Their ages ranged from 21 to 50 years of age. In order to explore different group timings, we chose, for practical reason, to include 6 participants per group. A strong statistical difference would surface and provide us with more evidence to look at specific situations in more detail, which was not the case here.

Study Setup. Our general study design varies very little from our previous methodologies. This is deliberate in order for the integrity of the data when it is compared to our previous findings. When comparing the data to our previous findings, we report on a limitation. In the study presented in chapter 3, which used the same documents, no eye-tracker was available to us. Therefore, only the participants' relevance scores can be compared directly. We can compare the visual area between the too studies, but our certainty as to the exact visual focus of the previous study is not accurately available. Our participants were initially briefed on the definition of the term *document triage*. Care was taken however, in not giving out information on how users are expected to act during document triage that may bias the behaviour of the participants. A pre-study questionnaire which mirrored the questions found in our previous questionnaires was given to the participants.

In the main study, we gave our participants two information needs, reflective of the ones in Chapter 3. This way, we are able to compare the findings with all the previous studies (since the other studies contained either all or part of the documents in Chapter 3). The same documents were also presented to the participants. In this study, instead of an open time-frame however, information seekers were given a limited time frame to make their relevance decisions and perform document triage on each document. We are interested in specific document features rather than surrogates or to alter the satisficing rules by allowing users to not look at documents. For this reason we limit the time on each document rather than the overall time. In order to vary the times, 6 participants were chosen to triage the documents using each time constraint condition (30, 60 and 90). Since there were two tasks, the participants were alternated as to which task they would begin their triage exercise with.

The documents were presented one at a time on a 19 inch screen (the same size screen as our previous desktop studies), in Portable Document Format (PDF). The participants' eye gaze

was tracked using the non-intrusive Tobii x50 eye-tracker, previously reported on (For a detailed specification of the eye-tracker see Section 4.3.4). The participants' screen was recorded using BB Flashback Screen Recorder Standard Edition. A dictionary was supplied for the clarification of unknown words and questions could be asked at any point in the study provided they were related to clarifying participant goals. Guidance regarding the relevance of the documents themselves was not given.

Following the main study, a semi-structured interview was performed in order to gain qualitative feedback from the participants. Parts of the study were played back to the participants and they were asked to explain their thoughts during interesting or confusing behaviours. Participants were rewarded with £10 (GBP) for their time.

6.3   Results and Discussion

This section begins by presenting the pre-study questionnaire feedback. Following the questionnaires, the actual behaviour and participants' navigational patterns and visual focus is presented. This data reports on which document features were viewed most and those which were scarcely or not viewed. We also investigate the participant relevance ratings for the three groups and compare the results with our open-time frame study (Chapter 3). The most distinct findings in each section are summarised and discussed.

6.3.1   Subjective Feature Importance

The results are unsurprising, and are in line with all our previous questionnaires. The main title is always subjectively considered to be the most important feature of the document. The features which occur on the initial page are also rated highly and headings are always ranked within the four most important features. Taking the average ratings given to the document parts we found the two most highly rated features to be the *Main Title* (7.6/10) and the *Abstract* (7.1/10). *Images, headings and tables* received the lowest score (5/10), lower than *plain text* (5.3/10). The *introduction* (6.6/10), *conclusion*(6.7/10), *emphasised text*(6.05/10) and the *headings*(6.85/10) received intermediate average scores. All the categories had a standard deviation between 1.5 and 2.5.

6.3.2 Navigational Patterns

When users perform document triage, they will navigate through a document following a specific pattern. These patterns were recorded and classified using visual inspection method earlier in our research [97] (Chapter 3). When users triage while under the pressure of time constraints, they are found to replicate these patterns. What changes is the number of times each individual pattern is replicated.

The most common navigational behaviour across all the groups, as well as in each individual group, was the "step up navigation". In total each group performed 96 triage instances. The **30 second group** had 4% of the instances performed using a begin and end pattern, 8% of the instances performed using the flatline pattern and 88% of the times using the dominant step pattern. The **60 second group** had 3% of the instances performed using the begin and end pattern, 7% of the instances using the flatline pattern and 90% of the instances using the step navigation pattern. The **90 second group** displayed the least amount of variation in terms of patterns. Only one of the triage instances (1%) was performed using the begin and end pattern while there were no instances from the 90 second group which undertook flatline pattern behaviour. 99% of the triage instances where executed with the step navigationl pattern.

Several behaviours from all three groups mirror findings from previous work [28, 97] (Chapter) 3. The navigational patterns used by the participants therefore, replicated those performed by users with no time limit. There is however, an overwhelming bias in the number of participants choosing to use the *'step up'* navigation rather than any other. 92% of the triage instances were performed using the step up navigation approach, a much larger percentage than any of our other studies. We can hypothesise that the triage performed by the participants was more opportunistic, in that the users would rely on locating relevant pieces of information, as they skimmed through the entirety of the document at a fast pace.

6.3.3 User Attention

Participants were allowed to examine every document for a limited amount of time; up to 5 seconds extra was permitted after the limit was reached for the decision to be orally given to the investigator. Participants with a 30 second limit, 60 second limit and 90 second limit are referred to as Group

30, Group 60 and Group 90 respectively. The experimental setup used a Latin square approach to reduce effects such as fatigue on the behaviour of the participants. A summary of the visual attention given to each feature can be found in Table 6.1. The figure also includes the area which each feature occupied for easier comparison.

	Viewing Area (%)	Group 30 (%)	Group 60 (%)	Group 90 (%)
Initial Page	10.8	42.7	27.0	46.7
Main Title	1.1	10.8	5.5	9.5
Abstract	1.2	16.5	15.1	21.1
Introduction	2.3	8.1	5.9	7.1
Headings	2.1	15.3	8.1	8.2
Emph. Text	1.9	2.2	2.8	4.6
Images	6.2	4.9	6.8	10.5
Diagrams/Tables	2.3	12.5	4.6	6.3
Conclusions	1.6	6.1	5.0	7.5

Table 6.1: Average document area per feature Vs. Average visual attention (%) between Groups.

Total Time Participants in Group 30 took an average of 33 seconds to complete triage on a document (SD: 11.8 secs), thus using their whole time on average to the full and had to be asked to give a rating. Participants in Group 60, took on average 45 seconds to complete triage per document (SD: 15.2 secs) and to give an answer. This is an average of 15 seconds less than the mentioned limited time. Group 90 Participants completed triage on average at 61 seconds per document (SD: 18.8 secs), an average of 29 seconds less than the limit time. An ANOVA test produces significant results between the average times taken and the amount of time given to the participants at the $p < 0.05$ level for the three conditions [$F(2,285) = 140.8768368$, $p = 0.286^{-42}$]. Post-hoc Tukey's HSD tests showed statistical significance between all groups at the 0.01 level.

Participants were asked to comment on when the relevance decision was made. The feedback presented the initial page as the point at which the decision was made, whereas the remaining triage was performed in order to verify their decision. When questioned directly about why they did not therefore spend all their time on the initial page, since the time constraint was strict, participants from all three groups suggested that they "just wanted to be sure that the document was what I think it is". In other words, they wanted to make sure that the abstract and main title were not misleading. Our data verifies the subjective importance that users attribute to the initial page. We also notice however, that being able to scrutinise the entire document rather than only a part

of the document, is also considered a benefit for the users. We remind the reader that previous work [28] gave us evidence that people would often not read the whole document, and with a speed restriction, it may be that this behaviour would be more common.

Initial Page The initial page of a document most often includes the main title, abstract and part, if not all, of the introduction. The average area taken by the initial page was 10.8% (SD: 9.99) of the document area. Group 30 and Group 90 spent almost half their viewing time scrutinizing the initial page (42.7% (SD: 49.28) and 46.7% (SD: 33.39) respectively). Group 60 participants had a lower average, with 27% (S.D: 9.09) of their time spent viewing the initial page. We are not certain of the reason why there is a drop in visual attention on the initial page by the 60 second group. We eliminated the effects of bias from previous triage operations by alternating the order in which the the individual time constraints are assigned to the participants. Although further testing will be needed to identify the cause for this phenomenon, we hypothesise a reason. We are reminded that participants follow a largely 'step up' approach in triaging documents when there is a time limitation. It may be the case that participants in Group 60 did not have the time to return to the initial page after scanning the remainder of the document. An ANOVA test revealed a statistical significance at the $p < 0.05$ level for the three conditions [F(2,285) = 13.9630965900541, p = 0.0000016]. Upon further scrutiny Tukey's HSD tests showed statistical significance between the 30 and 90 groups and the 60 and 90 at the 0.01 level. There was no statistical difference between the 30 and 60 second groups at the 0.05 level.

Figure 6.1 shows an example of a typical triage performed by an information seeker and eye-tracked on the initial page of a document. The figure shows an eye-gaze pattern, identifying the fixations of a typical user's view while looking at sections of the initial page. This visualisation represents all three of the groups.

In general, the initial page is scrutinised extensively compared to the remaining features, and this is reflected in the reduced attention of the remaining features. The increased attention of the initial page; specifically the abstract and introduction, creates a domino effect of reduced attention on the remaining features. The statistical relevance between the times, suggests that the 60 and 90 second groups are more inclined, and able, to scrutinise the remaining features, whereas the 30 second group mostly choose to focus on the initial page features, yet skim rapidly to verify their
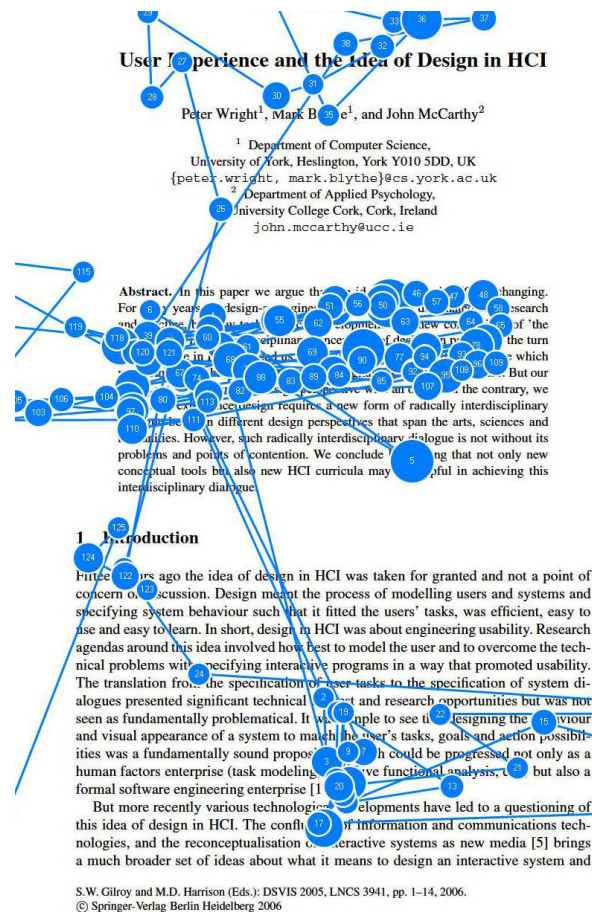
Figure 6.1: Gazeplot showing an information seeker's attention on the initial page.

decisions on the rest of the document.

Main Title The main title, which was subjectively rated as the most important feature to assist in document triage on a document, was viewed by all our participants. Focus times on these areas were 10.8% (SD: 15.44), 5.5% (SD: 9.28) and 9.5% (SD: 9.01) respectively for the three groups while the area of the main title occupied an average of 1.1% of the document area. A substantial amount of visual attention is given to the main title even with a limited time constraint, making it the second most visually important document feature across all three time groups. No significant differences were uncovered using an ANOVA at the $p < 0.05$ level [$F_{(2,285)} = 2.06141997633082$, p = 0.129166646] between the three groups.

Abstract Average viewing times are 16.5% (SD: 19.66), 15.1% (SD: 7.88) and 21.1% (SD: 8.64) for the groups respectively, while the average area occupied by the abstract was 1.2%. The abstract is therefore, the most visually processed feature on the document on all three time constraints. This

mirrors previous results on the desktop and the small screen from our previous chapters and verifies our model's validity (see Chapter 8) even with the constraint of time. An ANOVA test revealed a statistical significance at the $p < 0.05$ level for the three conditions [$F_{(2,285)} = 7.84885414936385$, p = 0.000480674]. A Tukey's HSD test reveals statistical differences between the 30 and 90 groups and between the 60 and 90 groups at the 0.05 level.

Introduction We remind the reader that the introduction is often located on, or partly located on, the first page of the document. As seen, the initial page is weighted high in viewing time to viewable area ratio. The average area taken by the introduction was 2.3%. Viewing times for the introduction were, respectively, 8.1% (SD: 5.91), 5.9% (SD: 7.86) and 7.1% (SD: 8.64) for the three groups. Within all three groups, the introduction rates moderately when compared to all the document features. The visual processing ratio however (3.52, 2.56 and 3.08 respectively), testifies to the fact that there is attention given to the introduction by all three groups. Using an ANOVA test we find strong significance between groups at the $p < 0.05$ level [$F_{(2,285)} = 12.25319907$, p = 0.00003467]. Upon further scrutiny, we find statistical significance between all groups at the 0.05 level using Tukey's HSD.

Headings Headings have been commented on as having "a modest impact on users' navigational behaviour" during document triage [97] (Chapter 3). When comparing the ratios during fast triage there is an exception to this. The average heading area on a document was 2.1%. Viewing time averages were 15.3% (SD: 6.47), 8.1% (SD: 6.47) and 8.2% (SD: 6.47) giving time to area ratios of 7.28, 3.85 and 4.68 respectively. These percentages create a visual processing ratio of double that reported on our previous studies for Group 60 and 90, and more than triple the time in Group 30 (7.28, 3.85 and 3.90 respectively). These findings testify to a very high level of attraction by headings while an information seeker is performing document triage with a limited time. Using an ANOVA test we find strong significance between groups at the $p < 0.05$ level [$F_{(2,285)} = 42.25319907$, p = 0.000000000000000007]. Using post hoc Tukey's HSD tests we find statistical significance between all groups at the 0.05 level. Headings receive increased attention when an information seeker performs within-document triage with a short time constraint compared to our previous findings. Our previous research gives headings visual processing ratios of 1.91 (Desktop Study - Chapter 4) and 2.14 (Small Screen Study - Chapter 5). Users with time constraints

produced area to time ratios of 7.29 (Group 30), 3.86 (Group 60) and 3.90 (Group 90).

Emphasised Text The average viewing area occupied by emphasised text was 1.9%. Average focus times were 2.2% (SD: 1.51), 2.8% (SD: 1.84) and 4.6% (SD: 3.09). This shows that, on average, emphasised text receives slightly elevated scrutiny during fast triage. As previously reported in our small screen and desktop studies, emphasised text does not receive a large share of the users' visual attention when compared to the other features. Using an ANOVA test we find strong significance between groups at the $p < 0.05$ level [$F_{(2,285)} = 6.757003504$, p = 0.001358066]. Using Tukey's HSD tests for multiple comparisons, the 30 and 90 groups and the 60 and 90 groups have statistical significance at the 0.05 level.

Images Images had an average area of 6.2% and an average viewing time of 4.9% (SD: 8.87), 6.8% (SD: 8.94) and 10.5% (SD: 7.74). They were considered the least visually important feature across all groups. Again, this correlates with our previous findings in that images are, although viewed, seldom focused upon to extract further information. Using an ANOVA test we find strong significance between groups at the $p < 0.05$ level [$F_{(2,285)} = 4.909462175$, p = 0.008012243]. A Tukey's HSD post hoc test reveals a statistical difference only between the 30 and 60 groups at the 0.05 level.

Diagrams and Tables The average area occupied by diagrams and tables was 2.3% Average focus from the participants were 12.5% (S.D: 5.94), 4.6% (S.D: 4.90) and 6.3% (S.D: 4.40). Our previous study data suggests that diagrams and tables do not present a visually significant feature regarding document triage. The time to area ratio of the three groups in this chapter, show elevated time to area ratios with values of 5.43, 2 and 2.73 respectively. We hypothesise that diagrams and tables capture the visual attention of the users, who although focus on them, are not able to extract sufficient information using the limited time available. Using an ANOVA test we find significance between groups at the $p < 0.05$ level [$F_{(2,285)} = 3.474228397$, p = 0.032304375]. Post hoc testing finds statistical significance at the 0.05 level betweent the 30 and 90 second groups, using Tukey's HSD.

Conclusions Section The conclusions section is considered a highly influential feature regarding users' visual focus during normal document triage (See Chapter 3). The average area occupied by the conclusion was 1.6%. Average viewing times however, were elevated with 6.1% (S.D: 6.88),

5.0% (S.D: 9.14) and 7.5% (S.D: 5.80). Our previous results rate the conclusion high in terms of ranking compared to the other features. When performing the eye-tracking study for example, the conclusion was rated as the second most visually important feature. In our small screen study, the conclusion ranked third in terms of time to area ratio. The time to area ratio of the conclusions section when a time constraint is applied, is considerably lower than our previous studies (3.81, 3.125, 4.6875). We hypothesise that since the conclusion is located towards the end of a document, this makes it harder to find. We also know that the attention of participants falls inversely to the page number of the document (See Chapter 3). The conclusion therefore, being close to the end can also be a contributing factor to the reduced attention of the information seekers. Using an ANOVA test we find significance between groups at the $p < 0.05$ level [$F(2,285) = 4.378387468$, p = 0.013400407]. Statistical differences, using a Tukey's HSD post hoc test, exists between the 30 and 60 second groups at the 0.05 level.

6.3.4 Document Relevance Ratings

Figure 6.2 shows the average rating and standard deviation for the three groups, as well as the average rating from our findings in chapter 3). A Mann-Whitney test for between subjects studies with non-parametric data gives us no statistical significance between the relevance ratings of the three groups.

	Group 30 (%)	Group 60 (%)	Group 90 (%)	Group ∞(%)
Average Score	5.10	5.82	6.02	5.99
St. Deviation	2.41	2.42	2.15	2.20

Table 6.2: Average ratings and standard deviation for Group 30, Group 60, Group 90 and Group $\infty(open time frame)$.

The above statement holds when comparing the results with users from our previous studies. This group of participants performed triage on the same set of documents with no time restrictions (See Chapter 3). The current evidence does not produce any statistical significance to indicate that there is any positive or negative influence of time on the relevance decisions. Since we are aware that information seekers are prone to making several erroneous relevance decisions (See Chapter 3), we can infer that the same is true when given a time constraint. Our findings reveal however, that statistically, this error rating did not decrease and therefore reducing time will not reduce

the efficiency of the triage task. Further scrutiny is needed into the exact cut off time between relevance ratings and time constraints. We remind the reader that the initial page is also a dominant part of the document and therefore can influence the relevance decision the most. As this is the first of the document to be viewed by the information seekers, this can also be the reason for the reported non-statistical change in relevance decisions. Efficiency is measured in many ways. Ben Shneiderman identifies one measurement of efficiency to be time [140]. Thus far our participants with a time constraint performed with the same relevance accuracy as our participants without a time constraint. We have evidence to suggest that applying a time constraint can improve the overall efficiency of the document triage process. Although this statement is not tested in its entirety, we can hypothesise that excessive time in relevance decisions can produce limited results. Using the satisficing theory, we can influence 'how much is enough' and measure the effectiveness of this constraint.

6.4   Summary

In this chapter the effects of limited time on document triage in a process we label *rapid triage* are reported on. These are presented in detail in the results and discussion section. Several findings, correspond well with our previous experimental data and we can verify their validity during rapid triage also. These include the subjective ratings of the participants on the document features and the extensive visual scrutiny of the initial page and features such as the abstract and main title. We are able to make further reports also on new evidence which contributes to our understanding on document triage.

Although the 'step up' pattern is the most popular navigation method for document triage, according to all our previous studies, we see an increased use of this pattern in this specific study. Given a limited time frame, information seekers will choose to navigate using a 'step up' behavioural pattern more frequently than without the constraint of time (Significant to the $p < 0.05\%$ level using a t-test, N = 416). We have reported that the initial page is considered by the participants, both subjectively and from the observational data, to be the most visually attractive page of a document for relevance decisions. We also know from all our previous experiments that users claim to make a relevance decision during the initial part of their triage process and most often from the

content of a document's first page features. Even given a limited time however, information seekers will most probably choose to navigate beyond the initial page in order to "verify their decision" using the rest of the document.

Previous work produces the hypothesis that, given limited time, readers would effectively focus on the most important parts of the text and sacrifice non-important parts of the text [47, 105]. Without speculating what features are cognitively important to the information seeker during triage, we are able to report on the features that are visually scrutinised during a limited time frame. The findings are similar to our previous studies in that the initial page is heavily scrutinised. A representation a the users' typical visual navigation can be seen in Figure 6.2.
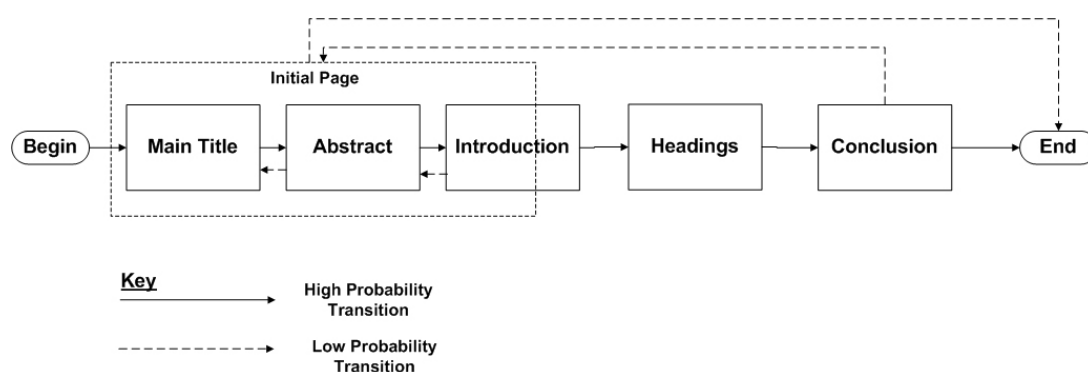
Figure 6.2: A Typical Navigation and Visual Attention of an Information Seeker During Rapid Triage.

In this model, we present the top 5 most visually processed features which attract an information seeker's visual attention during rapid triage. This behaviour allows us to hypothesise that, when faced with a time constraint, users will scan through the information quickly and focus on areas which are not devoid of visual features. Pages with complete plain text are ignored, while all other features are focused upon. We remind the reader that the most common navigational behaviour during rapid triage was the step up navigation (See Chapter 3) When compared to our previous chapters using a student's t-test, we find statistical significance between the participants in the rapid triage study and the one without a time constraint at the $p < 0.05$ level, N = 416.

From their limited time triage activity, our participants gave relevance decision scores on the documents. Our intention, was to compare the ratings given by the participants under rapid triage conditions to the ones without a time constraint. Our data produce an interesting find;

namely, that there is no statistical difference between the relevance decisions of our rapid triage participants and our unlimited time participants. If reducing time does not produce a different relevance rating then, according to Shneiderman's efficiency measures [140], we have a more efficient triage process. Application of this finding to a document triage software system could produce improved performance for the information seeker. Further study of this however, would be required to report on accurate design principles.

To conclude, our findings contribute to the previous data from our studies to produce key similarities and differences in behaviour, thus verifying or rejecting previous hypothesise such as the visual attraction of the initial page and the features which are scrutinised for the initial decision on a document. We are able to then use the data from all our studies in order to produce unified models of the document triage process as we have observed it in our investigation. By changing key variables, we have already begun to formulate conceptual models to incorporate the generalised behaviours during triage from our lab based studies.

Chapter 7

Diary Study of Multidisciplinary Document Triage

7.1 Introduction

This chapter presents a diary study in which the participants are required to keep records of their triage activities as they occur, over a period of time. To better understand the variation in document triage across a wider range of people and circumstances, participants from a wide range of backgrounds and disciplines are selected. This study aims to gather information from a richer pool of participants to the studies presented thus far and add to the richness of the results by incorporating real world data. We are specifically interested in seeing how real life behaviours match our lab-based studies and to identify further areas that require investigation which did not surface during lab-based tests.

Document triage is a complex task, with many variables presenting themselves (See Section 1.2). These variables, such as time and a user's experience, can all contribute towards the behaviour and actions of information seekers during their triage activities. It would be unrealistic to attempt to cover all the variables in one thesis. A suitable starting point was selected; one which would cover the most familiar domains according to the investigator's experience. Thus far, only controlled lab-based studies were undertaken. The documents that were to be triaged were always selected specifically by the investigator. Each study altered a variable, such as screen size and time con-

straints, to identify variations to triage behaviour. We have not reported or investigated however, how the document triage process would compare against users from a wider range of disciplines, who triage different types of documents, and specifically, on data which was extracted under real triage situations. One example, is that we always had a set amount of documents presented to the user or a time constraint. This limited the time which the users would spend performing triage and also the amount of documents that were 'returned' from a search query for them to triage. We can now report on further findings to apply to our models and measures (See Chapter 7).

7.2   Study Design

7.2.1   Background and Design

In order to begin exploring the document triage behaviour on a wider type of participants we needed to choose a suitable methodology. In order to elicit information from our participants, we chose to implement a diary study methodology. We selected this approach which is similar to Kulthau's methodology when formulating a model for the information search process [88, 89]. Kuhlthau's research included methods such as journals, search logs, questionnaires and short pieces of writing to gain insight into the actions, thoughts and feelings of users. Our study design was explicitly tailored in order to gain qualitative feedback from our participants in a way which was as non-intrusive as possible and borrowed parts of Kuhlthau's methods. Very little intrusion into the participants' document triage process was made, with short pieces of writing, guided by a questionnaire required. This method helped us to effectively extract information about the user's experience over time and it is in the actual context of use. The use of a diary study was chosen over other alternatives such as ethnography or field studies which contain direct observations.

Observation can be obtrusive and participants may alter their behaviour when an observer is present; something that we specifically did not want as our objective was to register real scenarios. Notes and videotapes could have been used and analysed after the participants' triage activities. Due to ethical constraints and technical limitations these would not be as useful as the diary logs and so this approach was not taken.

We chose to give participants logs to complete during their triage activities in their natural

environment. These were guided by example questions that should be asked by the participants when completing their logs. We included three interviews (one prior, one during and one after the diary study) to complement the information provided by the participants. The results were coded using an adaptation of Strauss et al [146]. This can be seen in table 7.1. The codes were created from a priori ideas from our studies and existing theories.

Name	Example
Behaviours	Stop search
Strategies	Looking at documents linearly
States	Frustration and hopelessness
Conditions or Constraints	In a hurry

Table 7.1: Coding Scheme - adapted from Strauss et al [146]

We begin by describing the participant selection process and proceed to describe the study format in detail. We then present the findings, grouped by categories. A discussion of the main points concludes the chapter.

7.2.2 Participant Selection

A wide range of information seekers were recruited, based on their general information needs. A total of 20 participants were chosen which included disciplines from engineering, psychology, criminology, law, computer science, finance, chemistry and design. Two high school students and an unemployed housewife were also recruited to allow for a non-specialised audience and a larger age range. Participants were aged between 12 and 38 years old. All participants were given a 10 pound (GBP) gift voucher as a token of appreciation for their time.

7.2.3 Study Description

A semi-structured interview was conducted before the main study took place. This interview questioned participants with regards to their searching habits. Questions included the average amount of time they spend looking for information (not reading) each day, the methods used to triage information, the means and what they find important when searching for information. Context gathering questions were also asked, such as how long they have used the internet for searching, what experience each participant has with searching for specific information and the

usual time constraints they are likely to face while searching and performing document triage during the study.

The participants where then instructed on how to complete the diary, the type of events to record, and the information to note about each event. Three main things were asked of the participants to be recorded. These were, physical actions, thoughts that participants would be thinking during the steps of the document triage process and feelings which surface from their information seeking task at all points.

Participants were given a series of questions to guide their entries as a whole rather than for each individual event. They were encouraged to read the questions carefully before beginning the diary logs. The participants were also allowed to use recordings if they so wished to record their triage process [1]. Questions included were:

- "Is the information you are looking for of a personal nature or work related?"

- "Is the search for a specific item or a broader search for knowledge on a subject?"

- "Comment on how important it is that you find the information you are looking for?"

- "What information need do you have? What is it that you are looking for?"

- "How well do you know the subject area of the item you are looking for?"

- "What were your time constraints in terms of getting the information needed?"

- "How much time did that search take you in total?"

- "What means (physical or electronic did you use to search?)"

- "Do you think the means I am using are good for searching?"

- "What seems to be the problem locating the information required?"

- "How do you feel during the steps of triage? Before during and after."

- "How many documents are returned on the query search?"

---

[1]none of the participants chose to follow this option.

- "How many and what documents did you open and why?"

- "If you are returned with a results list what do you look at?"

- "What do you look for during the within-document triage process? What do you look at first?"

- "What do you think of each of the scrutinised documents?"

- "At which part of your triage have you realised the potential relevance of the document initially?"

- "Do you keep performing triage after the initial relevance decision you have made? If yes why?"

- "Do you change your mind as to the initial relevance decision after you look at a document more?"

- "At which point do you stop the triage and why?"

- "How confident and pleased are you with a) the quality of the selected documents b) the amount of information you have potentially extracted from them?"

- "Will you repeat a search later for the same information need?"

- "Do you read the documents later and realise that the relevance decisions you have made during within-document triage were false?"

Participants did keep close to the guidelines although they were encouraged to include additional thoughts and information they thought was appropriate or wanted to include. They were asked to fill in the majority of the diary after a triage process was complete rather than interrupt their triage process and thus adding bias to the results.

Participants were permitted one month, at most, to complete the diary study and were requested to limit the diary entries to ten triage operations. The entries were limited so as to improve compliance and response. It would be unrealistic to ask them to record every triage activity undertaken. Furthermore, repetition could cause our participants to mirror most of their entries

rather than reporting in detail. Our participants were encouraged to randomly register entries and not just ones which simply occurred in strict sequence; perhaps by registering some same day searches and ignore some searches before registering another. Upon the participants completing two or three of their document triage process entries, an intermediate interview was conducted, in order to examine their initial diary entries, check whether the logs in the diary entries were kept correctly and to also ask the participants any questions about them while the search was still relatively fresh in their minds.

At the end of the study, a final semi-structured interview took place in order to again ask questions related to behaviour and the triage process while the triage searches were, again, relatively fresh in their minds. We specifically designed the interview to be flexible, so as to allow the investigator to focus on points which were considered important. Most questions related to explanation of the entries in the diaries in order for the entries to be easily and effectively analysed later.

7.3   Results

7.3.1   Pre-study Interview Results

During the **pre-study interview**, participants estimated that, on average, they spend 84 minutes looking for information per day. We had asked the participants to describe how they would spend this time and they reported that they would spend it looking for sources which were relevant to their information need. They would look for websites, papers and books which would contain material they required and would either skim or read the material in depth to extract the information they wanted. This part of the interview gave us interesting insight into a part of the document triage process that is becoming more and more evident. Information seekers do not stop performing document triage after they choose some documents, whether these are books or websites or academic documents. They simply narrow down the search results and subsequently rat the documents with further reading and in depth scrutiny. Thus far, the document triage process was considered to begin with the production of search results and end when the initial selection of the documents was made. We have further evidence, to argue that the triage process extends into the reading stage. This is further supported later from the participants' comments during the triage exercise. The

standard deviation on this subjective rating was 75.2. This variation testifies to the different types of search activities undertaken by the participants and the different search strategies used. This evidence was not reflected in our lab-based studies previously. We begin to understand how users triage beyond the linear setup in our previous studies. For example, one participant who was about to become a mother, would "leisurely browse" search results, taking breaks in between some of the documents to "take care of other things". Other participants reported a "lack of time" in order to make their decisions which caused them to simplify their search. Unfortunately, the data we have does not allow for a richer examination of this specific types of behaviour, but we know have enough evidence to justify further investigation on the effects of these. Participant 4 would comment on being very "judgemental" and "easily abandoned" documents or web pages that did not "have the information required jump out at you". Two participants (15 and 18) reported that the study time would be a particularly busy time for them search wise, as they would have increased search activity due to their work. Participants were asked whether they are likely to be using electronic means, physical material such as books or simply ask other people to obtain the information required. Responses indicated a heavy bias towards electronic means in that participants indicated that, on average, 70% of their searching time would be using electronic means, 22% using physical means, such as libraries and only 8% of their time would be conversing with other people for information. Two participants reported that they only use electronic means when looking for information. These two individuals further reported being given access to library material at university but that "there was never a need to actually look for books or journals in book form" since "everything they wanted was either online or could be e-mailed to them electronically". Our previous studies on the differences between electronic and physical document triage [28]

Participants described typical searches (material they would look for, features they look for to make their decisions) they are likely to undertake during the study. These included a broad range of material, either for work or for entertainment value, mostly over the internet. Participants described their triage activity in terms of their predicted navigational and visual behaviour, focusing mainly on the main titles of results lists before proceeding to traverse a potentially relevant page by skimming the contents for any relevant headings or links that would lead them to the relevant information. All participants reported moving on to in depth reading directly after recognising a

document to be relevant when performing triage on the internet. Participants 6 and 10 continued to talk about how their in depth reading stage was a further verification of their decision being right or sometimes "way off when actually reading the material". We hypothesise that triage does not stop at the results list level, or the skim reading, but continues further into the reading stage. This is later reported and verified in our main diary study comments.

When asked for what problems they can think of that makes their triage process more difficult, 8 participants reported having to triage through "too much information". Other difficulties included "annoying advertisements" which would distract the information seeker from their task at hand or portray irrelevant information and "waste time". Reliability of the information found on-line was also another factor which was thought of as a "danger" in searching, and was reported by 4 participants. Interestingly, yet not surprisingly, none of our under-age participant mentioned the reliability of the information that is found on the internet as a problem. When directly questioned about the subject in the interviews, they did not produce a strong opinion against the internet being unreliable in relation to the material they required. One of our younger participants commented: "as long as we tell them [teachers] where we go it from they are OK with it".

When the participants were asked why they preferred electronic searching, all participants commented that the speed at which they could access and find the information using electronic means was much superior to the other means available. When questioned more, we were made to understand that the process of ordering and waiting for a book or paper in physical format was often not an option for participants. The reasons went beyond the time limitation and took more of an impatience factor. Interestingly, one participant stated that "sometimes I'm lazy and will keep searching to find an electronic source from somewhere else rather than a book that I know has the information I need from a library". 15 participants also reported on successfully locating information when searching in a digital context since "there always seems to be information on something". Interestingly, 14 participants mentioned how the internet can also be "littered with unreliable information" which can often be a barrier to "quoting it" in their work tasks.

Participants were asked to give any positive aspects they think physical material, such as books have, both in the interaction as well as any intrinsic values they can think of relating to triage. The dominant feedback given from 12 of the 20 participants was that physical material such as

books taken from a library are "trusted sources of information" that "can be quoted reliably". Interestingly, 5 of the participants mentioned how searching for information through books "feels better". Looking at the data, we question whether the triage activities are influenced by this factor; namely the uncertainty of the reliability of the information. If the users are returned with the information needed through a digital source, would they continue triage to verify the information through further resources?

7.3.2   Triage Process and Search Results

We now present the key findings from the participants' diary studies. We present the data as reported by the participants, incorporating the feedback received during the mid-study interview. We categorise the findings to report on distinct triage findings.

7.3.2.1   Searches and Duration

In total, 178 searches were made. 78 of the searches were work related, 97 were of a personal nature and 3 were classified as for both. Searches for work were usually accompanied by a larger stress level by participants. Comments like "this needs to happen now" or "I had to get this seriously quickly" were common place in mostly work related questions. Due to the difficulty in accurately measuring the time required for every individual triage process to be measured, participants only registered the amount of time spent on the entire search process, to accomplish the goal they were looking for. This included the triage of all the queries searched for, during an information need, the traversing through the results lists, searching for the relevance of the documents by scanning through the full text and even the in depth reading which followed. On average, the complete search process took 51 minutes (SD: 145.0 minutes).

7.3.2.2   Search Constraints

84 of the searches were made on an information need that the user was not too familiar with, such as "finding fast track courses" which they have not looked for before. When asked, the participants performing triage for work related purposes under this category were the ones more anxious about whether they would find the result they required. One of the participants had an interesting quote,

saying that she needs to "be extra careful in order not to miss something important". The triage scrutiny level is again dependant on the initial attitude of the participants, a factor that needs more exploring. Thus far, only the time constraint was imposed in our work to emulate this effect. The remaining 94 searches where based on familiar topics to the information seekers. 86 of the searches had no time constraint, while only 15 of the searches were required as quickly as possible with a very limited (15 minutes or less) time frame. 125 of the searches were reported to be "important" in terms of the need to obtain correct and reliable information in order to satisfy their information need. 125 of the searches were considered *specific* while 48 were considered broad. The remaining 5 were classified as both.

We asked participants to clarify 'broad' and 'specific'. Participants commented that the specific searches were usually one item of information such as "the address to email fraud report to [BANK] branch". These were more information retrieval in nature in that the information was certain to be found, as reported by the participants. The broader searches reported were specifically browse strategy oriented and more information seeking in nature, meaning that the structure of the answer was not always clear. Furthermore, the answer to the information need was not always certain to be found. One example of this was look for "anything related to the quote [political apathy]". The participant reported that he did not know what to expect to "come up with in the end" due to the "vague request and vast information" available. It should be noted here that the searches which were classified as important did not constitute a one to one relationship with the searches reported as specific and that the number is coincidental. 163 of the searches were made using electronic means while 12 were using physical material such as libraries and books. Only 3 searches were made using one to one communication with other individuals. These searches were of particular interest to us since, beyond being the only ones to use interaction with other people, they were also always combined with electronic searching. For example, one of the searches found our participant requiring a technique to "how to normalize the y-axis in gnoplot". The participant mentioned that before attempting to look for this on the internet, "it was easier to ask the person close to him for suggestions". Another search constituted of the same requirement; namely, the solution to a technical problem. The participant in this case asked her supervisor for the answer and guidance rather than look for a likely answer on her own. She reports "less chance of wasting time and going

wrong than on my own".

All the participants that reported asking another person for information, reported doing so due to the complexity of the problem in conjunction to their own inexperience. Participants were generally certain and confident that the tools they were using were appropriate for the information seeking task. Only 8 of the tasks were reported as suffering uncertainty as to the ability of the tools to produce the required results. Interestingly, all 8 of those were conducted using electronic and not other means. When asked why they were uncertain of their tools, the participants reported two main reasons. The first reason was "this may not have the information that I need in it"; In other words, the database may not include the required data. The second reason reported is interesting from an interaction viewpoint. Our participants reported a lack of trust in themselves to effectively communicate their needs to the electronic system as well as a lack of trust in the system to be clever enough to understand their needs. As one participant commented: "... not sure if its getting what I want to tell it".

7.3.2.3 Triage Barriers

The participants were asked to report on barriers to the search and triage capabilities of the method they were using. The first barrier that was heavily reported on was that of time constraints. This is an interesting observation since only 15 of the searches were reported to be constrained by time at the beginning. When we asked the participants to explain, it was found that more time was needed to go through the information returned than a user originally allocated. We hypothesize that the complexity and information overload [18, 53] of the information need is not evident to an information seeker. There appears to be an illusion that using electronic means is a fast easy process for locating information. When the information seeker engages in document triage in order to locate relevance and then extract the appropriate information, barriers with interaction and the lack of supporting software is made evident. This can be characterised by the expression of one of our participants "so much rubbish to go through, why can't you just tell me what I want". Another participant also reported wanted "a simple answer rather than all this philosophy and [word removed]".

The second barrier to the interactive triage process that participants reported on was that of

"lack of knowledge". Participants were feeling that their own knowledge of where to look on the material presented to them was an issue which reduced their potential to locate relevant documents. Furthermore, not being able to identify relevant material when it is actually presented to them, such as a link to a good source of information, was also a concern. One of our participants, who was looking up the answer to a technical problem, put it nicely: "I may be staring at the answer but its not clicking cause of all the advanced mambo jumbo". This, coupled with reports on limited and hard to locate resources, made the document triage experience difficult for information seekers. It is interesting to note that none of the participants blamed or criticised the tools they were using but instead blamed themselves for the inability to triage effectively. A reported phenomenon in HCI studies is that a user will, more often than not, blame himself rather than a system or interface [114]. We hypothesise that the same is true in this case, as barriers to interaction, such as skimming plain text aimlessly, has been reported in previous studies in this thesis. When participants were asked whether they changed tools in order to attempt to provide themselves with better information, the majority (162 from the 178) reported continuing to use the same tool. This is an interesting behaviour that we scrutinised further, again to be given the same reasons; namely, "it's probably me not understanding what to look for or not able to find it". We questioned one participant further specifically asking the question as to whether they had ever thought it is the tool's problem and not their own. Their response was "even so, it's the tool that most people use and so I kept using that one cause its the best chance I have [at getting the information]".

7.3.2.4   Confidence Levels

Participants were asked to record their feelings, including confidence levels, before they begin the triage process, during the process and after they end the triage. Our source for obtaining this information was Kulthau [89] in which the information seeking model presented includes feelings as well as user actions. Care was taken to not report on the feelings of the users during the entirety of their information seeking process; something which has already been reported on [50].

42 triage activities began with a negative emotional start, such as uncertainty (as described by Kuhlthau [90]), worry and lack of confidence. Most of these were work related and our participants were under stress to find the correct data. One participant commented: "I need to find it, don't

really have an option not to and it needs to happen 10 minutes ago".

In contrast to the personal activities, work related triage appears to have a more directed and higher satisficing threshold than the personal triage instances. 63 triage activities began with a neutral emotional frame of mind and confidence level. The remaining 75 triage activities began with a positive confidence level and emotional cue. During their triage, 54 triage activities were reported by the participants to be negative in terms of feelings and confidence. Again, the majority of these were related to the work triage instances. "When time is running out and you got nothing it starts annoying you and you push harder and faster" said one of our participants who reported being extremely stresses and lacking confidence on a deadline.

52 activities were reported to be neutral while 72 of the activities contained a positive confidence rating. At the end of the triage activity 52 activities produced low confidence and feelings for the participants, 20 of the activities produced a neutral level of confidence. The remaining 106 were were completed with a positive level of confidence. It is interesting to note that the correlation between positive confidence and feelings at the end of the triage activities, does not produce a 'one-to-one' relationship with the activities which were successful in producing relevant information. Once again, upon further scrutiny we understand the reasons behind this to be the importance of the triage. In other words, some participants "didn't care about not finding exactly what I needed, I'll look again perhaps another time". Some also commented that "I found information which isn't perfect but needed to use it so not very happy about that". This produces another instance when satisficing is seen in our work [47]. The participant had a lower satisficing threshold on this particular activity.

7.3.2.5 Document Selection

On average, 1,674,357 documents were returned to the participants per document triage activity [2]. The distribution of the result amounts can be seen in Figure 7.3.2.4. Users are inundated with more documents than they are willing to, and are physically able to, triage on the majority of their search results [18]. Interestingly, our participants did not feel like the actual results retuned were an issue. "I only look at the first page so I don't care how many things its [the search engine] found"

---

[2]Most participants reported the number of results given by the search engine
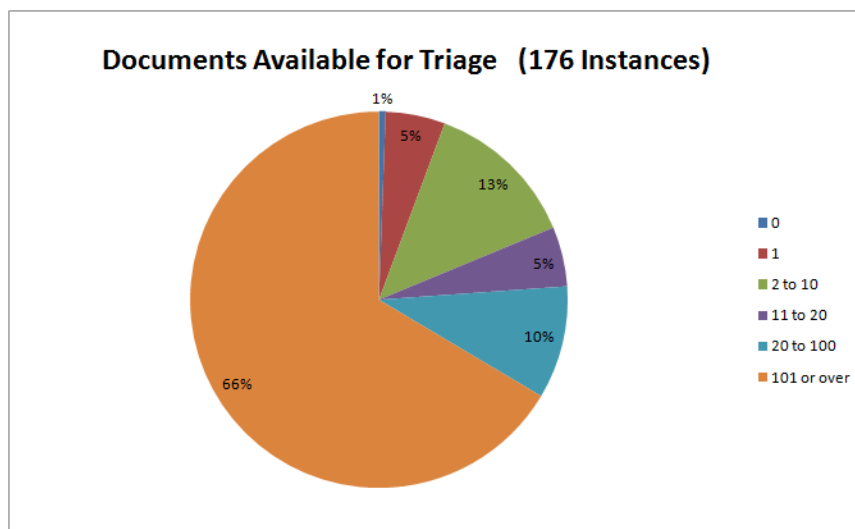
Figure 7.1: Documents Returned on Query Per Triage Activity.

was a natural response from our participants. It seems that the source of information overload is the actual duplicate information sources which present the data in a similar way. This confuses the information seeker into which source is correct and which of the information is relevant. This applies primarily to the broader search instances rather than the specific ones. The triage operation thus begins with the immediate expectation that the majority of the documents returned will be rejected. It is assumed that information seekers will not triage deeper than the second results list page [144]. Our participants commented rarely reaching the third page and generally not willing to navigate to the second page. "I will just change my search" was a natural response when asked what they would do in the interview. When asked for a reason, the participants would think and usually not have a specific reason for this behaviour. A typical results list will present 10 results per page. The typical information seeker will therefore actively triage a maximum of 20 documents during his or her entire document triage process. By taking the average documents returned on a search this equates to merely 0.001% of the documents likely to be within the visual attention of the information seeker. The remaining 99.998% of the documents will be rejected by default.

Our participants chose, on average, just under six documents (5.85) to open to view them in full text. Upon asked, participants commented that they would open something if they were (a) familiar (b) interesting or (c) the answer is evident from the surrogate level. The diary entries contained entries describing the behaviour and visual attention of the information seekers during their triage activities. Most of the documents triaged by the users were websites. Unlike academic

documents, these do not have a set structure but contain features which are almost universal to a web page layout. For example, a web page will contain headings, images and most likely an introductory text. Participants were asked to include features that they looked at or helped guide them to the information they looked at in their entries.

7.3.2.6 Reported Behaviour

The participants commented on the first action they took upon receiving the results. 59 of the search diary entries report that headings or titles were the main feature that the participants looked at. We asked the participants what they remembered looking at first and why. The response gave evidence of the need for textual information and so the focus was mainly on textual information. Furthermore, headings would "make it easier to understand the contents and if I'm right or not". Headings were the feature most frequently reported by our participants. Only 11 searches reported actively looking at images or pictures while performing triage at the first stage after receiving results. 40 participants reported on skimming or "scanning" for information as a first stage in their triage, rather than reading anything in depth. Only 24 of the participants reported on reading in detail as a first action. When questioned further the reasons were "to get an introduction to the topic". The excerpt or snippet of any given document was only reported to be scrutinised on 11 of the searches.

The majority of the searches (123) reported on using a linear approach to performing document triage within the full text of a document. The main reason for this was that the most common form of target document our participants were looking for was text based. The fold of the websites were reported to be scrolled for subsequent information regularly. Most participants would comment on usually looking at the entire page "from start to finish but only at a glance". When asked to elaborate, our participant commented that he did not want to miss any information if it was obvious. Furthermore, he wanted to see the whole of the document "even just for a bit". Similar comments were made in Chapters 3 and 8 were participants favour looking at the whole document to verify their decisions, even if they are convinced by the first part. 79 of the searches described headings and titles being used to navigate and extract relevance. Interestingly, participants commented only in 22 searches of reading material in depth when looking at the full text of the document. This is

less than the reported number of times in depth reading occurred during the results list stage. We hypothesise that information seekers when given less information visually, will attempt to read in more depth. In other words, when the participants where returned with mostly headings and an excerpt of the document on the results lists, they were more willing to spend more time reading than when faced with more information within the full text.

There were 18 mentions of images or figures examined. Interestingly, the participants did not mention the images as being important in their triage decisions when it came to textual information needs. 23 searches reported a scanning behaviour over the information provided in order to locate information that was relevant. Participants were often frustrated at the amount of information presented to them. Participant 1 commented on having to "trawl through the rubbish" that was present before finding some useful guidance or information. Other participants commented on being presented information they had not requested, which added to the problems of triaging through the documents. Participant 5 mentioned that is for this reason that he "hates pop ups". Conclusions were not mentioned in the triage activities by any of our participants. We remind the reader however that most triage activities were based on web pages which do not contain conclusions. Looking at the author names was reported twice, references were reported once, while emphasised text was never reported on. On asking the university student participants of their triage activities, they commented that the author does not play an important role in their relevance decisions, but is important in making the decision of whether to triage the document. Hypertext links within the full text of a document was mentioned only two times. References were important in creating further triage instances but not in the actual document decision. Participants only reported twice on reading the plain text of a document in any detail other than the simple scanning behaviour mentioned. Participants during the interview elaborated that they would only read the text if they thought they had found the required information. Again this confirms our theory that further reading also enforces triage decisions and can overturn them, as participants also reported rejecting documents after they had read parts of them in more depth.

The document triage process ended for several different reasons. 150 of the triage activities ended because the participants decided that they had either found the information they needed, or that the information was not there. Interestingly, 15 of the activities ended due to the participants

reporting themselves to be tired of the search, frustrated or bored. 10 of the searches were stopped due to the time limit while the remaining three reported on external factors such as hunger, expense of printing and a "phone ringing". 26 of the triage activities reported participants not retrieving any relevant information, while the remaining 162 activities returned at least some relevant material for the information seekers. The information seekers did extract a portion of information which contributed to their overall search and information need. Participants commented that "at least I know where not to look now" or "got something out of it". 75 of the triage activities reported that there would be no further in depth reading, while 81 reported that there would. The remaining 12 were uncertain.

7.4 Post-Study Interview Results

After the main study, participants were asked a series of questions in a semi-structured interview in order to assess certain features of the document triage process. Some of the questions were ad hoc, and based on the logs they had handed in. These are incorporated in the previous paragraphs. Beyond the standard questions, any other queries regarding inconsistencies of diary studies were also asked. The main questions had three main goals:

1. How long did the participants think they were searching for during the study period?

2. Which medium they preferred performing triage on and why?

3. To list the benefits and disadvantages of the tools they used to perform document triage.

18 out of the 20 participants reported on having a normal schedule in terms of information seeking during their user study period. One participant reported a busier than normal schedule, while the remaining participant commented that this period was not as busy as normal. On average, participants reported a total of 72 minutes of seeking for information per day, without in depth reading. This figure varied greatly with a standard deviation of 67, giving a large spread of searching times. We found it difficult for the participants to distinguish between where the triage started and where the reading began. We again, notice a link between in depth reading and triage instances, testifying to the fact that in depth reading is related and in fact proving to be a valid part of the triage process in itself.

Participants were asked to rate three methods of searching, with regards to their triage useful-ness and capability. In other words, how easy is it to triage the information using one of the three ways and which would you prefer to use. The three available options were a) Electronic b) Physical Libraries and c) Information obtained from another individual while talking face to face.

Unsurprisingly, electronic document triage received a vote for being preferred 82% of the times, making it the prefered means of document triage. Using a physical library space was subjectively prefered 11% of the time while face to face conversations proved the hardest to establish relevance of the suggested materials (7%). Figures 7.2 and 7.3 give us a detailed ANOVA test that rejects the null hypothesis.

Groups	Count	Sum	Average	Variance
Electronic	20	1633	81.65	183.92
Library	20	207	10.35	83.50
Face to Face	20	140	7.00	37.89

Table 7.2: Summary of ANOVA calculations test for the three triage mediums.

Source of Variation	SS	df	MS	F	P-Value	F crit.
Between Groups	71116.9	2.00	35558.45	349.38	1.02^{-32}	3.15
Within Groups	5801.1	57.00	101.77			
Total	76918.00	59.00				

Table 7.3: ANOVA test for the three triage mediums.

When performing a student's t-test between the three groups we find that the electronic mode produces significant differences compared to both the other groups at the 5% level. The library and face to face statistics do not produce a significant result at the 5% level.

In the final part of the interview, participants were asked to report on the positive as well as the negative features of each medium, regarding document triage. All the participants reported the speed of electronic tools during triage as a key benefit of the medium. Fast triage contributed to "effective work when you need some information quickly" but also to the general emotional interaction of the participants. As one of the participants reported, "I hate having to waste time looking for a book in a library when I could get it from the internet as a PDF". Electronic triage was reported by 19 participants as difficult to report on due to the large amount of unreliable in-formation. 6 of the participants were concerned about the amount of information which is returned

after a query is executed. As one participant commented: "...the fact that you can get so much information on-line is a blessing and a curse. You have a lot of information but that means you have to go through it to find things that are not (useless information) [3]".

When commenting on triage with physical media such as print, contrasting advantages and disadvantages were reported. Performing triage in a library situation for example, is reported by most of our participants to be "not as fast" as when being performed on electronic means. Reliability of the information and access to library material was not reported as a problem, but a benefit to the information seekers. Physical constraints were also reported by two participants, such as carrying the books to a table to be able to triage through them all and then returning for more. It is worth mentioning that none of our participants were disabled.

7.5 Discussion and Conclusions

In this chapter we undertook a diary study, to report on the document triage behaviour of a wider group of participants. Although diary studies are accepted as being not reproducible in the way that an experiment is, the aim of the study was to examine real world scenarios and formulate a better understanding of the thoughts, and subjective feedback reported. The greater the amount of participants reporting on an individual behaviour or giving feedback about a specific element, the more common the pattern is and the reliably it holds. We can confirm several of our hypothesese about how actual user behaviours match our lab-based studies and highlight variations.

Before the main study was conducted, a semi-structured interview took place to record subjective feedback from the participants. This feedback is used to elicit requirements as well as possible bias in the experiment. The first finding which correlates not only with the previous studies, but also with generic HCI principles, is that of human error and blame. In other words, it is more natural for the **users to blame themselves for the shortcomings of electronic tools and systems than to identify that the system which they are using is not supportive enough** [114, 149]. This is a valuable contribution to our model formulation from an emotional point of view, similar to Kuhlthau's information seeking model [89]. A user's feelings, as argued by Kuhlthau, can affect a user's seeking behaviour.

---

[3]Actual word substituted.

Barriers to interaction included a perceived lack of knowledge on behalf of the users when they could not locate the information needed. There was also a tendency to triage information with a low satisficing threshold. When asked, participants commented on relying heavily on automatic search systems "to do the work". As one participant put it, "it's just easier to do a better search (reformulate his query) than to look for information myself". Another barrier included too much information to triage.

Confidence levels were mixed at the start of the triage activities. These slowly rise as the triage process continues until the end.

It was impossible to record the **users' exact behaviour, including their visual attention** during the main study. The participants however, reported heavily on some features of the documents as having some influence to their triage behaviour. On average, 6 documents were reported as being opened for in depth viewing. A reported linear approach to viewing structured documents was dominant, while a more unstructured view was reported for web pages and other less structured documents. Reading and skimming behaviour alternated, with skimming being reported as the majority of the time consumed for the location of likely relevant information.

The main title in a results list as well as in a full text document (be it a web page or academic document) was heavily scrutinised. Relevance decisions began to form from the visual extracting of information primarily from this source. Headings and hyper links [4] were used to navigate the user to specific information. This information was considered by the users to possibly contain clues as to the relevance of a document to the information need. Introductions and abstracts were not thoroughly mentioned in the study. When the participants were asked about whether they would read the abstract or introduction, they would answer by simply saying that "websites do not have abstracts". Introductory material on web pages did influence the participants' visual attention to a small degree. Snippets of a web page's text was reported as being skimmed, by only a few of the participants on their searches. These results correlate well with the previous findings, mainly in terms of visual attention and behaviour. We can therefore be more confident in the ecological validity of our lab-based study results. The same features are reported on attracting the attention of an information seeker in situ as a lab-based equivalent study. Items such as the abstract which

---

[4]participants referred to links as headings to other material on the website

is not reported on by every participant. We remind the reader however that several documents triaged in this study did not contain an abstract, as such, for the user to look at. None of the participants reported attempting to use a specific tool for the triage of documents, either on the results lists or within the full text of a document.

In the **post study interview** the main findings which surfaced from the participants' comments was that of speed, reliability and availability of materials. We can hypothesise therefore that a suitable criterion with which to rate potential supportive document triage systems is that of speed. The reliability of the material searched for was also in question for a multitude of the triage activities. Whether fact finding or browsing, the information seeker would often be uncertain about the accuracy of the information he or she received. In this case, if the information was important, the data would be cross-examined with other information from other documents for clarity. Participants reported how the benefit of a physical library over an electronic triage tool is that it is often clear what information is available immediately to them.

In conclusion, the study findings report a positive correlation to the findings in the previous studies which increases the ecological validity of our findings. We also introduce new data that our lab-based experiments were not able to uncover, such as the further reading stage in document triage. These will contribute to a more unified model of triage which we present in the next chapter. By verifying the behaviours of users during the previous few chapters of this thesis and with the diary study reporting on the complete triage behaviour, we can formulate conceptual models of the document triage process (See Chapter 8). The models provide a conceptual framework for understanding the current data from our studies. They will also incorporate related work to include other findings alongside our work. This will contribute to the overall research of document triage and give a foundation upon which further software and study designs can be based.

Chapter 8

Modelling Visual Attention During the Document Triage Process

8.1 Introduction and Motivation

Document triage is largely a human cognitive process and has not been thoroughly researched, hence this process is not yet fully understood. In order to understand the effect that document triage has on information seeking we focus on the part of the information seeking process that document triage influences. Information seekers are reported on as making erroneous decisions on the relevance of documents during triage [28]. By observing the technicalities of their competence through our user testing, we produce a model of the behaviour. When influencing factors, such as document features, are altered the behavioural patterns of the users are also likely to change. If the resulting behaviour produces better relevance ratings, then we will be able to focus on the part of the behaviour that needs to be influenced to produce a more effective triage process. By following the exact steps someone took to achieve a certain goal, others can achieve it too.

Well established models exist that provide frameworks for information seeking (e.g. Gary Marchionini's model [100]). However, there is a need for more detailed scrutiny of more specialised activities within the information seeking process. Bates, critiques large-scale information seeking models [17] and agrees with Kuhn [86] in that "major models that are as central to a field as this one is, eventually begin to show inadequacies as testing leads to greater and greater understanding

167

of the process being studied". Ellis recognises that "there is a need for more micro evaluation of the activities and environments of the users of information systems in order to develop an understanding of the relation of information services to those activities and environments" [51].

In this chapter, we present two models for document triage. These models synthesise the key elements from our findings. To describe these models we constrain their internal workings by talking about sets. We think about triage as an activity where we have sets of documents. By doing this we are able to produce a universally understood mathematical explanation of our workings. This will give other researchers a common framework to compare our work to theirs. Furthermore, by using the theoretical stages in our models, we are able to derive measures (see subsection 8.4) for recall and precision at different levels of triage. These measures allow a more detailed evaluation of users' and software effectiveness during different stages of the triage process. Although these measures are based on our results, they help us to tie in our work with any future experiments relating to our own. These descriptions are a brief aside before we go into the conceptual representations of the models themselves. The two models discussed in this chapter are detailed conceptualisations of parts of the information seeking process. For this reason, we will be referring back to the general information seeking models (See Section 2.2.2) in order to position each stage of our triage models within the seeking process.

8.1.1 Mathematical Definition and Constraints.

Using basic set theory we can logically constrain the boundaries and limits of the document triage process using formal notions. Formal notions such as these help us achieve two purposes. Firstly, our work is directly comparable with further work from other researchers using a common mathematical language rather than simply lexical definitions. Secondly, we can refer back to these notations during our model descriptions to produce a clearer understanding of actions and behaviour and their effects. Figure 8.1 shows a Venn diagram of the possible sets the documents can be located during document triage.

Document triage is a part of the information seeking process. In other words, the elements of the document triage process (P) comprise a proper subset of the information seeking process (I).

$$P \subset I$$

Relations During Document Triage
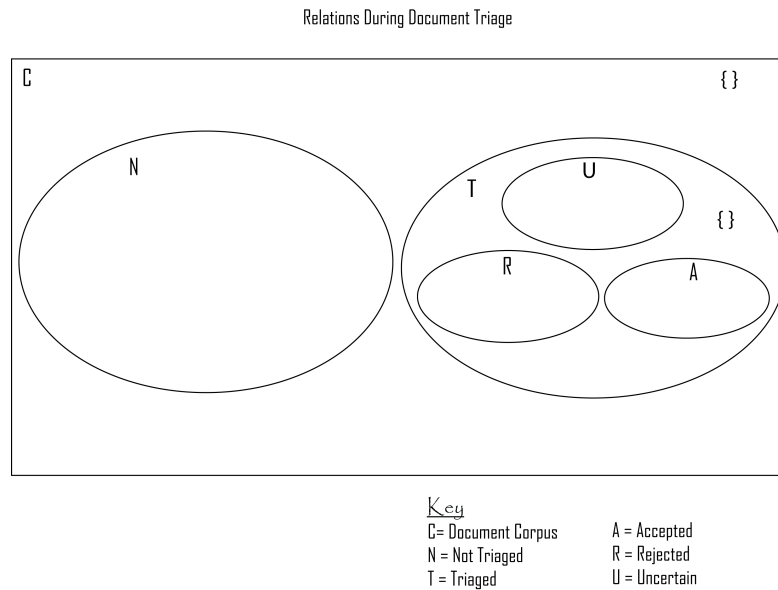
Figure 8.1: Venn Diagram Describing the Possible Status of the Documents in any Given Time of the Document Triage Process.

In order for document triage to begin, a set of potentially relevant documents (C) must exist. This scenario occurs after an automatic information retrieval operation. One practical example of this is a web search engine results list. The resulting web page set are potentially relevant documents. When these are presented to the information seeker, the document triage process initiates. The first stage of the document triage process begins after the "execute query" stage within Marchionini's model [100]. We can also encapsulate the initiation of the triage process after the "Differentiating" stage in Ellis' information seeking model [50].

The documents may either be triaged (T) or **not** triaged (N).

$$C = \{N, T\}$$
$$C = N \cap T$$

At the beginning of the triage process all the document elements (d_i) belong to the **not** triaged set (N). The set (N) is therefore, <u>initially</u> equivalent to the Universal set (V).

$$\forall d_i \in N$$

At the beginning of the triage process, none of the documents are triaged. Therefore, there are no elements in the Accepted, Rejected and Uncertain sets. These sets begin empty.

$$A \cup R \cup U = \emptyset$$

There are three possible outcomes when a document is triaged (t). The information seeker either: deems the document relevant and accepts it (A), does not deem the document relevant to their information need and rejects the document (R), or the relevance of the document remains uncertain without a decision made (U).

$$T = \{A, R, U\}$$

Sets U, R and A can never intersect.

$$U \cap R = \emptyset$$
$$U \cap A = \emptyset$$
$$R \cap A = \emptyset$$

As the triage process progresses, the documents are triaged and a decision is made on each. This stage is equivalent to Marchionini's "Examine Results" and "Extract Info" stages. This is where the majority of the document triage takes place. This process can also be seen in the "extracting" stage of Ellis' model, where an information seeker "selectively identifies relevant material".

$$t : N \mapsto T$$

At *the end* of the triage process all documents are either part of: the accepted set or the rejected set. Even if a document is never reached, or if it is still in the uncertain set, at the end of the triage process it is implicitly rejected.

$$N \cup U = \emptyset$$

We can relate this to Marchionini's model's last stage of "Reflect and Stop". Marchionini's definition reflects the end of the entirety of the search process rather than one set of documents, but does include the documents being triaged. Similarly, Ellis identifies this stage as the "verifying" and "ending" stages, where the seeker will "check the accuracy of the information" and "Tie up loose ends" [51].

During the triage process, information seekers may have a change in their states of knowledge [22], causing documents to move between the accepted, rejected and uncertainty sets. However, if the information need were to change as a result of these changes of knowledge states then the current triage process ends and a new triage process begins (even if the same documents are present).

To date, researchers not had a well defined framework with which to categorise the documents which underwent document triage. Our models, grounded on our experimental empirical data, can produce the foundation for these frameworks. We can begin to identify the relationship between document relevance and the efficiency of document triage during each stage of the process. In other words, how much "effort" or how many actions does it take for an information seeker to move a document from one relevance set to another? Can the time, effort and integrity of the decisions made by the users be justified as to the efficiency of the process as a whole; whether that be time or correct decisions? We can begin to identify ideal sets based on an information need and compare these to actual sets produced by the seekers. Given two different situations, which is more efficient in terms of recall and precision? These are some of the measurements which we will cover during the description of our models.

In our literature review (See Chapter 2), several information seeking models were presented and explained. Each model represents different qualities of the information seeking process. For example, Marchionini's model, provides the reader with a conceptualisation of the procedures undertaken during information seeking. Ellis, provides a behavioural view of the users, and Kuhlthau reports on the cognitive and emotional elements while users are seeking for information [51]. Our two models provide the reader with procedural and behavioural elements, which are closer to the models mentioned rather to focus on cognitive definitions such as Kuhlthau's information seeking model. Our first model, presents document triage from a high level of abstraction and describes the procedures undertaken. Our second model, although borrowing the visual style of Marchionini's model, presents a behavioural representation of triage from our within-document triage studies.

8.2 Document Triage Funnel Model

8.2.1 Overview

In this section, a high level model, dubbed *the funnel model*, is presented, which conceptualizes document triage as a whole (see Figure 8.2). We present the processes by which the user performs document triage and identify the production of artefacts caused by these actions. In order to inform this model, we combine data from related work and our study findings. Our studies in chapters 3 to 6, inform the middle level (Within-document triage). Throughout the studies, but especially in chapter 7 we report on behaviour which testifies to the inclusion of a third stage; namely, in-depth reading to the document triage process. Thus far, in-depth reading was not considered a part of the document triage process. The model contains three levels which information seekers undergo during document triage. We first explain the model and its different levels as well as the horizontal and vertical transitions that occur between these stages. We then refer to the related literature and findings which correspond to each stage. Finally, we complement our model with our set calculations to produce scenarios where the two can be utilised to produce tangible measurements of triage efficiency.

The document triage process begins after a set of documents are presented to an information seeker. The process ends when all these documents are evaluated by the information seekers, regarding their relevance to the information need. Even if documents in a results list are never reached, they are still considered to have been implicitly rejected with regards the specific triage instance. There are three stages at which a document can be evaluated. In Figure 8.2 we can see a model representing the document triage process and the how these stages are connected. The stages of the process are now explained in more detail.

8.2.2 Surrogate Level Triage (Level 1)

The first contact that an information seeker is likely to have with documents, when looking for information, is at a surrogate stage (such as a results list). Often, this stage occurs when information seekers perform a query based search on a search engine. Most of the documents are either rejected or accepted at this stage. The depth of triage a user is willing to undertake when looking for
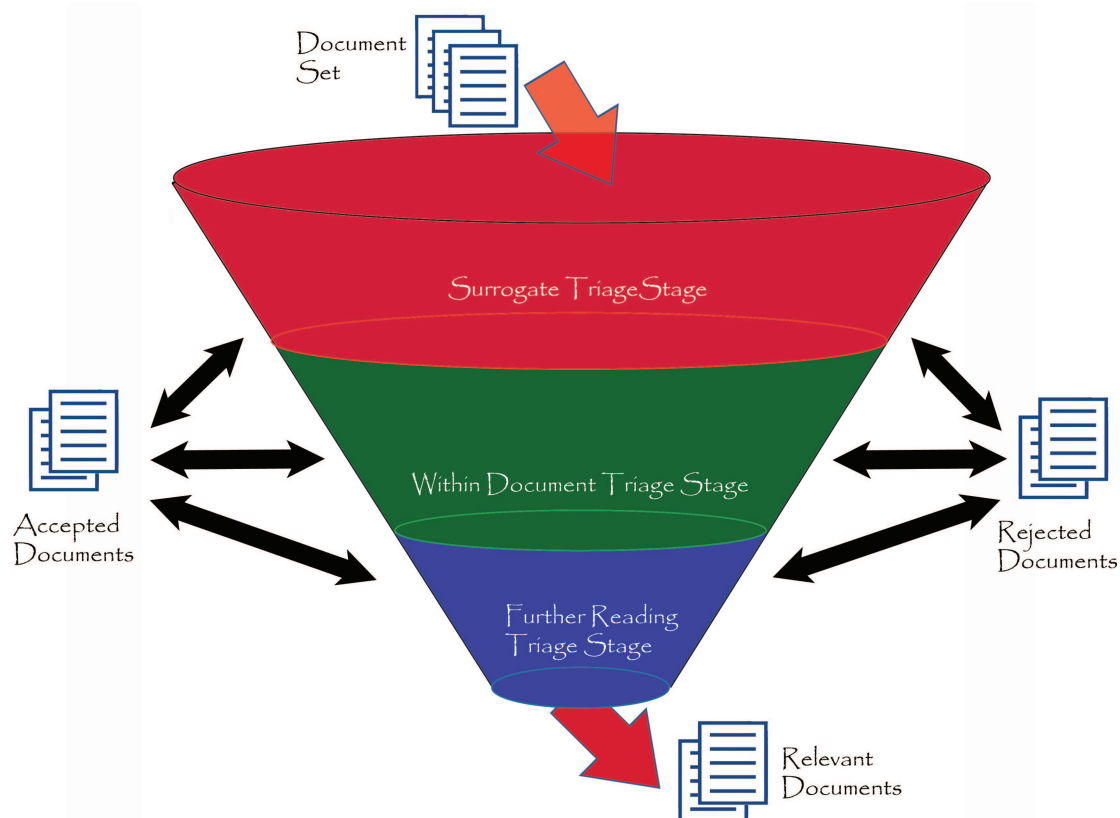
Figure 8.2: High Level Abstraction Document Triage Model

information through a search results list is limited [144]. Therefore, it is common for the majority of documents to be rejected without any viewing. We also know from our previous studies that users are likely to make a relevance judgement on a document without opening the full text [28]. Users mostly have a linear approach to traversing elements at the surrogate level.

8.2.3 Within-Document Triage (Level 2)

The stage of within-document triage begins when the information seeker first comes into contact with the full text of an individual document. During this stage the information seeker is likely to scrutinise some parts of the document at a fast pace. There is limited in-depth reading, with skimming being the dominant behaviour. Users are searching for areas of a document that are likely to communicate the relevance of the document to their specific information need (See Chapters 3 and 4). Factors which determine how extensive the within-document triage process is include: the

complexity of the information need, the obscurity and layout of a document and the information seeker's triage skill.

8.2.4 Further Reading Triage (Level 3)

A stage that is often neglected and not thoroughly researched is that of triage during the further reading stage. When information seekers pass the first two levels of triage (Results List Triage and Within-Document Triage) they are left with the set of documents which are deemed to be relevant. The true relevance of the remaining documents will only likely surface after an in-depth reading process occurs. This is the point where the information seeker will actually extract the relevant information to satisfy the information need. To date, there are no research studies which look into the document triage at the reading stage specifically. A user study is described in Chapter 7 performed which gives us insight into document triage process during the further reading stage.

8.2.5 Transitions

The model in Figure 2 presents the three levels of document triage. Information seekers **transition vertically** between these levels as they progress through their triage process. During these levels, documents are being judged by the information seeker in terms of relevance. These are represented as **horizontal transitions**.

Vertical Transitions

The information seeker will begin from level one, where a set of potentially relevant documents are presented and complete the triage process at the end of level three. The document triage ends with a set or set of documents which are not only accepted but also relevant to the information seeker's need. No further transitions exist after this takes place. During document triage the information seeker will be migrating between these levels, making the transitions bidirectional. These transitions occur when the information seeker advances his or her knowledge state while gathering information [19]. During document triage the information seeker attempts to decide the relevance of a set of documents to his or her information need. As we have reported in the last section, this is generally a fast-paced process. The speed of decision making at each level decreases.

To illustrate, we take the example of a typical hypothetical search task. A user types in a

query into an on-line search engine and is returned with 100,000 results. This is the beginning of the document triage process where the user is required to select documents from the search result list, our first stage in our model. Typically, a seeker is only likely to view the first two pages of a results list [144]. A typical results lists include 10 documents per page and so we can infer that the seeker is only likely to view 20 documents at most. With these factors in mind, even if the information seeker chooses to view the full text of every document on those two pages, then they are implicitly rejecting 99,980 documents (99.98% of the documents). Subsequently, depending on their information need, the query terms and the user's triage experience, only a fraction of these 20 remaining documents will be selected to be read in-depth [42, 49]. This will lead to the second stage in our model, that of within-document triage.

Horizontal Transitions

As information seekers gain knowledge about the perceived contents of a document, they make relevance decisions. Equally, they may have changed their relevance decisions based on the knowledge that they gain through the material they read. As the information seeker gains knowledge into the contents of the documents he formulates a better understanding of his or her information need. Through their triage stages (vertical triage) seekers will judge documents based on their understanding of their need and also based on their evaluation of the document relevance. As the information seeker gradually gains more knowledge, the value of each document being triaged may increase or decrease. The ratings seekers will give to documents are generally tacit in that they do not explicitly rate the documents on a scale but are tentative in nature. If that rating is above a subjective threshold, which is determined by the information seeker unconsciously, then the document is no longer rejected and moves to the **accepted documents set**. $(t : R \mapsto A)$. What contributes to the decision of an information seeker's decision on a cognitive level is not well understood. This set can be physical such as a printout or an actual set, virtual, such as an electronic bookmark, or simply a memo in someone's memory that the document is relevant. There are also documents that may be borderline, in terms of rating, or unclear as to their content. These are usually moved in the uncertainty set (see Figure 8.1) and are recycled through the funnel for further triage at a later stage when more knowledge is accumulated [97]. Documents that are well under the relevant threshold are perceived as irrelevant and moved to the **rejected documents**

set. These horizontal transitions are therefore a complex and important part of the model. They can be used to assess the interactions as well as the relevance rating of users. We have seen the different category sets that the documents can be classified in the previous section (also See Figure 8.1). The model in Figure 8.2 presents a higher level view of the process from start to finish and presents the state of the information seeker rather than the documents.

Horizontal transitions are bidirectional, although they are more commonly outwards away from the funnel. The reason is that it is common for documents to be accepted or rejected and placed on a set, but it is less frequent for an information seeker to re-triage a rejected or accepted document after judgement [28]. It is however, common for the seeker to gain knowledge while performing triage which will give a new perspective on previous documents triaged. When this knowledge affects the relevance threshold, and if the information needed is critical enough, then the information seeker is likely to retrieve documents from the accepted or rejected set and recycle them through part, or all of, the triage process again. Although our participants' feedback during the diary study was clear on this topic, the quantitative data we have is not sufficiently detailed to produce an accurate description of all the transitions in within-document triage. This needs further testing over the three mediums as we focused mainly on within-document triage. Further study is beyond the scope of the work as we focus on visual features and not cognitive evaluation.

8.3 Supporting Evidence

So far, we have presented a procedural model for document triage. Part of the evidence to support the model comes from related work and part from our findings in chapters 3, 4, 5, 6 and 7. We will now step through the triage process focusing on the key literature, rather than an exhaustive list of related work, to represent the primary argument for each stage. We remind the user that our model is built heavily on Marchionini's model [100]. The triage process is undertaken during the "Examine Results" and the "Extract Info" stages. The triage begins when "semantically related surrogates are scanned to identify a more comprehensive relevance assessment". The information is *extracted* when judged to be relevant by the seeker, and may be revised "based on what other documents were added to the relevant list and what events the information seeker experienced since the previous relevance judgement". Finally, the information seeker will decide to "reflect" on

whether the triage process is worth continuing and "stop" if not. In Marchionini's model, we are presented with several cycles between stages. Marchionini's transitions mirror our own transitions in that the user will move to different levels and make further relevance decision on new and already triaged documents. We begin describing our model at the "Examine Results" stage. These results are most often a surrogate view of the documents, such as a results list.

8.3.1 Surrogate Stage

The first document triage level is that of the surrogate stage. This is most commonly a results list view and is the usual point when an information seeker first comes into contact with the document set that needs to be triaged. The first few seconds of an information seekers visual attention are critically important to the selection of a document for within-document triage. The evidence presented by Saracevic's early work [128] implies how decisions are unlikely to change with the addition of further information. Previous work [28] and [97] (see also Chapter 3) gives qualitative feedback from users attesting to the fact that an information seeker's first impressions of a document, may formulate the relevance decision. This is even more so if the information need is fact finding; in other words, a short specific answer is all that is required. An example can be something like: "What is the average life span of a sheep?" It is likely that the first document selected, provided it has the answer, will be accepted and the other documents will be rejected. [144] describes how most information seekers are not likely to view documents located further than the second page when performing triage at the surrogate level. Eye-tracking data show us visual patterns and behaviour that users engage in during a each specific triage process. Granka et al present data supporting the fact that the first result has the highest probability to be selected for further (within-document) triage [58, 119]. They continue to mention how abstracts presented at the surrogate level, ranked one and two, receive almost equal attention and reading time. "After the second link, fixation time drops off sharply". Another interesting observation is that there is a "dip around result item 6/7, both in the viewing time as well as in the number of clicks. Unlike for ranks 2 to 5, the abstracts ranked 6 to 10 receive approximately equal attention". They account this to the "fact that typically only the first 5-6 links were visible without scrolling. Once the user has started scrolling, rank becomes less of an influence for attention". A similar study was undertaken

by Buscher et al, where participants' eye-gaze on contemporary search engine results was recorded [31]. Furthermore, it was hypothesized by Nicholas et al that during the surrogate triage stage, when an abstract web page was accessed before the document itself could be downloaded, users would often not continue to the download stage [113]. This was proved by Buchanan and Loizides when participants, although having the options to open a document and view the full text, chose to perform a relevance decision without proceeding to the within-document stage [98]. The same research shows how information seekers are less likely to re-triage documents even when their knowledge states change.

8.3.2 Within-Document Triage Stage:

In chapters 3 to 7 we undertook a series of experiments looking at triage within a document triage. The main focus of this thesis has been to uncover the processes of within-document triage; a topic that was previously poorly understood. By focusing on this part of the triage process we identify the main contribution of this thesis. We also identify a third level in the triage process which no research has been conducted on.

The second stage (Within-Document Triage) occurs when a seeker has selected a document from the surrogate level to visually scrutinise the document. There can be a prior intent from the user which causes him to select specific documents. The seeker may have an expectation of the positive relevance of a document or the document's relevance may be unclear. Furthermore, the seeker may wish to extract the document in order to dismiss it as non-relevant.

Early work by Saracevic, began to explore the extent to which titles and abstracts effect users' relevance judgements [128]. Saracevic found that relevance decisions were quite consistent when users judged based on a)title alone, b)title and abstract and c) the full text. This lay the foundation to our hypothesis that when information seekers perform triage, they make a cognitive decision early on in their triage, even if they subsequently seem to triage a document further beyond the title and the abstract. If there is no difference between decisions when users are given more information we can infer that there are problems at this level, in that the information given to the users is not being assimilated by the seekers in an effective way. Marshall and Shipman explored user behaviour during document triage [103]. They identified the concept of organizing the documents

to make sense of them, such as organising them into sets. They also began exploring how software can be used to assist users in their triage activity, by allowing the transition between the physical and the digital with the same affordances. A follow up study some 10 years later introduced how activities such as reading, browsing and the use of several pieces of software can affect as well as predict user interest [13]. By using our set theory guidelines recall and precision measurements for within-document triage (see Subsection 8.4), we are now able to begin investigating this further.

Interaction with documents does not necessarily need to be in an electronic mode. It may be the case however, that certain characteristics of the process done physically are adapted to the electronic environment. Buchanan and Loizides investigated the differences and similarities between participants performing document triage using physical and electronic means[98]. Evidence was given to support that users do categorize documents into three or more sets when performing triage in a physical environment, but they are constrained to using either their memory or an electronic note taking piece of software. Our models can be used to represent both physical as well as digital modes.

To date, different types of users' navigational patterns have been reported on [9, 115]. The results produced by those studies include documents that were being read rather than triaged. Some generic results showed how "navigation was found to be irritatingly slow and distracting" and some also correlated with our results in chapter 3 by saying that "assessing document length was difficult to do in any incidental or implicit way" [115]. We are able to report on the navigation and model the behaviour of participants specifically when performing triage. Due to privacy issues, the contents of documents in the above mentioned studies were not disclosed or recorded. It was not until a year later that an empirical study of navigational patterns during the triage process was performed [97]. This study identified four dominant navigational patterns. Using a tailor made document reader, the researchers were able to extract more detailed information of the attention of the information seekers, regarding the document elements present on screen. It was identified again however, that simply looking at the contents on the screen was not detailed enough and a more precise means of capturing users gaze was needed. A subsequent study used eye-tracking equipment to reveal the precise gaze of the information seekers during within-document triage [publication under review]. Interestingly, results showed that the most viewed part of the document was the

abstract, followed by the Conclusion, the Introduction and the Main title[1].

These results were later tested using small screen displays [98]. The same patterns of navigation and attention were also noticed and replicated on small screen displays with the only difference being a faster triage time using small screens. In all the cases of triage during within-document triage, a skimming, rather than reading behaviour is dominant. A representative heat map can be seen in Figure 8.3. The lighter areas represent heavy fixations while the darker areas less attention.
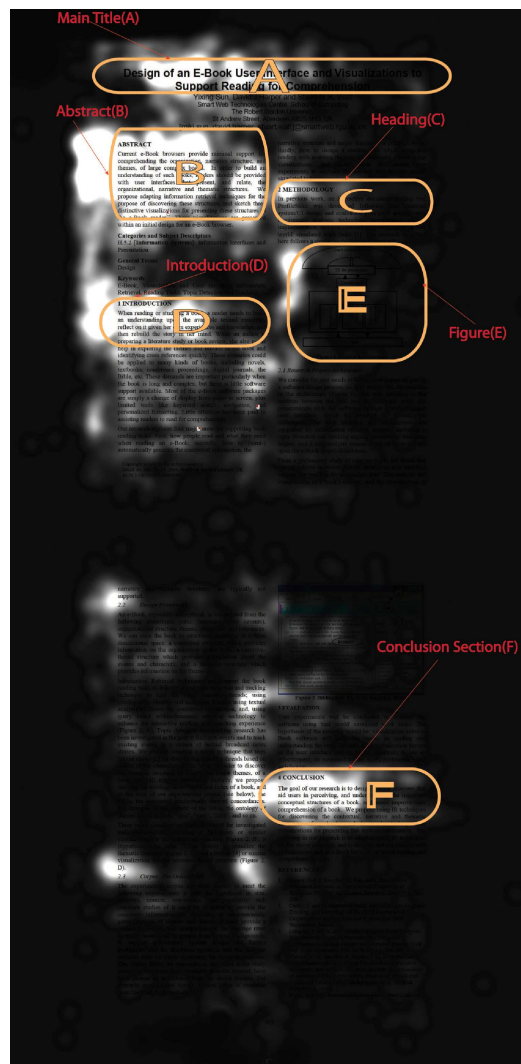
Figure 8.3: Eye-tracking heat map of document elements

Our model encompasses the data from all the studies and presents a typical triage process. The behavioural actions most likely to be undertaken by the users are presented and explained.

---

[1]Calculated by ratio: [Gaze time (%) Area occupied (%)]

Although visual attention time may vary on the document features, our results allow us to model a generic representation of document triage, based on the results that are in agreement throughout our studies. We use it as a guidance to explain the entirety of the document triage process and map the different stages that it can occur. We do not present a detailed analysis of each stage during this thesis.

8.3.3 Further Reading Triage Stage:

Information seekers decide on a perceived relevance for each document early on in their triage process. It is not until the document has been studied in-depth using active reading [7] and some information extracted that the true relevance of the document to the information need is made clear. Upon selecting a document that appears to be relevant (by using either the surrogate triage or the within-document triage methods), the information seeker will then focus to extract the necessary information from the documents selected by more focused reading rather than skimming.

To the best of our knowledge, there is no research into the triage activities that take place during this level. If the research from the work in this thesis were to continue, the further reading stage of triage is an area which would be in need of study. Due to the nature of the observations required, the methodology would defer from the ones presented in this thesis. A more intrusive solution for gathering information would be required. We would suggest a direct observational study of participants in their work surroundings, or by recording the whole search process electronically would be a suitable starting point. These methods, would extract the required information to report on the actual behaviour of users. Care would be required, so as to not bias the participants' behaviour by disrupting his natural environment as little as possible.

8.4 Evolving Information Retrieval Measures

In this chapter, we have presented a model that represents the document triage process and the different stages that the triage can occur. Using our model and the mathematical framework that we have created we can begin to report on measures for each stage of the triage process. Two of these measures are the recall and precision of the documents selected by the information seekers. To help us do this we employ the same standard measures as used in Information Retrieval to evaluate

results from a search engine. In our case however, the processing and selection is not carried out by a search engine, but by users.

We begin to evaluate the process after a set of documents are given to a user. Marchionini describes this stage as the "Execute Query" stage [100]. The subsequent three stages in his model; namely the "Examine Results" and "Extract Info" stages, identify the human element of relevance decision making. In recent years people have been moving to interactive information retrieval and have thought about the users behaviour at the surrogate results list [156]. At each stage in the document triage process a new Recall and Precision score is produced. As information seekers progress through the stages of the triage model, they are likely to reject documents, therefore changing the values for the answer sets and the actual relevant documents within those sets.

The common Recall and Precision scores are identical to the **surrogate triage stage**. This is calculated by:

$$Recall_1 = \frac{|Ra_1|}{|R|} \ , \ Precision_1 = \frac{|Ra_1|}{|A_1|}$$

where $|R|$ = Relevant Documents, $|Ra_1|$ = Relevant Documents in answer set and $|A|$ = Answer Set [15].

The Within-Document Triage stage has a unique Precision and Recall Score (denoted by a subscript number 2):

$$Recall_2 = \frac{|Ra_2|}{|R|} \ , \ Precision_2 = \frac{|Ra_2|}{|A_2|}$$

where $|Ra_2|$ = Relevant documents in within-document triage answer set, $|A_2|$ = New within-document triage answer set.

Similarly, the Further Reading stage also has a unique Precision and Recall score (denoted by a subscript number 3):

$$Recall_3 = \frac{|Ra_3|}{|R|} \ , \ Precision_3 = \frac{|Ra_3|}{|A_3|}$$

where $|Ra_3|$ = Relevant Documents in Further reading stage answer set, $|A_3|$ = New further reading stage answer set.

By synthesising scores between different stages the cumulative efficiency and effectiveness of a tool (or indeed a user) can be measured. Furthermore, by analysing the scores of the same stage between different tools, we can also compare the efficiency of the different tools. We can also bring together all the measurements presented above to create a sum of efficiency across the whole document triage process. We want to optimise this as much as possible as a whole, by breaking apart the different measurements and identifying points of weakness within the triage levels and points of failure. By using the above formulas as a basis we can infer further measurements to rate the transition effectiveness between the three levels of document triage. To illustrate, we take a hypothetical scenario:

An information seeker runs a search query which returns 10,000 results. Out of these 10,000 results only 5,000 are relevant; therefore making precision at the surrogate level 0.5. The information seeker decides to proceed to the within-document triage stage but only selects 8 documents to triage at this level. If we assume that 6 of these documents are relevant then the precision at this level now becomes 0.75. During the transition between the surrogate and the within-triage stage, the information seeker has been influenced by external and internal factors to make the decisions about the documents that will be viewed in full text. These factors usually constitute the interface by which the information is presented. An empirical method for rating these information retrieval interfaces is not available. We can however, infer formulas which can give us a rating for the effectiveness of the transition. For example, if we were to calculate the change between $Precision_1$ and $Precision_2$ we can then assign a score to the interface used. If we were to use the same exact document set, we can then compare and contrast different interface scores between the two levels. A similar approach is very common for assessing information retrieval algorithms in TREC [29].

By using this assessment we can now evaluate users and tools at different parts of the triage process. Our model and mathematical framework ties together to produce measures that can also be used by others to compare users, software tools and methods. We aim for future work to employ these measures to assess software tools that will be informed by our within-document triage model (See next Subsection).

8.5 Within-Document Triage Model

No one model of an activity can capture all the aspects of document triage. In this section, we introduce a specific model of within-document triage. It should be noted here, that this part of the article does not present a detailed presentation of the data extracted from the user studies, but identifies the framework which characterises the process. This model is based on all the data we have reported in chapters 3, 4, 5 and 6.

The model we present here, is similar to Marchionini's information seeking model [100] in that both have a structure which represents processes. Marchionini's model however, focuses on technologies and stages, explaining the actions and responses of information seekers to different artefacts. Most of Marchionini's model, presents the stages through a computer system interacting with users at the surrogate level. We present a model which conceptualises the subsequent stage of within-document interaction through a visual search. We are aware that this stage presents a large proportion of the errors made by the information seekers' relevance decisions. What is not clear however, is where the errors occur within this visual search. Our model is specifically built to represent academic documents, which are semi-structured. Other models may emerge which can be used to present further document formats, such as audio documents or novels.

Firstly, the conditions to *begin* and *end* the document triage process are explained. We will then look at the importance of the *initial page* of a document and the elements contained within this page. Finally, the impact of the features found subsequently on the document and the navigational behaviours of the information seeker is reported on. The within-document triage model can be seen in Figure 8.4.

Figure 8.4: within-document Triage Model

8.5.1   Begin and End

In order for the document triage process to begin, an information seeker must have an information need and receive a set of documents. The information seeker will then go through the documents (usually in linear order (Chapter 3) and try to make a relevance decision between each document with regards to the information need. The process of within-document triage modelled in this chapter begins when the information seeker first comes into contact with the first individual document. Marchionini presents this stage as "examine results and Extract Info". The model ends when a relevance decision is made, and no further scrutiny occurs on the document ("Reflect and Stop" stage in Marchionini's model). This decision may be positive, negative or uncertainty [28, 158].

8.5.2   Initial Page

The *initial page* is considered as an important part in the document triage process. 3 studies in this thesis have demonstrated the critical role of the initial page (see Chapter 3, 4 and 5). In our studies the initial page receives about half of the information seeker's attention when triaging the document. This was also evident from two of the three Groups presented in Chapter 6. Qualitative feedback from the participants in all the studies mentioned how an information seeker's decision is "pretty much made on the first page". Subsequent viewing of the document acts either as a confirmation of the decision or in isolated cases a further scrutiny for the formulation of a decision.

The initial page on an academic document is highly standardized and commonly includes: the main title, abstract, keywords, introduction, and authors. Earlier work by Saracevic indicated how the *abstract* and *main title* play an important part in the information seekers' decision making process [129]. Viewing times, including eye-tracking visual focus (Chapter 4), and qualitative feedback on post study interviews on all study conditions, attest to the visual impact of these features. An information seeker will begin by looking briefly at the main title, continue to triage the abstract extensively and continue to the introduction. While the introduction is rated highly in terms of subjective feedback, the empirical findings indicate that the users do not remain focused on this feature for long. The main title receives most attention on the left hand side, showing that it is not always fully read. The abstract receives the most attention to the first few lines. These lines are the most intensively viewed area of the document during triage. The same limited

focus applies of the *introduction*: although it is not as studied thoroughly as a whole, the first few lines receive a larger proportion of attention than the text that follows them. The remaining features on the initial page such as the authors and keywords do not receive considerable attention. Qualitative feedback from all our studies, give us a more detailed picture regarding the reasons for the behaviour of the information seekers. Although the main title may be looked at, users often realise that it is often not a good indication of the document's contents. Titles and abstracts may be misleading.

8.5.3 Further Document Features

Beyond the initial page the features of a document's pages include: headings, images, tables, diagrams, emphasized text, conclusions section and references. An information seeker will scan rapidly through the main body of the document, alternating from scrolling past the plain text, to fixating on features of potential interest. The average time a page would receive visual attention was just over 5 seconds (See Chapter 3 - not including the initial page). We have clearly distinguished two features that impact the visual attention of the information seeker: namely the *headings* and the *conclusions section*. These were visually the two most important features, with an visual processing ratio[2] of 2 and 4 respectively. When viewing these features, the user mostly looks at the first few lines after the heading; again, similar to the behaviour noted in Section 8.5.2 on the abstract and introduction. When looked at closely, the first few lines of the abstract receive the most attention (calculated by our visual processing ratio) than all the other features (See Chapter 4) Qualitative feedback indicated that users are likely to read text, only if it gives a good indication of the overall document content. Furthermore, our participants commented how these features primarily act as a verification to their relevance decision rather than helping in formulating it. The remainder of the features receive moderate attention and are unlikely to be scrutinised to a great extent as a general principle. Our participants commented that "although sometimes an image might catch my attention it is unlikely to tell me much about the overall meaning of the document". For this reason, the remaining documen features, were not included in our model as unique visual elements where we characterise them cruidly as "other features".

---

[2]Average Area Occupied (%) Average Visual Attention Time (%) - See Chapter 4

8.5.4 Navigation During Document Triage

When performing document triage, users have distinct navigational patterns (See Chapter 3). There are 4 presented in chapter 3. Only three of these patterns are dominant over all of our studies. We will now implement our within-document triage model to demonstrate the visual attention of users employing each of these navigations. We can then conceptualise the visual patterns of our information seekers based on the navigation they choose to undertake. The ability to adapt our model in such a way helps to present different user types.

8.5.4.1 Step Up Pattern

The 'step up' pattern describes a navigation where the initial part of the document is viewed for a prolonged period without any movement and the other parts of the document receive periodic spans of attention, during rapid scrolling through the document. This pattern was the most dominant in our findings, sometimes being used in 92% of the triage instances (See Chapter 5).

Information seekers begin their triage by looking at the first page (usually for around half of their triage time). Their focus initially fixates on the main title. We then see a transition to the abstract, where substantial scrutiny takes place. Although there is a possibility for the user to observe the title again, the next most probable transition is to the introduction. After a brief viewing, the user continues to triage the remainder of the document. The majority of skimming, is over plain text, with within-document features such as the headings and the conclusion, receiving the focus of the user. Once a feature is sufficiently scrutinised (with regards to the user's need), the common skimming behaviour is again, performed and the user returns to superficially viewing plain text. When information seekers use this pattern, the entirety of the within-document triage model can be used to represent their visual attention and navigational transitions.

8.5.4.2 Flatline Pattern

The 'flatline pattern' is a pattern which shows seekers simply scrutinizing the first page of the document. There is no subsequent scrolling to later document content. The information seeker is likely to view the main title, abstract and part of the introduction while using this pattern of triage. In this scenario, the within-document triage model can be adapted to represent the seeker's

behaviour (See Figure 8.5).
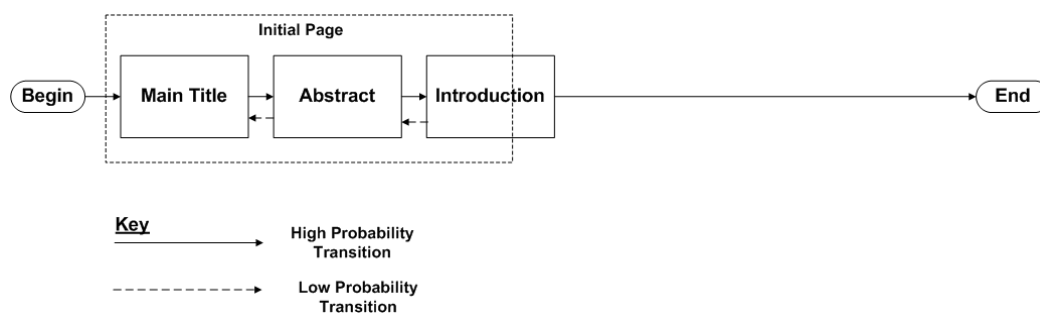
Figure 8.5: within-document Triage Model - adapted for Flatline pattern

8.5.4.3 Begin and End Pattern

A 'begin and end' pattern, sees the seeker scrutinizing the beginning of the document before directly scrolling to the conclusion of the document without any significant pause to viewing other document parts. An information seeker employing this pattern is likely to view the same features as a seeker using the flatline pattern, but also scrutinise the conclusions section before ending his triage. An adapted model for this pattern can be seen in Figure 8.6.
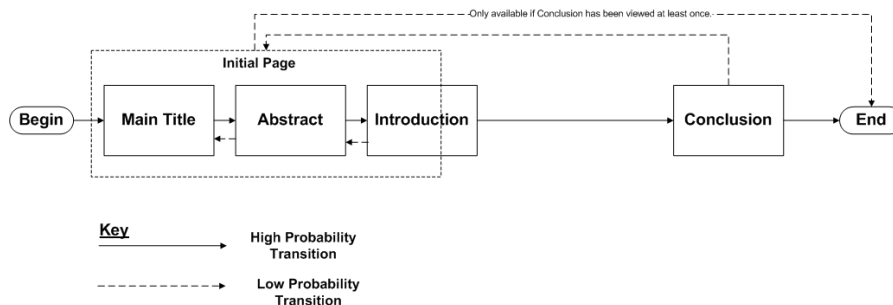
Figure 8.6: within-document Triage Model - adapted for Flatline pattern

8.6 Towards an Evaluation and Design Principles Based on a Document Triage Model

The work presented in this thesis aimed at investigating the visual attention of information seekers during the document triage process. We have recognised a gap in the knowledge in this area and

produce a theoretical foundation and models based on empirical data to address this knowledge need. Although there is still scope for further research, we produced a basic guide of the foundations in order to inform the reader and the researcher into the internal workings of the document triage process based on our findings. It is up to the individual's judgement to use the models in conjunction with their own work so as to benefit them. Like other information seeking models [91], we welcome the evolution of our work by further scrutiny and the expansion of the model by acquisition of new data.

From our model, we can also inform the design of further tests and experiments which achieves two things. The first, is to build software and designs based on what we have learned and the second, is to feed data back from these studies to evolve our model. We briefly demonstrate how we can use our model to facilitate the design of software and test for the validity of the model. The aim of this user study here is not to report fully on the data, but to introduce the reader to the way that our models can be used to inform the design of software and also on how to gain feedback into the workings of the model. Although mostly exploratory, our pilot user study reports encouraging findings and using the feedback given from the participants, we aim to produce further versions of the software and explore the triage process further. We begin by describing the internal workings and design of the software tool. We then describe a study design and finally the results from a preliminary participant group. The software will often be referred to as the "results list tool".

To begin with, we focus on some of the principles we have learned from our data and the initial stages of the within-document triage model; namely, the main title and abstract. We also take into account feedback given to us which states that users would like to have the entirety of the document available. This blends the two first levels of the funnel model (surrogate level and within-document triage level). The results list tool is aimed to "support the rapid assimilation and assessment of new material" [103]. The tool is a manipulation of a results list interface in order to provide added functionality and interactive capability to the information seeker. The modified results list can be seen in Figure 8.8. The results lists were written in Hypertext Mark-up Language (HTML), with Cascading style sheets (CSS) for the formatting and JavaScript code for the expanding functionality.

Each section, separated by a horizontal line, represents a document. The title, amount of pages, authors and headings (elements that have been interacted with visually and subjectively by
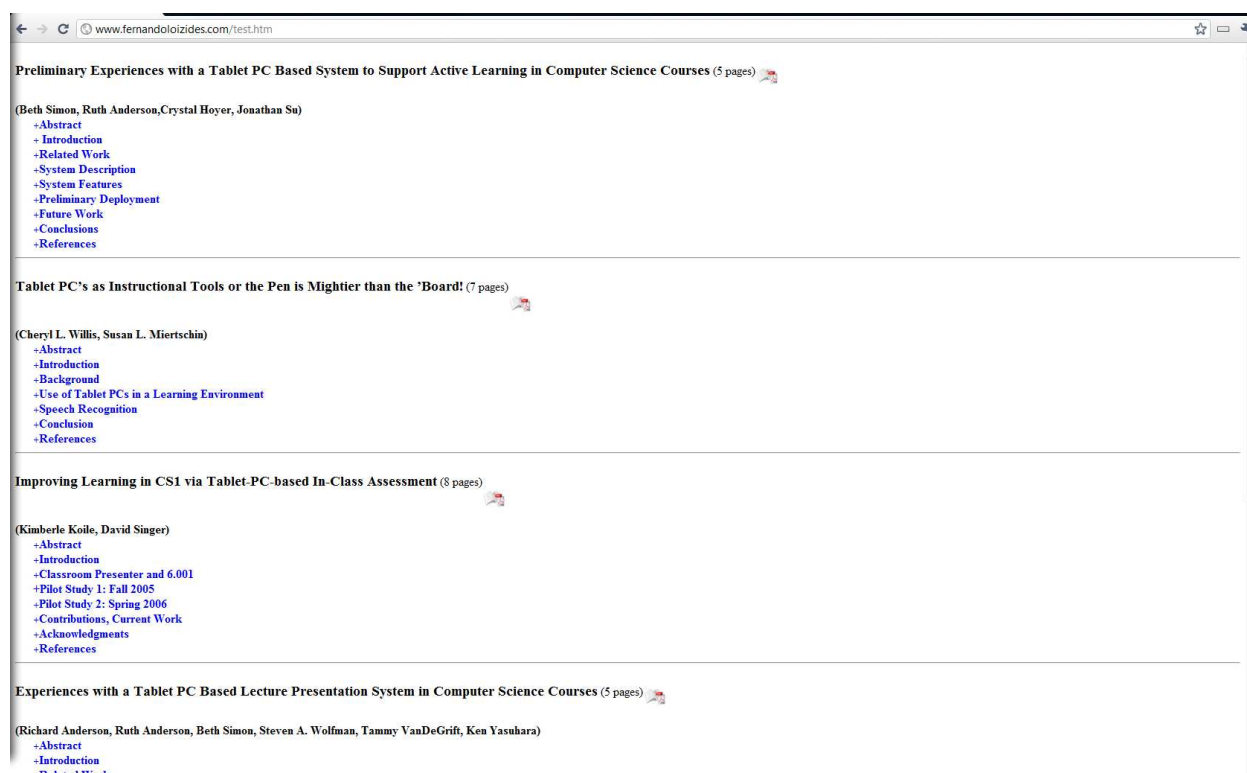
Figure 8.7: Modified Results List - Contracted

our participants) are clearly visible (See Figure 8.7). Each Section heading is clickable. Upon a user clicking one of the section headings, the section will expand and expose the plain text area of the specific section clicked (See Figure 8.8). Multiple sections can be expanded at once. This creates a novel concept which allows a seeker to transition between the 'surrogate level' and the 'within-document triage level'.

An icon of a PDF document is available on the right of the page indication for the information seeker to access the full text via PDF form if they choose to. From the author's related work, we are aware that information seeker subconsciously and cognitively require the full text of a document to be available in a familiar format [117]. Images and emphasised text was not included in the system. These features rated low in visual scrutiny to the participants in our studies and we were interested to see if the users would need to interact with them if they are not made available. The document length, in pages, is presented to the right of the main title. This can inform the information seeker of the amount of information that is possibly located inside the document. We are aware that seekers' relevance decisions can be influenced by the length of the document when
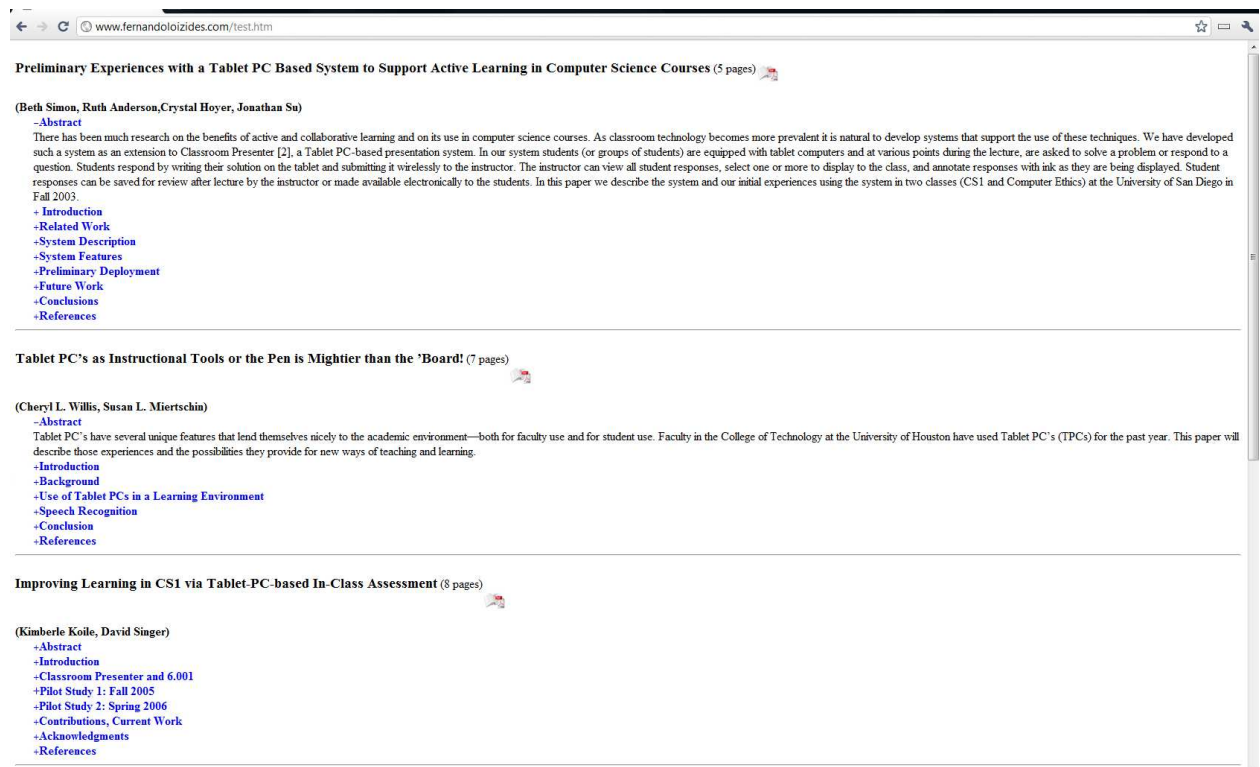
Figure 8.8: Modified Results List - Expanded

making relevance decisions. [28]. Since the results lists will mask the actual document length, we investigate whether having the document length displayed will effect the participants' cognitive judgements of the documents.

Using an eye-tracker we are also able to record the exact eye-gaze of of the participants. By using a similar methodology to our previous studies; namely, giving the same documents to the participants with an information need and asking them to triage. For the purpose of this study only 12 participants were used. The amount of participants (10+) is appropriate to compare statistically with the findings from our other studies. Although reporting on the data in detail is beyond the scope of this chapter, we summarise the main results that help to enforce or inform the validity and variation to our within-document triage model.

Overall, participants took 75 seconds, 20 seconds less than the average participant time in the eye-tracking study (See Chapter 5) and 17 seconds less than the small screen study (See Chapter 6). The predominant pattern followed by users on a desktop or small screen application was that of 'step up' navigation. By using the results lists prototype, users opted to use the 'flatline' approach
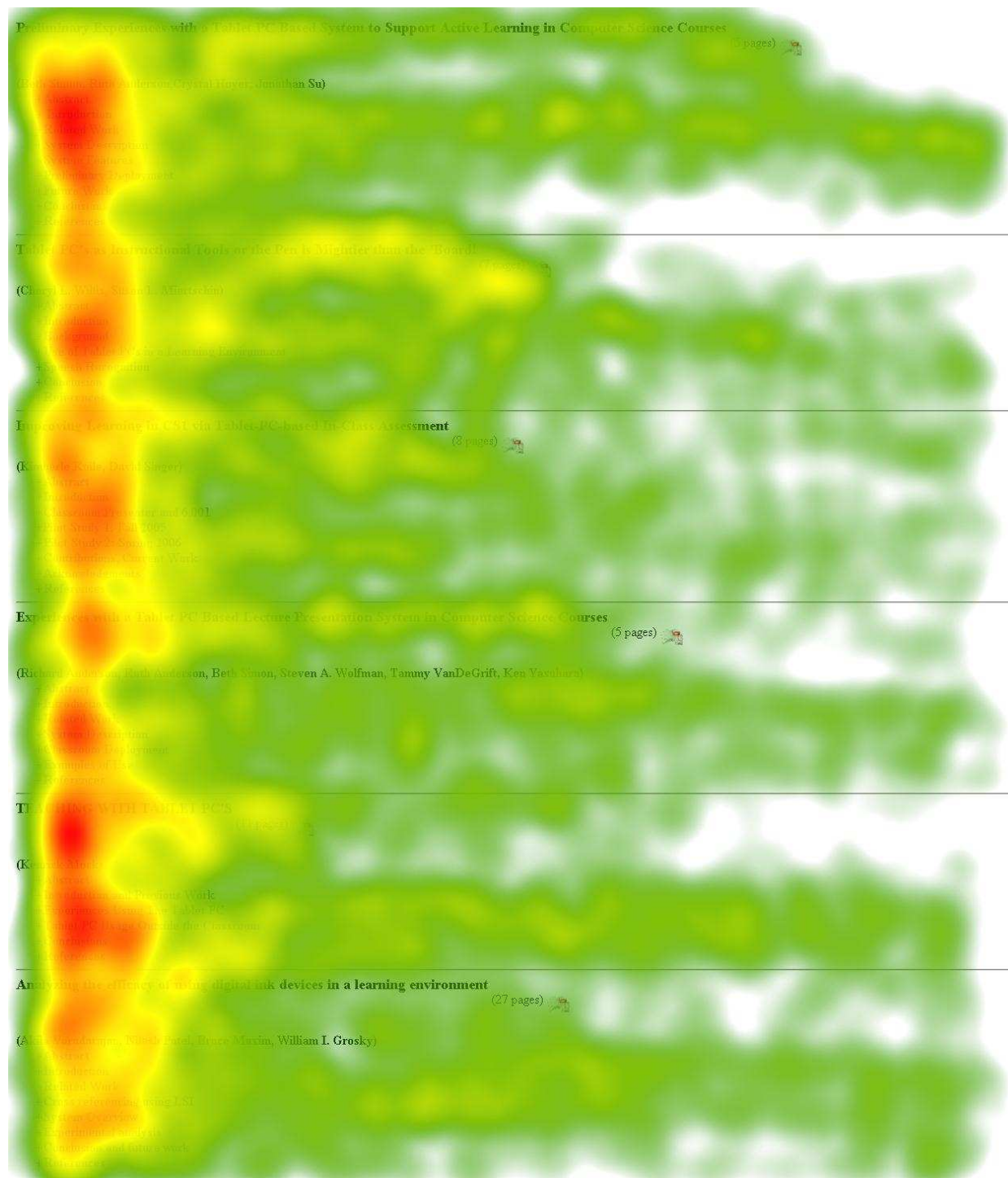
Figure 8.9: Eye-tracking Data from Results List

to triage 61% of the time. The 'step up' pattern was followed 36% of the time, while the 'begin and end' pattern was followed the remaining 3% of the time. Interestingly, although the main

title, abstract, introduction and headings were visually scrutinised heavily, which correlates with our previous findings, the conclusions section rated low with only 18/68 conclusions sections being viewed. We remind the reader that the documents presented had deliberately no emphasised text and no figures or images. This was a decision made in order to test their actual usefulness to the participants, as previous studies did not find evidence to suggest that they were highly scrutinised during document triage. Unsurprisingly, 11 of our participants did not notice that these features were missing. When asked directly if he noticed that images were missing, participant 1 commented: "no, images didn't effect my search actually". A screenshot of the visual attention of the participants can be seen in Figure 8.9. In summary, these preliminary findings can enforce the validity of our models. They also demonstrate how the models can successfully be used to inform the creation of specialised document triage software. We are able to include tests such as these in future work in order to further investigate the effects of altering the document structure regarding the triage behaviour of users.

8.7 Summary

Overall, we produced a global picture of document triage, local models, and a schematic set of measures that can be used to evaluate an entire chain systematically. Furthermore, we know the visual route through the within-document process that an information seeker is likely to take and which variations may occur. These are reported on with a higher level of abstraction in this chapter, in order to conceptualise the triage process. Practical applications of these are also presented throughout the chapter into how our models can be implemented to benefit the information seeking field.

We began this chapter by introducing sets as a framework for containing the documents being triaged. These sets assist us in explaining our models using mathematical rather than lexical notation. These definitions allows us to explain the different states that documents can be in when the user is performing document triage. The actions performed by information seekers, determine the relevance judgements on documents. These documents are then assigned to 'virtual' sets according to their relevance. After identifying and modelling the different stages in triage, we are able to scrutinise the resulting document sets and use our precision and recall measures

together. We can use the sets presented to model optimum relevance decisions. By comparing the difference from these nominal 'perfect' decisions to the users' actual relevance decisions, which are often inconsistent and inaccurate [28], we can measure the effectiveness of the triage process at each of its stages. Measuring triage effectiveness at different levels equips us with the capability to compare software tools and user groups to each other.

Our first model of triage depicts the levels at which document triage can occur. It also models the transitions which can be made to the documents, depending on the users' actions and relevance decisions, which account for the changes in our sets. The first level, or surrogate stage, is the initial stage of the triage process, where the information seeker is presented with a set of documents in a surrogate form. This can be in the form of a results list, where the documents are not presented in full text, but instead are represented with surrogates such as titles and abstracts. Research from the author, as well as other researchers has been undertaken, reporting on user behaviours at this level. We learn that the majority of documents are rejected at this stage and that users request to be able to move into the within-document triage stage and view a document in its entirety, before making a relevance decision. The second level, or within-document triage stage, is the point at which the information seeker will open the full text of the document and triage the contents. This stage has been the main focus of this research and a further model on this stage was also presented later in this chapter. Document features such as the abstract and headings are visually scrutinised by the information seeker before a relevance decision is made. We were able to uncover typical patterns that users undertake in order to visually examine these features and report on variations between groups. The final level, called the further reading stage, represents the point in the information seekers' triage process in which the seeker reads the documents in-depth, after a relevance decision has been made. It is only at this point, that the seeker can decide if a document is truly relevant. To the best of our knowledge, there has been no research on document triage during the further reading stage. Our results have simply verified the fact that triage does take place during this stage. Triage during the third stage is a complex and challenging topic which would warrant further investigation, should the research in this thesis continues. Of course, a completely different methodological approach would be required. Our model can inform the further reading stage in that we are able to report on the effects of the previous two stages. Residual effects from

the initial two stages may be an influencing factor to further reading and perceived document relevance, as well as behaviour.

During all the three stages, seekers can make relevance decisions on documents, either positive or negative, and move documents between our relevance sets. We implemented measures from information retrieval research, which allowed us to merge the automatic information retrieval process of a search engine, with the human processes of triage and produce universal measures between human and computer performance. What was less understood until now, is the visual attention that takes place, during the second level of triage. In other words, where to the information seekers look at during their highly visual triage operations. By answering this question, we are able to also report on which document features also also likely to influence the triage process and which are not. During the thesis, we have produced empirical evidence reporting on information seekers' visual triage patterns. We also showed variation to the model which can be adapted to can take, depending on the users' most common navigational patterns. These adapted patterns produce predicted user actions and behaviours that uncover which document features are not looked at when a user employs a specific navigational behaviour. Finally, we showed how software tools and further comparative studies can be created using guidance from our model. This chapter conceptualises and brings together all our results. It can now be used as a foundation for further work can use these variations to identify the integrity of relevance decisions of seekers based on their behaviour.

Chapter 9

Discussion

9.1 Introduction

The aim of our work was to uncover behavioural and procedural methods employed by information seekers during their document triage activities. The motivation behind this was the limited research in the understanding of this vital process. A gap in the scientific knowledge of within-document triage activities in particular, steered our research towards gaining deeper understanding of this stage. Information seekers make their relevance decisions through the document triage activities. It was thus far not understood how information seekers visually triaged the documents returned from their searches; a process which is mostly unaided by automated software tools. Our software engineering methodological approach gained empirical data rather than rely on subjective self-reported evidence.

The work presented in this thesis covers a range of user studies, investigating user behaviour during a selection of scenarios. These scenarios were chosen specifically in order to form a foundation of knowledge on the document triage process. From our data, patterns which coincide between the studies were identified, ultimately leading to the modelling and conceptualisation of the document triage process. In this chapter, we report on the key findings from all the work presented. We also discuss user behaviour which varies between participant groups and suggest reasons for these. Finally, we hypothesise on future directions for the field of electronic document triage and suggest areas of research which are beneficial for research and development.

The initial section discusses the document triage process through our high level triage model. We also discuss the measurements which were created to evaluate document triage effectiveness. This way we present an overview of the document triage process and a summary of all the stages it encompasses. The chapter continues to discuss our within-document triage findings, beginning with the initial bias participants have towards document features when they begin their triage process. Once this bias is established we continue the chapter by discussing the visual attention of information seekers. This contains the core of our findings and is a substantial piece of this thesis. The findings are compared between the studies, using our within-document triage model. Our chapter then continues to discuss the navigational patterns users employ during triage. Again, we use our within-document triage model to assist us in conceptualising these patterns. Finally, the chapter continues to discuss the relevance decisions between the groups.

9.2 Modelling the Document Triage Process

By investigating individual document triage scenarios (Chapters 3 - 7) we produce evidence which report on users' activities to a fine level of granularity. Our main aim however, is to conceptualise the document triage process, firstly in its entirety and specifically looking at the within-document triage stage. We do this by comparing our studies and identifying similarities and differences between user behaviour and procedures. By creating a conceptual framework, we can scientifically make assumptions about the users actions and analyse the document triage process. We can also predict likely variations that will occur in information seekers' behaviour from variations to their triage process. Finally, we can use these models to inform the design and implementation of software tools to assist with document triage.

Chapter 8 presented two models that represented the document triage process. In this section we focus on the procedural document triage model, dubbed the "funnel model" (See Figure 8.2). This model conceptualises the entire triage process and sets the boundaries of where document triage begins and ends. Having a clear boundary set allows us to distinguish the triage behaviour from other stages in the information seeking process. We identify, for example, that the triage process begins after the "Execute Query stage" of Marchionini's information seeking model [100].

Before undertaking the research in this thesis, there was existing evidence on how document

triage occurs on a surrogate level (usually a results list) [28]. Our work investigated within-document triage, where a highly visual interaction occurs for the seeker to make sense of the information within a document (See Chapters 3, 4, 5 and 6). These first two stages, are represented in the model as the "Surrogate Triage Stage" and the "Within-Document Triage Stage". Upon scrutiny of participants' behaviours, outside of a lab-based study (see Chapter 7) we were able to identify and add a third stage of triage. This is dubbed the "Further Reading Stage". and occurs when the documents are selected for in depth reading after they have been found to be, initially, relevant by the users. It is an indirect stage of triage, in that users do not actively engage in a triage activity to assess relevance. If however, upon reading the document further, the reader finds little or no relevance then the document is rejected. The sets in which a document can belong to during the triage process, such as 'relevant' and 'not relevant', can be seen in Figure 8.1. These sets are populated during all the levels of the triage process and are represented and explained by the transitions on our funnel model.

By modelling these transitions and having used set theory to represent them, we can mathematically produce measures which can compare and rate the relevance decisions of users. These measures expand the already existing information retrieval measures, giving us a way to measure human and computer interaction performance subsequent to the automatic information retrieval stage. Using these measurements we can assess how the document triage process is performing at any given level of the information seekers' search. Where the measurements rate the triage process poorly, we can focus on . This method of evaluating the human-computer interaction process of document triage, was thus far not possible without using traditional user study techniques.

9.2.1 Contribution to Literature

Ellis observes the need "for more micro evaluation of the activities and environments of the users of information systems in order to develop an understanding of the relation of information services to those activities and environments" [51]. So far, the information seeking procedure, has been documented and modelled extensively in its entirety. Ellis presents an information seeking behavioural model which was refined in several versions (See Section 2.2.2.3). Ellis touches upon all the three levels presented in our funnel model. 'Chaining' is described as a process of looking

through sources recommended from an initial source, such as the ones described at our surrogate level. This surrogate stage, which is often a results list, has been identified and looked at more closely by other researchers[60]. Aula et al identify different user types for searching through results lists and associate different users to having different behaviours [10]. Cutrel and Guan enhanced this research by exploring "the effects of changes in the presentation of search results" [42].

In his model, Ellis describes 'Browsing' and 'Differentiating' as the search in areas of potential interest, such as headings and abstracts, and the selection of relevant information. This is the equivalent to our second stage in our funnel model. Interestingly, users which are characterised as 'exhaustive' are more likely to proceed to scrutinising information, as they require more to make a relevance judgement [10]. By using our model we can identify this further scrutiny to occur as a separate (second) level in a seeker's document triage process. This stage in the process is explained in detail later in this section.

Our further reading stage (level 3 in our funnel model), is not discussed or researched thoroughly in any literature we are aware of. Models such as Ellis's behavioural model describe possible scenarios where this stage exists such as 'Monitoring' [51]). Due to Ellis' Model not having a stage layout, the features described can be applied at any stage of the information seeking process. Using our funnel model, we can explain the behaviour and where each occurs thus producing a structure which conceptualises the order and procure in which the document triage process can occur. Marchionini explains the 'reflect and stop' stage as ending where 'motivation' stops. The information seeker will ultimately stop examining the document only when he or she has finished reading it and has either extracted enough information to satisfy the information need (satisficing [47]). Kuhlthau's 6th stage is the conclusion of search information [90]. We argue that while the user still has the opportunity to change his relevance decisions on material (a process which was identified in Chapter 7) then reading is within the triage and information seeking domain. From these examples, we can identify how literature on reading activity is important to relevance decisions and so with our model, extend the information seeking and specifically the triage process to include reading. As Adler et al report, people can 'read to search for particular information' [6]. This is a typical 'passive' way of searching for information, as validated by Bates [17]. We in turn, model how relevance decisions can be made at all three levels of triage, including that of further

Chapter 3	Chapter 4	Chapter 5	Chapter 6
Main Title	Main Title	Main Title	Main Title
Headings	Abstract	Abstract	Abstract
Abstract	Headings	Introduction-Conclusions	Headings
Introduction	Conclusions	Headings	Conclusions
Emph. Text	Introduction	Figures	Introduction
Conclusions	Emph. Text	Emph. Text	Emph. Text
Plain. Text	Plain Text	Plain Text	Plain Text
Figures	Figures	References	Figures
-	References	-	-

Table 9.1: Ranked Subjective ratings from our Studies in Chapters 3 to 6

reading.

9.3 Subjective Ratings for Document Features

Information seekers will develop individual browsing strategies based on their experiences [100]. This information seeking behaviour is heavily influenced by cognitive factors, such as user perceptions of objects and natural biases as to the importance of document features. An example of one such bias is found in the way that user may choose to rate longer documents using a negative constant in their subjective perception of its relevance [28]. It is important then, to identify the subjective importance of the individual features of a document to the information seekers. The important features rated high by users, are hypothesised to be highly influential in a seeker's relevance decisions as well as gaining favourable visual attention during triage. We have reported on earlier work which gives evidence on how certain features are highly influential on information seekers' attention and bias the relevance decision processes [128]. We therefore need to know the features of documents which are cognitively important to the information seekers and can effect their visual patterns.

During all the lab-based studies presented in this thesis, a subjective rating was required of the participants before engaging in the information seeking task. These ratings were then compared and contrasted to the actual time and fixations the participants spent engaged with the same features. Table 9.1 shows the subjective ratings in order of perceived importance. Each column represents one study.

From the consensus, the 5 features which are subjectively considered most important for triage

are:

1. Main Title
2. Abstract
3. Headings
4. Introduction
5. Conclusion

These are the features that are likely to have an initial bias on the personal searching strategies [100] of an information seeker during their triage activities. They are perceived to have an important visual impact during within-document triage. It is interesting to note, due to our findings on the importance of the initial page (Chapters 3, 4, 5 and 6) that three out of the five items mentioned, are usually located on the initial page of a document. The Headings are usually spread throughout a document, and the conclusions sections is located close to the end.

During one of the studies (Chapter 5) presented in this thesis, the participants were asked to rank the features again, subsequent to their triage process. The order of subjective post study interview ratings were:

1. Main Title
2. Abstract
3. Headings
4. Figures
5. Conclusions
6. Emph. Text
7. Introduction
8. Plain Text
9. References

The interesting variation in these ratings from the previous ones is the drop in the partici-pants' views of the introduction. The remaining four features retain their subjective importance throughout the triage process. We hypothesised whether this change was a side effect of the small screen space provided in the study. Verbal feedback from the participants in the post-study inter-view however, indicated that the reason behind the low confidence in the introduction, was "the introduction not having enough useful information".

Although subjective feedback indicated that the priorities of the order of importance of the features has changed, the empirical data suggests otherwise. This is a content effect that we have previously reported on [28]. Our studies report that information seekers' visual attention is not focused on the figures of a document as much as they claim. The conclusions section however always receives a large amount of visual attention. Further details on the visual attention of information seekers during triage is discussed in the next section.

We are not aware if this subjective ratings bare any relevance to the relevance decision making process in terms of relevance decisions, but we hypothesise that the visual behaviour of seekers while triaging is guided by these criteria. Wang et al reported on "user criteria and their relationship with document values[158]. Wang identified some document features, such as "recognition of the author", that are important to seekers in order to assess the document relevance. Their study, like our subjective data, was based on participants' feedback rather than visual data. Wang et al identify some document information elements (DIEs) which are reported as important to the document selection process. These are reported after the triage activity is undertaken however and not before. We complement their findings by reporting on the data extracted from before the search activity occurs.

9.4 Within-Document Triage Behaviour

In the previous section, we looked at how users rate the features within-documents subjectively as to their importance in the document triage process. Our previous work suggested that what users say they do and what they actually do, is not always the same [96]. Previous literature has also reported on the impact of some of these features to relevance decisions. For example, Saracevic began reporting on the effects of titles and the abstract on the relevance decisions. His work suggested that the main title and abstract inform a seeker into making a relevance decision, more than any other feature.

There has thus far been limited research into the actual visual attention and processes of information seekers performing within-document triage; the second stage of triage process. Most of the findings which we presented in our related work chapter (Chapter 2), was taken from subjective feedback rather than empirical quantitative findings. Four of our studies presented in this thesis

Chapter 3	Chapter 4	Chapter 5	Chapter 6(30)	Chapter 6(60)	Chapter 6(90)
Initial Page	Abstract	Abstract	Abstract	Abstract	Abstract
Headings	Conclusions	Main Title	Main Title	Main Title	Main Title
Conclusions	Introduction	Conclusions	Headings	Headings	Conclusions
Images	Main Title	Introduction	Diagrams	Conclusions	Headings
Emph. Text	Headings	Headings	Conclusions	Introduction	Introduction
Diagrams	Emph. Text	Emph. Text	Introduction	Diagrams	Diagrams
-	Diagrams	Figures	Emph text	Emph. Text	Emph. Text
-	Images	-	Images	Images	Images

Table 9.2: Feature Importance for studies in Chapter 3: Empirical Desktop Triage Study, Chapter 4: Eye Tracking Triage Study, Chapter 5: Small Screen Triage Study and Chapter 6: Fast triage. (Number in brackets denotes the time group)

(Chapters 3 - 6), focused on the study of the visual attention of information seekers during within-document triage. The main aim was to identify the visual patterns associated with information seekers during their scrutiny of the full text of a document, including features such as headings, introductions and conclusions; features which were previously unreported on. Table 9.2 shows the within-document features, ranked by their visual importance, as presented in the chapters. We will now report on the key findings and compare the different study groups.

9.4.1   Initial Page

A recurring finding that is observed systematically throughout all the studies in this thesis, is that of the importance of the **initial page** of a document. The visual attention devoted to the initial page of a document surpasses the attention given to any other page of the document. The studies found in chapters 3, 4 and 5 (apart from group 60) all concur with this finding. To give an example, in our initial study it was found that the initial page showed an average of 23.21 seconds of visual attention, a contrast with the average for all other pages of 5.28s. This finding coincides with earlier related work which also recorded the same initial page importance [28]. The one exception to this rule was found to be one of the group of participants which were required to triage documents with a limited time constraint. This group was allowed 60 seconds to triage each document and devoted an average of 27% of their time to the initial page. There was no statistically significant reason for this behaviour. We hypothesise that this finding may either be an outlier, or the discrepancy is an artefact of the specific time limit given.

It is not surprising that the initial page receives the majority of the attention. The abstract is the most viewed feature during all our studies. Other features which rate highly in terms of visual attention are the main title, and the introduction, all of which are found on the initial page of a document. The decision making process which is affected by the initial, fast interaction with documents is recorded in by Spink [144] and Saracevic [128]. Subjective feedback from the majority of the participants also recorded how decisions are usually made within the first few seconds of scrutinising a document and therefore attest to the importance of the initial page. The subsequent reading and searching behaviour constitutes a verification role, rather than a formulation of the information seeker's decision.

9.4.2 Further Document Features

We have seen the importance of the initial page features. These features however, are not the only ones which receive attention during triage. From our data, we identify two other features which are consistently highly rated with regards to their importance ratio [1]. These are the headings and the conclusions section. Typically, an information seeker will skim the full text of a document and identify headings which may give the seeker information as to the documents' content (Chapter 3). Headings which are considered potentially relevant by the information seeker, will receive in depth scrutiny on the first few lines of plain text (See Chapter 4). Although the time taken to triage all the important features, as presented above, makes up for some of the seeker's triage time, there is still a large proportion of the triage time of which visual attention is not focused on these elements but is focused on the remaining document features.

Further features on a document include *figures*, such as images and tables, *emphasized text*, and *plain text*. Although these features, receive some attention from the information seekers, they do not rate highly (using our visual processing ratio measure). Plain text receives a large proportion of the information seekers time, but the amount of focus does not allow the seeker to scrutinise the information in detail. The information seeker simply skims the text looking for the above mentioned 'important features', such as headings, rather than fixate on the actual plain text and extract information from it (See Chapters 3, 4 and 5). An exception to this rule are the few lines

---

[1]See Chapter 4 for an explanation of importance ratio

of plain text under a heading deemed relevant to the triage exercise by the seeker (See Chapter 4).

9.4.3   Contribution

Ellis, produces a behavioural model with features to explain the users actions during the information search. An overview of this model can be found in section 2.2.2.3. In his 'Browsing' feature, Ellis explains how a seeker will simplify browsing by looking through features of the document, such as tables of contents, lists of titles, subject headings, abstracts and so on. There is no detailed analysis however, of which of these features are visually observed and to what extent. Marchionini also encapsulates the within-document triage process in his 'Extract information' stage of his information seeing model[100]. Marchionini describes how the user will scan a document for relevance but does not elaborate further on what features of the document are scanned. We can now contribute data to support the extraction stage of Marchionini's model.

Granka et al present research which gives "insight into how users browse the presented abstracts" [59]. Grank et al identify the need for a more detailed assessment of a document beyond the mere surrogate level but do not report on other features which are scrutinised. Saracevic also reported on the main title and abstract of a document as being a deciding feature for the relevance judgements of documents. We can confirm that visually, these two features are important to the seekers and also extend the knowledge on which parts of these features specifically, are scrutinised. An example of this is our findings on the first few lines of the abstract, which are focused on more than any other part of the document. Furthermore, we report on further document features with objective data, to verify subjective reports of important document features [158].

Literature reporting indirectly on the document triage process, has mostly been evidenced on self-reported data. We give empirical observational data, some of which originate from eye-tracking results, to model the exact behaviour of users. A cognitive/behavioural model which resembles our within-document triage model is presented by Andrew Dillon [43]. Dillon presents the triage process under the category of reading. Since the users will 'read' while searching for information this term is technically true but it does not capture the intent or reading style. The term information seeking or document triage are more accurate and well established ways of explaining this particular reading behaviour. Dillon identifies several of the within-document elements which will influence

an information seeker's decision making process.

He writes: "Graphics in articles (for example,tables, figures, etc)" were generally viewed positively as seven readers explicitly stated a dislike for academic articles that consisted of pages of straight text. Our previous work, confirms that images are considered subjectively as useful to the triage process [28]. Our in depth scrutiny of the visual patterns in the studies in this thesis however, can verify that, although figures and images can attract the attention of information seekers during triage, they are not as informative as one may speculate (See Table 9.3 and Chapters 3 - 6).

Dillon also reports on document (academic journal) size as a 'disincentive' to the reading process. Although we can confirm that users do prefer to read documents which are shorter [28], we also report further findings which evidence that the bias to length can often be overcome when the document is identified as very relevant to the information need and if the information seeker is limited in terms of time (Chapter 6)[2].

Dillon continues to describe the effect of features within a document to the readers' relevance decision making. On page 110 [43] a "generic model for journal usage is presented". The model describes the process as follows (simplified):

1. Skim Titles and authors.
2. Scan abstract and main sections.
3. Non serial read of main section.
4. Detailed serial read.

Our model, can confirm parts of the reported behaviours, add detail to features which are not mentioned in Dillon's model and, evidence variations to this model from our studies. Our participants (Chapter 4) read the main title, authors and abstract, like Dillon's model describes. We can also extend the model to include that the abstract's first few lines are the ones that are heavily scrutinised. The model presented does not include the introduction, although Dillon comments on how "the start of the introduction" is also reported as being read. We can empirically, using quantitative evidence, confirm that the introduction is viewed on by information seekers, but again, they focus on the first few lines of the text (Chapters 4 and 5). The section headings are also reported as viewed by Dillon. We can complement his findings by adding that not all the

---

[2]Here we refer to the relevance judgement and not the actual further reading process which may account for the difference in Dillon's statement with our findings.

headings are viewed. Furthermore, any headings thought to be 'relevant' to the seeker is likely to be further scrutinised for a few lines (Chapter 4). The conclusions section is also reported by Dillon as "a common method of extracting central ideas from the article and deciding on its relevance". We can confirm that the conclusion is viewed in a triage operation, on a number of occasions (See navigational patterns in Chapter 3 and eye-tracking evidence in Chapter 4). Furthermore, the first few lines capture the visual attention. We can also report on how our qualitative feedback from all our studies suggests that the conclusions verify the relevance decisions rather than formulate them.

9.4.4   Summary

In this section, we reported on the visual behaviour of our participants, during our four studies. We identified features within a document that receive increased visual attention, such as the abstract, headings and conclusions section. We also identified features such as images and figures which do not rate highly in terms of visual attention by our participants. Our contribution to the related literature was discussed, as we contribute to models and self-reported studies with quantitative empirical data. We also complement and expand existing work to produce more in depth and accurate results on users' actual behaviours.

9.5   Within-Document Triage Model and Navigational Patterns

Earlier in this chapter, we discussed our *funnel model* of document triage. This model identified three stages to the document triage process. The initial stage, dubbed the surrogate stage, has been previously researched in terms of visual attention and user behaviour [144] (For a detailed related literature list, see Chapter 2). We also have data from related literature which give us limited insight into how users perform triage on the second stage (within-document triage). This has so far been subjective in nature and limited in scope [128]. By identifying this knowledge gap, we enhanced the scientific research in this area by focusing four of our studies directly on within-document triage. By comparing the data from our user studies, we formulated a model to represent and conceptualise the information seeker's behaviour during within-document triage (See Figure 8.4). We present the navigational actions and visual cues which are likely to receive the attention of the seeker during triage. The document triage process is highly visual and driven

Chapter 3	Chapter 4	Chapter 5	Chapter 6
Step Up	Step Up	Step Up	Step Up
Flatline	Flatline	Flatline	Flatline
Begin End Page	Begin End	Begin End	Begin End
Mountain	Mountian	Mountain	Mountain

Table 9.3: Feature Importance for studies in Chapters 3- 6. (Number in brackets denotes the time group)

by the features presented to the user. We do not attempt to provide a cognitive explanation of the seekers' relevance decisions, but we can provide evidence of which document parts effect the decision making process. By dissecting the within-document scrutiny into individual stages, future research can then proceed to analyse each part separately and using our measurements found in Sections 8.1.1 and 8.4, isolate areas which hinder or assist the document triage process.

The model begins with the seeker triaging the features in the initial page of a document. These were reported as rating high in visual importance (See Chapters 3, 4 and 5 and 6) with a high probability of visual focus by the seekers. Beyond this point, the seeker is likely to skim past the plain text looking for headings or the conclusions section, occasionally stopping at further document features such as figures. The within-document triage process is recognised to be iterative, based on the seekers' reflections, similar to Marchionini's information seeking process model [100]. The triage process ends when the seeker reformulates his query and a new set of documents is presented.

Using this model we have formalized current understanding the internal workings, scope and boundaries of within-document triage. We have contributed an effective way to communicate users' actions when interacting with documents within a specialised sub process of information seeking; that of document triage. We can proceed to predict information seekers' behaviour based on his model. One example of this is the integration of our navigational patterns found in Section 3.4.2. We have identified the most common navigations by the participants during our studies. Table 9.3 shows the most common navigational behaviours in order of the most common for all our studies.

Clearly, the 'step up' navigation is the most popular. This navigation occurs on average, more than 70% of within-triage instances. The 'Mountain effect' navigation is uncommon among all our studies. Both these navigations however share one thing in common. Informtion seekers will follow all the stages found in our within-document model. Seekers will therefore, view beyond the initial page and will likely view headers and the conclusions section. Which of these features the

information seekers' use cognitively to inform their relevance decisions, is not yet clearly understood.

Information seekers who use the 'flatline' or the 'begin and end' navigational patterns will not scrutinise the main document area and will not benefit from images, headings and other features located in the centre of the document. Future work looking at the effects of these navigational patterns on relevance decisions can help uncover the benefits or disadvantages of each. One such example, was found in the previous chapter, when our model helped inform a prototype software tool. Although the software is still under development, we are able to demonstrate how our findings can act as guidelines to further explore document triage software design.

9.5.1   Contributions

We have previously reported on two models which directly relate to the document triage process (See Sections 2.5.2 and 9.4.2). Our findings complement these models by providing quantitative evidence to support the users' navigational and visual patterns. Wang et al reported on document information elements (DIEs) which were reported by participants to be important in their decision making process [158]. We can verify the features which are visually scrutinised, rather than reported to be important and describe the effects of each of the DIEs to the information seeker's triage path. Wang followed up the model and reported that "ten of the eleven criteria of the original study are used in reading and citing decisions, but four of these: novelty, expected quality, reading time, and availability, were not used in citing decisions" [159]. The focus of this research differed from our investigation in that document features were not reported on at this stage. Personal correspondence with Wang, uncovered that the data has been since abandoned and no further research was to be carried out. Unfortunately, the data was not available for ethical consent reasons but will be available at a future date.

The second behavioural/procedural model presented in our work was that of Andrew Dillon [43]. His model reports on "three levels of reading" which we classify as triage. The model, although valid in certain circumstances, does not encapsulate the entirety of the behaviours of information seekers during triage. For example, Dillon states that the second stage of triage is a "non serial scan of major sections". From our studies and navigational patterns reported, we are aware that information seekers will most likely follow a linear approach to triage (See 'Step Up' navigational

pattern in Chapter 3). Dillon's model does not identify transitions related to decision making at all the stages of the model. A user will either quit, file a document or read it in depth only after reading major sections of the document. We can complement the findings reported by Dillon and introduce our navigational patterns to represent a relevance decision making process during any point of the document triage process. Furthermore, we report on features such as conclusions and the first few lines after relevant headings which acquire a substantial amount of reading by a seeker.

9.6 Ecological and External Validity

The majority of studies presented in the thesis, are lab-based and controlled (See Chapters 3 - 6 and Example study in Chapter 9). In this section, questions of ecological and external validity are addressed with regards to our studies. External validity refers to the generalisability of inferences made from experiment results [108]. Ecological validity similarly dictates the validity of the results obtained to the 'real world' equivalent situations [24]. Our studies are aimed at investigating the visual attention of information seekers during document triage on semi-structured documents; in this case academic documents. From the results we inform models to produce a foundation of the process. In order to replicate the results or evolve the models we overview the ecological and external validity of our results.

In order to address the issue of ecological and external validity, the experimental methodologies employed were based on existing methodological approaches as carried out by authors of previous literature on document triage. The methodology of giving participants similar documents in different scenarios and with different screen sizes, for example, followed closely to the methodology employed by Bae et al [14]. In their experiment however, small screens were not investigated due to them not being available at the time of publication. Wang et al [13] presented documents to participants and asked them to report the areas they would use in order to make judgements with regards to an information need. Using the data, they formed a model of document selection. However, Wang et al were not able to identify detailed results as to the actual validity of these finding; findings which were self-reported. By employing the same methodology and an eye-tracker (See Chapter 4) and specialised document reader software (See Chapter 3) we were able to a) verify the results by Wang et al and b) report on further detail into the actual triage behaviour of partici-

pants. Furthermore, by keeping the study design similar throughout the investigation, we are able to reduce the possibility of external and ecological validity variations between experiments.

Beyond being in line with previous research methodological approaches, we review the adherence of our studies to a heuristic evaluation which will guide us as to the external and ecological validity of our experiments. We review heuristics as outlined by Trochim [151] and relate the weaknesses and strengths of our studies with relation to these. The main study section of our experiments were mostly 'self sustained'. In other words, we could report on the data without the need for further information such as pre-study and post study questions. Due the fact that we need visual attention and not a cognitive understanding of what the participants are thinking, this is easier to be achieved. Any further information that was reported by the participants to compliment the data was recorded as self reported data which needed to be further investigated. An example of this is the comment that the participants would judge the relevance of a document in the first few seconds and the first page. This was reported on but not claimed by our data observational data. This method addresses the issue of Pre and Post Test effects.

One limitation which our studies, and lab studies have in general, is the effect of situational specifics. The location, was not in the natural surroundings and the timing of one of the studies had induced conditions which would not have always been replicated in the real world. These were required however in order to extract specific information. In order to make the electronic environemnt however, as natural as possible, all the software, including the custom made reader was made to replicate the exact working environment of the participants. Furthermore, the documents have not been altered in any way (with the exception of the user study in this chapter, done for further methodological suggestions) thus to not influence the natural behaviour of the participants (or to alter the dependable variables).

Using real users with realistic goals is another measure for testing the ecological validity of results [99]. The observers and facilitators need to also be familiar with the process. Care was taken in order to select participants for the documents triage process which would accurately reflect users that would perform document triage in real life. The types of documents accurately reflected the documents that our participants would triage for real in order to increase the ecological and external validity of the experiments. Our participants also varied in skills and experience to reduce

aptitude effects. One limitation that we can identify here is that there are many disciplines and we were only able to investigate a few of these. This was due to the expertise of the research team, which was able to understand and plan the studies better in this discipline.

Overall, our studies are mostly lab-based. In order to have dependant variables which we could influence and due to time and privacy reasons we had to do perform the studies in this fashion. Further studies like the diary study and direct observation can further address issues of ecological validity. We simply provide a foundation for further research and a basis for exploring document triage.

9.7   Summary

We began our research into document triage, by identifying a gap in the scientific knowledge of the information seeking process. Thus far, findings have been reported on relevance decisions, but the process of document triage, which is the manual process by which an information seeker goes through documents to make these decisions has been neglected. Related literature informed us of a need for research on the visual attention of users, specifically looking at the within-document triage stage. A conceptualisation of the document triage process did not, thus far, exist.

In this chapter, we have highlighted findings which correlate between all the studies in this thesis. We began by discussing the 'funnel model' and how it now gives us the tools to analyse the effects of each triage stage further using software engineering and information retrieval measures. These measures allow us to go beyond the evaluations of automatic computer-based of the search engine systems and allows us to measure the effectiveness of the human-computer interaction also. We reported on the natural bias with which the information seekers will begin the document triage process and how this may effect their cognitive decision making process. Some document features are dominant in terms of capturing an information seeker's attention during document triage. These features will not only visually capture the attention of the seeker but also affect the relevance decision making process. We report on which features are the most relevant and continue to present these within a behavioural model for triage. Features which are visually important, such as the initial page, as well as less important, such as emphasised text were described.

As our research progressed we uncovered how document triage is a complex and multi-faceted

process. One example is that of the effects of navigational patterns on decision making. Another example, is identifying the 'further reading stage' in the document triage process. The literature relating to our studies was discussed and our contribution was made evident. It was not feasible to examine all the factors of document triage in our thesis. Our intention was, instead, to focus on establishing a foundation upon which further research can be based. We will now see future work that will be undertaken to investigate the document triage process further.

Chapter 10

Future Work and Conclusions

10.1 Introduction

In this thesis we have produced a series of findings on the document triage process. Document triage is a popular process which is undertaken daily by information seekers. Using this activity, seekers make relevance decisions on documents. We have identified that there has been a large scientific knowledge gap in research specifically looking at the triage process. Studies have reported on the surrogate level of document triage of users and their navigational patterns but there was limited work on the stages occurring after the surrogate level; namely the within-document triage level (and as we later discovered a third level - further reading). Our research aimed to fill that knowledge gap by investigating the behaviour of users, and their visual attention during within-document triage. The two most important contributions that this thesis has offered are the document triage models (see Chapter 8), based on our empirical data and the detailed reports on visual behaviour during document triage (see Chapters 3 - 7).

Due to the lack of related work on the subject, we decided to begin our research by covering areas which were familiar to us in terms of our expertise and related work. We therefore chose semi-structured academic documents and academic users for the majority of the user studies. Our findings contributed significantly to the understanding of document triage as we expanded our basic studies to include variables such as time limitations and small screen real estate. By testing different scenarios, each time changing a limited amount of variables, we were able to generalise

typical triage behaviour of users while also reporting on distinct behaviours specific to each of our scenarios. This thesis has produced an extensive volume of data on document triage (Chapters 3, 4, 5, 6 and 7). During our research, it was found that there are more variables that could be realistically covered in one thesis. We therefore decided to lay the foundation, whereupon further research could be built upon to expand our findings.

In this chapter, we discuss further directions that the research can expand on, recognising knowledge gaps that still remain in the document triage process. Although there are several areas which we can expand on, we focus on work which directly links to our research. We suggest two areas as a next step in the development of our research. The first, includes research into our final third stage of document triage; namely that of *further reading*. During our research, it was discovered the document triage is a much more complicated process than we thought. Due to the process being multi-faceted, it was impossible therefore, to report on all the variables which could affect information seekers' triage activities. The second piece of work suggested in this chapter identifies further variables which can be investigated.

10.2   Investigating the Further Reading Stage

During our research, we identified three levels in which document triage occurs (See Chapter 8). The first level occurs at the surrogate level and has been researched by our earlier work [28] as well as by others [49, 100]. The second level is that of within-document triage and has been covered extensively by our research (See Chapters 3, 4, 5 and 6). The third level of 'further reading' has, to the best of our knowledge, not been researched. This stage in the process occurs after the information seeker has identified a document as potentially relevant and continues to read it in depth. After this in depth reading occurs, the seeker may still choose to reject the document as irrelevant.

Investigating how this process works will require a methodology different to the one used in our within-document triage studies. Further reading occurs over a longer period of time and cannot be studies in a lab-based environment. For this reason, we suggest diary studies or bespoke software to monitor users' actions. Jason Alexander uses a similar approach, although the software presented was limited specifically designed to record user navigation [9]. Other approaches have been used

to model the reading stage. Although some studies have reported on the searching side of in depth reading, they present the findings from a largely different perspective to us. Adler et al present a diary study on work related reading [5]. By using this method they are able to elicit requirements and reports on participants' activities without the use of specialised equipment. Furthermore, Wang and Soergel describe an interesting study in which they interview participants after they had selected documents [158, 159]. The task set to them, that of selecting documents to cite, was a faster process than reading, but their methodology was successful in identifying features which are confirmed by our results to be important. Such features are titles, headings and conclusions. We identify a limitation to this method in that some of the information reported by the participants as useful, does not get scrutinised visually. We identify that there are actions which do not agree with participants' reported behaviour [28]. It is therefore vital to be able to collect quantitative results from instrumented accurate sources, in order to identify patterns which may either not be reported on, or mistakenly self-reported. To the best of our knowledge, this has not been researched.

By understanding the final level of document triage we can produce a conceptualisation of the triage process which includes this final step. Furthermore, we can present all three of the levels in our model and begin to identify how they relate to each other. Our horizontal transitions (document relevance judgements), are dependant on user actions which at this third level, are not clearly understood.

10.3   Effects of Model Variation

In this thesis, we have identified the features within structured documents which can capture the visual attention of an information seeker (See Chapters 3 - 6). We have also modelled the visual behaviour a seeker is likely to undertake and reported on navigational patterns that the user will undergo while performing triage. While we have modelled the information seeker's visual attention during triage on structured documents. We have reported on how our funnel model can be more thoroughly reported on through further reading triage data. We now suggest other areas of research which will complement our data.

10.3.1   Document Types

In this thesis, we reported the within-document triage activities of scholars on academic publications. We have produced extensive data and modelled the visual attention of users while looking through the features of a document. Academic documents are semi-structured and are usually 5-30 pages of length as reported by Dillon [43]. Dillon also reported on how users will "read[1]" in search for information inside manuals. Other types of searching have looked at finding information in on-line newspapers [154]. These studies were more directed in that there was an information task to be accomplished rather than a relevant document found and read.

There are an abundance of media types which now represent documents. One such media type is video. Neema Moraveji reports on how "meta data are being created to improve video information foraging. However, the problem of video information retrieval remains unsolved [110]". The use of surrogates is dominant in the triage of video documents and has been thoroughly researched [69]. Moraveji however, suggests how video triage can be assisted using more visual rather than textual features [110]. Audio media suffers from the same problem; namely that of requiring surrogates for users to identify relevance. It would be interesting to expand our knowledge on how users can identify relevant audio recordings by reporting on within-audio document triage behaviour.

10.3.2   Cognitive Effects

In our studies, our focus was directed at the behavioural and procedural models of document triage. We followed on from Marchionini's procedural information seeking model to provide a more in depth analysis of two of the stages ('Examine Info' and 'Extract Results') [100]. Our within-document triage model, presented transitions and actions, specifically relating to the visual and navigational patters of users. We did not however, explain the 'why' information seekers do what they do. Andrew Dillon touched on the subject in describing different purposes for reading a document [43]. He also reports on reasons for the users going through stages of their reading process such as "getting a feel for the documents contents". Our research has reported subjective feedback from our participants. We recognises the same effect as Dillon reported when our participants commented on getting an initial opinion on the documents relevance by first impressions. Although some of

---

[1]Dillon's phrase 'read' can also imply that a users is reading while looking for information to to make a relevance judgement.

these cognitive elements were reported in our research, our goal was not to investigate the cognitive effects or behaviours of information seekers.

There have been models which report on the cognitive side of information seeking. An example of this is Kuhlthau's model[91]. In this model, Kuhlthau describes not only the actions of users, but also their thoughts, feelings and strategies. Apart from our diary study in chapter 7, we do not present any of the cognitive elements as presented by Kuhlthau. We can extract some of Kuhlthau's findings which relate to the triage process, such as the confidence levels after seekers skim a document, but there has, to the best of our knowledge, not been a direct study of the cognitive effects of document triage. This is an area which would complement our model and worth investigating further.

10.3.3 Methods

Our methodological techniques were mostly lab-based controlled experiments. Blandford et al, critique controlled experiments by saying that it is often not possible to "isolate and control the variables that are pertinent to the interactive behaviour, and it is difficult to design experiments to eliminate all confounds"[34]. In order to minimise any of these variables, we chose a methodology which is iterative in nature of the study designs. Our studies mirrored each other, changing one variable at a time to ensure that any external variables which cannot be controlled were eliminated. Furthermore, our research was compared to previous work by both the author and other researchers, in order to identify variations to any findings.

We are confident that our methodology produced accurate and concrete results which contributed to the understanding of document triage using empirical methods. Dillon recognised the need for "more formal analyses" of interview data and that "one issue that appears difficult to tackle directly through observation or interviewing is the structure of information space, which underlies the navigation issue". Although we have covered these requirements, we would welcome a more qualitative approach to elicit further cognitive data as well as a more exploratory methodology (similar to the one in chapter 7), to introduce new complementary findings to our data.

10.3.4   Model Expansion

Our two document triage models in chapter 8, presented new measurements to rate the efficiency of the human computer interaction during document triage. The measures, which were adapted from the information retrieval domain, are vital to evaluate users, systems and results, at every part of the triage process. These measurements were not implemented in a working scenario but were only theoretically explained and validated. One future aim, is to introduce the document triage models into the search process and use our measurements to evaluate every stage of the models. This can produce weightings, which will identify features within a document which are important. Furthermore, we can also implement our measurements on a larger scale and report on the efficiency of every part of the triage process.

We do not however, need to limit the measurements to our models alone. Be applying our calculations to each stage of Dillon's "Generic model of journal usage" or "Generic model of manual usage" we can create evaluations of the benefits of each of these stages. Most models which are associated with searching tasks, such as Dillon's and Marchionini's are iterative in nature. Our measurements can provide readings to rate each iteration and identify the positive and negative effects of each. By using this methodology on we aim future work to evaluate all the aforementioned models as well as others.

Conclusions This thesis was motivated by a gap in the existing scientific knowledge about the information seeking process; namely the activity of document triage. Document triage is the process by which information seekers go through to make relevance decisions based on an information need. There had previously been limited research on the document triage process. A series of user studies were conducted, each examining the triage process with a different set of variables. From these studies, this thesis was able to present patterns and behaviours that information seekers employ, in order to make their relevance decisions. These behaviours were modelled in two ways. The first model, presented document triage as a whole and presented the stages at which document triage occurs. The second model, presented triage at the within-document level from a navigational behavioural and visual perspective. Barriers to interaction were identified and the visual importance of features was reported on. The results were cross-examined with previous related finding to ensure validity and progressed the research into information seeking and information architecture

significantly. The thesis has made three major contributions to the research knowledge in the domain of information seeking.

1. **Models of the document triage process** are produced, grounded on the results from our user studies to produce theoretical frameworks for the document triage process. The first model describes the stages which an information seeker goes through when performing document triage. The second model reports on the conceptualisation of the within-document triage behaviours of information seekers. These models are created as a summary of all the work produced in this thesis. They guide the reader into thinking about the triage process and gives some design guidelines for further scrutiny. It is our hope that with future work, these models can be further refined and evolve to cover further areas and variables within document triage.

2. **User studies** presented empirical evidence to inform of the behaviour of information seekers during a set of different information seeking scenarios. A detailed analysis of information seekers' visual attention identifies patterns and areas of visual importance on documents. This type of empirical data was thus far not reported on and the area of document triage was underdeveloped. The evidence presented covers, again, a small amount of the scenarios which could be used. The work presented begins to uncover data which produce empirical validations of hypotheses relating to the document triage and information seeking field. These detailed findings, lay a foundation for further work to be carried out and uncovers gaps that require further scrutiny.

The research reported on provides a foundation for other researchers, designers and developers in the area of information seeking. There have been some limitation in the scope of studies that our time permitted. The focus of the research is placed on academic documents which are semi-structured. Although a general study was presented in which further results were analysed a more in depth research into individual topic domains is encouraged.

Furthermore, Chapter 9, discussed the ecological validity of the studies in the thesis. Most of the studies in the thesis have been lab based controlled studies. The reasons for this is the need for exact detail into the exact interactions and visual interests of the information seekers. We tailor

made the user studies in order to reduce the amount of threats to ecological validity, by replicating the software tools and equipment that the participants would use in real life situations. We also made the tasks as realistic as possible to actual triage activities that could have been undertaken by our users. Finally, we used a diary study to verify some of our findings and we identify further areas that need to be researched further for validity.

Researchers can use our findings as a starting point upon which to build further research scope and observations. Interaction designers and software developers can be informed by the our behavioural models presented to create software tools with advanced triage capabilities. Information seekers can also use the research presented to create better personal browsing strategies when performing document triage. Our adaptation of information seeking measurements give a set of tools to measure performance of individuals, software and methods during the entirety of the triage process.

The contributions made by this research have raised a number of possible directions for future work. One example, presented in Chapter 8, has demonstrated how we can create software for document triage in the future. We can also think of different genres of writing, and a larger user base, such as disabled users. We have also began research looking directly at textual documents. Our research can be used as a foundation for the further investigation of other types of documents such as multimedia. An investigation in these areas of document triage is encourages, which would further refine the theoretical and practical understanding of the digital library, and human computer interaction domain.

Appendix A

A.1 Guidelines and Heuristics

In this chapter, we have seen the constraints of the document triage process. We have also been presented with models, conceptualising the behaviour of information seekers during their document triage process. Using the models presented in conjunction with our empirical data, related work and subjective feedback from the studies described in this thesis we are able to provide heuristics in order to facilitate three user groups:

- Software Designers.

- Academic Authors.

- Information Seekers.

A.1.1 Software Designers

One of the aims of this thesis is to conceptualise the document triage model, in order for the subsequent informed creation of supportive software. Following are 6 isolated high level guidelines which will assist a software designer in the creation of document triage software.

In summary, these are:

1. Minimal Transitions.

2. Summarised Visible Important Features.

3. Full Text Availability.

4. Clear Document Length.

5. Easy Document Comparisons.

6. Easy Re-access to Already Triaged Documents.

Minimal Transitions. Bae et al report on how "overview and reading applications is a profound source of interruption, especially since users had to rearrange or reorder windows at almost every transition [14]. "Transitions in general can be a distraction, increase the cognitive load of an information seeker who interrupts the accumulation of information in order to focus on the navigation or wait for the next screen to load. A simple rule for more effective document triage software is to minimise transitions from one document to the other, and from reading activity to skimming for information.

Summarised Visible Important Features. From the previous studies, it is evident that not all document features carry the same weight in terms of document triage importance. As document triage is a fast-paced activity it is important that the most prominent of these are made available to the information seeker in a concise and summarative format. The four main features include the *main title, abstract, headings* and the *conclusions section.*

Full Text Availability. Information seekers are less likely to want to open the full text of a document when there is a summary presented to them [28]. There is however a sense of need to have the full document available to the information seekers. This has been subjectively evident from user feedback as well as empirical evidence justifying the the rating of systems which obfuscate the full text [118].

Clear Document Length. Document Length can be a deciding factor in the decision making process of document triage. Longer documents can be considered of higher value due to the amount of information that can be found within them. They can also however, produce difficulties to the informations seekers when they are unable to locate important parts of the document that contain useful information. This can result in the rejection of documents which are relevant to the information need. A good document triage system will be able to inform the user of the document length, while encouraging and supporting the location of relevant information within the document.

Easy Document Comparisons. One of the most difficult tasks that an information seeker

may face, is the assessment of documents in terms of the extent of their relevance to their information need. When the information need is not very familiar to an information seeker he or she often sets a standard based on the material read during the triage. The first document for example, can influence the threshold of acceptance for the remainder of the documents. The second document may move that threshold and so on. It is vital for easy comparisons between documents from the presentation of the document set. This will allow the information seeker to create a threshold using a broader range of documents rather than a limited amount.

Easy Re-access to Already Triaged Documents. It is found that information seekers who reject documents, will seldom return to re-assess them. Even when new insight into the intrinsic relevance value of the document is established, the user will simply reject the document and hope that it will re-appear in another search, rather than return to it [28]. If a user is allowed fast, easy access to all the documents, including the ones already triaged, there is more likelihood of the documents being re-assessed.

A.1.2 Academic Authors

In the thesis thus far, the patterns and behaviour of information seekers during within-document triage have been recorded. Using eye-tracking technology the exact visual attention range of information seekers was extracted and modelled. Although performing document triage is not the only way a document can be selected for reading [1], it is the most common. It is therefore reasonable to empower the authors of the documents with guidelines into writing to be effectively triaged. Using these guidelines, users will be able to understand the content of the documents faster and with greater accuracy as to their information need. The authors will in turn benefit themselves by making sure that information seekers who could benefit from reading (and perhaps citing) their work do so. There are five guidelines authors should take into account when writing an academic publication. These are:

1. Clear and Descriptive Main Title.

2. Clear Summarative Abstract.

---

[1]Another example may be a colleague's recommendation

3. Clear Descriptive Sentences of Introduction.

4. Clear Section Headings.

5. Clear and Concise Summary Sentence After Every Main Section Heading.

6. Clearly Summarise Results/Findings/Key Points Early in Conclusion.

Clear and Descriptive Main Title. Information Seekers will always read the title of a document during their triage activity. Subjective feedback has indicated that although useful, the main title can be either misleading and/or uninformative. Although there are limited words that an author can use to represent a heading, an accurate description of the contents of the document is preferable than an abstract "catchy" title.

Clear Summarative Abstract. The abstract is empirically classified as the most important part of a document during the triage process. Subjective feedback often criticised the abstract of "being too abstract". Although this may seem like a contradiction, further enquiry as to the meaning and reasons for such phrases uncovered the intrinsic value of the abstract to an information seeker. An abstract should outline the areas of research covered in the document and be less descriptive of the actual problem that needed solving or of the general area of research within which the problem lies.

Clear Descriptive Sentences of Introduction. According to our eye-tracking data, the introduction is viewed but not in its entirety during document triage. The most scrutinised part of the introduction is the first few lines. The information seeker is less likely to read further than the first two sentences (about 5 lines). Qualitative feedback concluded that the introduction, although being subjectively important to all the participants, was deemed to produce limited information on the document contents from the first few lines of the introduction. As one participants put it: "usually the beginning of the introduction starts with the author's philosophies and beliefs rather than getting to the point".

Clear Section Headings. Headings are heavily scrutinised by an information seeker performing document triage. During the triage process the skimming behaviour during within-document triage is interrupted by the appearance of a section heading, during which a user will likely read the heading to discern the content of the section. If the heading is not specific or detailed enough,

the entire section can easily be ignored. One example, would be to keep the conclusion section as *Conclusions* rather than renaming the section *Summary*.

Clear and Concise Summary Sentence After Every Main Section Heading. It was mentioned of how the first few lines of the introduction are scrutinised more thoroughly than the remaining section. The same behaviour holds true for headings. When (and only when) a section's heading is deemed to be relevant to the information seeker's need, the first few lines under the heading. Again, this has to do with the content of the section, and how effective the first few lines of a section are at explaining the content of the section.

Clearly Summarise Results/Findings/Key Points Early in Conclusion. As with the Introduction and the individual sections, the conclusion is only scrutinised for a few lines under the conclusion heading. Although users seem more willing to traverse the conclusion to a greater extent than the introduction or section content, a conclusion needs to be a fast, summary of the Results/Findings/Key Points of the document as soon as possible. Contributions and further material is advised to be added towards the end of this section.

A.1.3 Information Seekers

With the models presented in this thesis, it is clear that information seekers have set patterns in the way the perform document triage. All the user studies described concurred in the majority of the features and patterns that the users follow. Here we present heuristics that can act as guidelines to an information seeker performing within-document triage on academic documents without the support of specialised software.

These are:

1. Ignore Document Length.

2. Target Less Information.

3. Read in More Depth.

Ignore Document Length. Document Length can often be a deciding factor in an information seeker's relevance decision process. A long document may be considered to cumbersome to read, and therefore rejected for that reason alone. On the other hand, a document which contains more

information may be considered more favourably for that reason, irrelevant of the actual relevant information. An information seeker would do well to ignore the length of the document and focus on the content instead.

Target Less Information. It was found that plain text is often viewed but not read, while the information seeker skims through the document. Skimming is largely a location targeting activity in order to absorb any areas of a document which may betray the relevance of the document to the information need. Information seekers are advised to avoid the long skimming technique as it was found that most decisions and time is spent reading specific features of a document. By having a triage strategy, an information seeker can go directly do document features which are predetermined and therefore save time and effort while making the same decision. An example of such a strategy could be: Read main title, read abstract, read conclusion.

Read in More Depth. It was found in the studies that during skimming, information seekers do not have enough time to read any of the material being skimmed. It was also found that there is minimal reading when the information seeker does begin a reading activity, thus missing information located a little later in text of the introduction, conclusion and abstract. Information seekers are encouraged to substitute skimming time with more in depth reading during within-document triage.

Appendix B

All Formatting and logos are removed.

B.1 Bill of Rights Form

B.2 Chapter 4 Study Consent Form

B.3 Chapter 5 Study Consent Form

B.4 Chapter 6 Study Consent Form

B.5 Chapter 7 Study Consent Form

B.6 Chapter 8 Study Consent Form

Research Participant's Bill of Rights

The following is a list of your rights if you participate in a research project organised within the Centre for Human Computer Interaction Design.

As a research participant, you have the right:

- To be treated with respect and dignity in every phase of the research.

- To be fully and clearly informed of all aspects of the research prior to becoming involved in it.

- To enter into clear, informed, and written agreement with the researcher prior to becoming involved in the activity. You should sense NO pressure, explicit or otherwise, to sign this contract.

- To choose explicitly whether or not you will become involved in the research under the clearly stated provision that refusal to participate or the choice to withdraw during the activity can be made at any time without penalty to you.

- To be treated with honesty, integrity, openness, and straightforwardness in all phases of the research, including a guarantee that you will not unknowingly be deceived during the course of the research.

- To receive something in return for your time and energy.

- To demand proof that an independent and competent ethical review of human rights and protections associated with the research has been successfully completed.

- To demand complete personal confidentiality and privacy in any reports of the research unless you have explicitly negotiated otherwise.

- To expect that your personal welfare is protected and promoted in all phases of the research, including knowing that no harm will come to you.

- To be informed of the results of the research study in a language you understand.

- To be offered a range of research studies or experiences from which to select, if the research is part of fulfilling your educational or employment goals.

The contents of this bill were prepared by the University of Calgary who examined all of the relevant Ethical Standards from the Canadian Psychological Association's Code of Ethics for Psychologists, 1991 and rewrote these to be of relevance to research participants.

Descriptions of the CPA Ethical Code and the CPA Ethical Standards relevant to each of these rights are available at http://www.cpa.ca/ethics2000.html and http://www.psych.ucalgary.ca/Research/ethics/bill/billcode.html if you would like to examine them.

The complete CPA Ethical Code can be found in Canadian Psychological Association *"Companion manual for the Canadian Code of Ethics for Psychologists"* (1992).

Research Consent Form

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Research Project Title

Exploring Document Triage Empirically.

Researcher

Mr. Fernando Loizides.

Experiment Purpose

The purpose of this experiment is to observe and record the ways that computer users search for information and triage documents using a PDF reader.

Participant Recruitment and Selection

Computer Science related participants at the appropriate level (Masters or above).

Procedure

This session will require about 30 minutes of your time. You will be asked to answer a pre-study questionnaire and fill in a scoring sheet during an observational study. None of the tasks is a test – our objective is to find out how you approach the tasks. There is no right or wrong method.

Data Collection

As you work on the task your mouse and keyboard actions will be recorded by the computer using screen capturing software and customised PDF Reader software. The interviewer will take notes. There will be no video or sound recording of you in the physical world sense, only of the contents of the screen.

Data Archiving/Destruction

Data will be kept secure stored in the Usability Laboratory archive under the control of the Usability Laboratory Manager.

Confidentiality

Confidentiality and participant anonymity will be strictly maintained. All information gathered will be used for statistical analysis only and no names or other identifying characteristics will be stated in the final or any other reports.

Likelihood of Discomfort

There is no likelihood of discomfort or risk associated with participation.

Researcher

Mr. Fernando is working towards his Ph.D research on Document Triage. His supervisor is Dr George Buchanan.

Fernando Loizides can be contacted by his email address: Fernando.loizides@gmail.com

Finding out about Results

The Participants can find out the results of the study by contacting the researcher after January 1, 2009.

Agreement

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to take part as a participant. In no way does this waive you legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact the researcher.

_____   _____

Participant   Date

_____   _____

Investigator/Witness   Date

A copy of this consent form has been given to you to keep for your records and reference.

Research Consent Form

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Research Project Title

Exploring Document Triage.

Researcher

Mr. Fernando Loizides.

Experiment Purpose

The purpose of this experiment is to observe and record the ways that computer users triage documents using an eye-tracker.

Participant Recruitment and Selection

HCI related participants at the appropriate level.

Procedure

This session will require about 60 minutes of your time. You will be asked to answer a pre-study questionnaire and fill in a scoring sheet during an observational study. None of the tasks is a test – our objective is to find out how you approach the tasks. There is no right or wrong method.

Data Collection

As you work on the task your mouse and keyboard actions will be recorded by the computer using screen capturing software and an eye tracker. The interviewer will take notes.

Data Archiving/Destruction

Data will be kept secure stored in the Usability Laboratory archive under the control of the Usability Laboratory Manager.

Confidentiality

Confidentiality and participant anonymity will be strictly maintained. All information gathered will be used for statistical analysis only and no names or other identifying characteristics will be stated in the final or any other reports.

Likelihood of Discomfort
There is no likelihood of discomfort or risk associated with participation.

Researcher
Mr. Fernando is working as a Research Assistant at HCID in City University. This study will contribute to his research on Document Triage. His supervisor is Dr George Buchanan.

Fernando Loizides can be contacted by email : Fernando.loizides@gmail.com

Finding out about Results
The Participants can find out the results of the study by contacting the researcher after February 1, 2010.

Agreement
Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to take part as a participant. In no way does this waive you legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact the researcher.

_____   _____

Participant   Date

_____   _____

Investigator/Witness   Date

A copy of this consent form has been given to you to keep for your records and reference.

Research Consent Form

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Research Project Title

Exploring Document Triage.

Researcher

Mr. Fernando Loizides.

Experiment Purpose

The purpose of this experiment is to observe and record the ways that computer users triage documents using a mobile small screen device.

Participant Recruitment and Selection

HCI related participants at the appropriate level.

Procedure

This session will require less than 60 minutes of your time. You will be asked to answer a pre-study questionnaire and fill in a scoring sheet during an observational study using a small screen device. None of the tasks is a test – our objective is to find out how you approach the tasks. There is no right or wrong method.

Data Collection

As you work on the task your mouse mobile device screen will be recorded using specialist mobile testing equipment. The interviewer will take notes.

Data Archiving/Destruction

Data will be kept secure stored in the Usability Laboratory archive under the control of the Usability Laboratory Manager.

Confidentiality

Confidentiality and participant anonymity will be strictly maintained. All information gathered will be used for statistical analysis only and no names or other identifying characteristics will be stated in the final or any other reports.

Likelihood of Discomfort

There is no likelihood of discomfort or risk associated with participation.

Researcher

Mr. Fernando is working as a Research Assistant at HCID in City University. This study will contribute to his research on Document Triage. His supervisor is Dr George Buchanan.

Fernando Loizides can be contacted by email : Fernando.loizides@gmail.com

Finding out about Results

The Participants can find out the results of the study by contacting the researcher after December 1, 2010.

Agreement

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to take part as a participant. In no way does this waive you legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact the researcher.

_____   _____

Participant   Date

_____   _____

Investigator/Witness   Date

A copy of this consent form has been given to you to keep for your records and reference.

Research Consent Form

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Research Project Title
Exploring Document Triage During Time Constraint Activities.

Researcher
Mr. Fernando Loizides.

Experiment Purpose
The purpose of this experiment is to observe and record the ways that computer users triage documents, under time constraints, using an eye-tracker.

Participant Recruitment and Selection
HCI related participants at the appropriate level.

Procedure
This session will require less than 60 minutes of your time. You will be asked to answer a pre-study questionnaire and rate document relevance during an observational study. An eye-tracker and screen recording software will document the procedure None of the tasks is a test – our objective is to find out how you approach the tasks. There is no right or wrong method.

Data Collection
As you work on the task your mouse and keyboard actions will be recorded by the computer using screen capturing software and an eye tracker. The interviewer will take notes.

Data Archiving/Destruction
Data will be kept secure stored in the Usability Laboratory archive under the control of the Usability Laboratory Manager.

Confidentiality
Confidentiality and participant anonymity will be strictly maintained. All information gathered will be used for statistical analysis only and no names or other identifying characteristics will be stated in the final or any other reports.

Likelihood of Discomfort

There is no likelihood of discomfort or risk associated with participation.

Researcher

Mr. Fernando is working as a Research Assistant at HCID in City University. This study will contribute to his research on Document Triage. His supervisor is Dr George Buchanan.

Fernando Loizides can be contacted by email : Fernando.loizides@gmail.com

Finding out about Results

The Participants can find out the results of the study by contacting the researcher after July 1, 2010.

Agreement

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to take part as a participant. In no way does this waive you legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact the researcher.

_____   _____

Participant   Date

_____   _____

Investigator/Witness   Date

A copy of this consent form has been given to you to keep for your records and reference.

Research Consent Form

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Please take the time to read this form carefully and to understand any accompanying information.

Research Project Title
Exploring Document Triage Activities.

Researcher
Mr. Fernando Loizides.

Experiment Purpose
The purpose of this study is to record everyday document triage activities of users using a diary.

Participant Recruitment and Selection
HCI related participants at the appropriate level.

Procedure
This study will take place over the period of a month. You will be asked to fill in a diary account of 10 of your information searches. None of the tasks is a test – our objective is to find out how you triage information. There is no right or wrong method.

Data Archiving/Destruction

Data will be kept secure stored in the Usability Laboratory archive under the control of the Usability Laboratory Manager.

Confidentiality
Confidentiality and participant anonymity will be strictly maintained. All information gathered will be used for statistical analysis only and no names or other identifying characteristics will be stated in the final or any other reports.

Likelihood of Discomfort
There is no likelihood of discomfort or risk associated with participation.

Researcher

Mr. Fernando is working as a Research Assistant at HCID in City University. This study will contribute to his research on Document Triage. His supervisor is Dr George Buchanan.

Fernando Loizides can be contacted by email : Fernando.loizides@gmail.com

Finding out about Results

The Participants can find out the results of the study by contacting the researcher after February 1, 2009.

Agreement

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to take part as a participant. In no way does this waive you legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to not answer specific items or questions in interviews or on questionnaires. You are free to withdraw from the study at any time without penalty. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact the researcher.

_____   _____

Participant   Date

_____   _____

Investigator/Witness   Date

A copy of this consent form has been given to you to keep for your records and reference.

Appendix C

Age: Sex:

How long have you used document readers?

How many academic documents on average would you say you triage per day?

How long would you say on average do you take to triage these documents?

Rate the following from 1 (not useful) to 10 (extremely) useful according to the usefulness in helping deciding the relevance of a document.

Main Title

Abstract

Headings

Introduction

Plain Text

Conclusions

References

Images and Figures

Highlighted text

Do you prefer to triage on paper or electronically? Paper ____   Electronic ____

Task 1

You are a teacher. You are thinking of integrating a tablet PC/touch screen environment for teaching in your classroom. You want to find different studies/instances which give information on cases already implemented and find out a) how receptive and positive(or negative) the students were to them b) limitations and problems discovered. Rate the documents in the TABLET FOLDER giving them a score from 1 to 10 (1 being the worst rating for the job and 10 being the best). You should not read the whole thing simply find an answer as quickly as possible in order to read it in depth "later" so to speak.

Tablet =   Tablet3 =

Tablet1 =   Tablet4 =

Tablet2 =   Tablet5 =

TASK 2

In every discipline, when designing a product for the public, uses certain techniques for the evaluation of the product. These techniques are often reported in papers as an indicator of good evaluation for the product. As a new Human Computer Interaction researcher you are interested in discovering different types of techniques and see how good they are. Rate the documents in the HCI FOLDER giving them a score from 1 to 10 (1 being the worst rating for the job and 10 being the best). You should not read the whole thing simply find an answer as quickly as possible in order to read it in depth at a "later date" so to speak.

HCI =   HCI5 =

HCI1 =   HCI6 =

HCI2 =   HCI7 =

HCI3 =   HCI8 =

HCI4 =   HCI9 =

How certain are you about the decisions you have made?

What features about documents create a negative impression about the document?

What features about documents encourage you to select the documents for readings?

How did the following features help / deter you from performing your document triage?

Headings

Introduction

Plain Text

Conclusions

References

Images and Figures

Highlighted text

Age: Sex:

How long have you used document readers?

How many academic documents on average would you say you triage per day?

How long would you say on average do you take to triage these documents?

Rate the following from 1 (not useful) to 10 (extremely) useful according to the usefulness in helping deciding the relevance of a document.

Main Title

Abstract

Headings

Introduction

Plain Text

Conclusions

References

Images and Figures

Highlighted text

Do you prefer to triage on paper or electronically? Paper ____ Electronic ____ Please Justify your answer.

Task 1

You are a teacher. You are thinking of integrating a tablet PC/touch screen environment for teaching in your classroom. You want to find different studies/instances which give information on cases already implemented and find out a) how receptive and positive(or negative) the students were to them b) limitations and problems discovered. Rate the documents in the TABLET FOLDER. You should not read the whole thing simply find an answer in order to read it in depth "later" so to speak.

Tablet = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet1 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet2 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet3 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet4 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet5 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

TASK 2

In every discipline, when designing a product for the public, uses certain techniques for the evaluation of the product. These techniques are often reported in papers as an indicator of good evaluation for the product. As a new Human Computer Interaction researcher you are interested in discovering different types of techniques and see how good they are. Rate the documents in the HCI FOLDER. You should not read the whole thing simply find an answer in order to read it in depth at a "later date" so to speak.

HCI = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI1 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI2 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI3 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI4 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI5 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI6 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI7 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI8 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

HCI9 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Age: Sex:

What Small Screen Devices do you use?

How long have you used small screen readers?

How many academic documents would you say you triage on them per day?

On average how long would you say you take to triage these documents?

Please rate the following from 0 to 10 in terms of how important you think they are to deciding the relevance of a document while performing document triage. (0 = not relevant, 10 = highly relevant).

Main Title	
Abstract	
Headings	
Introduction	
Plain Text	
Conclusions	
References	
Images and Figures	
Text that Stands Out	

Do you prefer to search on small screen or paper? (please mark with an X)

Paper |__|__|__|__|__| Both the same |__|__|__|__|__| Small Screen

Rate the following from 1 (not useful) to 10 (extremely) useful according to the usefulness in helping deciding the relevance of a document. Explaining the reason for the score.

Main Title ☐

Abstract ☐

Headings ☐

Introduction ☐

Plain Text ☐

Conclusions ☐

References ☐

Images and Figures ☐

Highlighted text ☐

Age: Sex:

How long have you used document readers?

How many academic documents on average would you say you triage per day?

How long would you say on average do you take to triage these documents?

Rate the following from 1 (not useful) to 10 (extremely) useful according to the usefulness in helping deciding the relevance of a document.

Main Title

Abstract

Headings

Introduction

Plain Text

Conclusions

References

Images and Figures

Highlighted text

Do you prefer to triage on paper or electronically? Paper ____   Electronic ____

Task 1

You are a teacher. You are thinking of integrating a tablet PC/touch screen environment for teaching in your classroom. You want to find different studies/instances which give information on cases already implemented and find out a) how receptive and positive(or negative) the students were to them b) limitations and problems discovered. Rate the documents in the TABLET FOLDER. You should not read the whole thing simply find an answer in order to read it in depth "later" so to speak.

Tablet = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet1 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet2 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet3 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet4 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Tablet5 = Relevant - Somewhat (Mostly) Relevant - Somewhat (Mostly) Irrelevant - Not Relevant.

Please describe the triage operations, after you have completed your search. Keep the following questions in mind when you are filling in your diary entries.

Triage Number.

Means used for Triage: Internet, Book, Library, Other.

Is the information you are looking for of a personal nature or work related?

Is the search for a specific item or a broader search for knowledge on a subject?

Comment on how important it is that you find the information you are looking for?

What information need do you have? What is it that you are looking for?

How well do you know the subject area of the item you are looking for?

What were your time constraints in terms of getting the information needed?

How much time did that search take you in total?

What means physical or electronic did you use to search?

Do you think the means I am using are good for searching?

What seems to be the problem locating the information required?

How do you feel during the steps of triage? Before, during and after.

How many documents are returned on the query search?

How many and what documents you open and why?

If you are returned with a results list what do you look at?

What do you look for first and during the within-document triage process?

What do you think of each of the scrutinised documents?

At which part of your triage do you initially realise the potential relevance of the document?

Do you keep performing triage after the initial relevance decision you have made? If yes why?

Do you change your mind to the initial decision after looking at a document more?

At which point do you stop the triage and why?

How confident and pleased are you with a) the quality of the selected documents b) the amount of information you have potentially extracted from them?

Will you repeat a search later for the same information need?

Do you read the documents later and realise that the relevance decisions you have made during within-document triage were false?"

Bibliography

[1] Jones, Matt and Buchanan, George and Mohd Nasir, Norliza An Evaluation of WebTwig - A Site Outliner for Handheld Web Access. In *Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing*, pages 343–345, 1999.

[2] Keenan, Stacey A. Effects of chunking and line length on reading efficiency. In *Visible Language*, pages 61–80, 1984.

[3] Anne Adams and Ann Blandford. Digital libraries in academia: Challenges and changes. In *Proceedings of the 5th International Conference on Asian Digital Libraries: Digital Libraries: People, Knowledge, and Technology*, pages 392–403, 2002.

[4] Anne Adams and Ann Blandford. Digital libraries' support for the user's 'information journey'. In *Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*, pages 160–169, 2005.

[5] Annette Adler, Anuj Gujar, Beverly L. Harrison, Kenton O'Hara, and Abigail Sellen. A diary study of work-related reading: design implications for digital reading devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 241–248, 1998.

[6] Annette Adler, Anuj Gujar, Beverly L. Harrison, Kenton O'Hara, and Abigail Sellen. A diary study of work-related reading: design implications for digital reading devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '98, pages 241–248, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[7] Mortimer J. Adler and Charles L. Van Doren. *How to Read a Book*. Simon and Schuster, 1996.

[8] Denise E. Agosto. Bounded rationality and satisficing in young people's web-based decision making. In *Journal of the American Society for Information Science and Technology*, pages 16–27, 2002.

[9] Jason Alexander and Andy Cockburn. An empirical characterisation of electronic document navigation. In *Graphics Interface*, pages 123–130, 2008.

[10] Aula Anne, Majaranta Pivi, and Raiha Kari Jouko. Eye-tracking reveals the personal styles for search result evaluation. In *INTERACT*, pages 1058–1061, 2005.

[11] Susanne Askwall. Computer supported reading vs reading text on paper: a comparison of two reading situations. pages 425–439, 1985.

[12] A.D Baddeley. Is working memory still working? In *American Psychologist*, pages 851–864, 2001.

[13] Rajiv Badi, Soonil Bae, J. Michael Moore, Konstantinos Meintanis, Anna Zacchi, Haowei Hsieh, Frank Shipman, and Catherine C. Marshall. Recognizing user interest and document value from reading and organizing activities in document triage. In *IUI '06: Proceedings of the 11th international conference on Intelligent user interfaces*, pages 218–225, 2006.

[14] Soonil Bae, Rajiv Badi, Konstantinos A. Meintanis, J. Michael Moore, Anna Zacchi, Hao wei Hsieh, Catherine C. Marshall, and Frank M. Shipman III. Effects of display configurations on document triage. In *INTERACT*, pages 130–143, 2005.

[15] Ricardo A. Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[16] Carol L. Barry. Document representations and clues to document relevance. In *Journal of the American Society for Information Science*, pages 1293–1303, 1998.

[17] J Marcia Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, pages 407–424, 1989.

[18] David Bawden and Lyn Robinson. The dark side of information: overload, anxiety and other paradoxes and pathologies. *J. Information Science*, 35(2):180–191, 2009.

[19] N. J. Belkin. Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, (5):133–143, 1980.

[20] N. J. Belkin, R. N. Oddy, and H. M. Brooks. Ask for information retrieval: part ii. results of a design study. pages 145–164, 1982.

[21] N. J. Belkin, R. N. Oddy, and H. M. Brooks. *Ask for information retrieval: part I.: background and theory*, pages 299–304. 1997.

[22] N.J Belkin, R.N Oddy, and H.M Brooks. Ask for information retrieval: Part i. background and theory. In *The Journal of Documentation*, volume 2, pages 61–71, 1982.

[23] Ann Blandford, Anne Adams, Simon Attfield, George Buchanan, Jeremy Gow, Stephann Makri, Jon Rimmer, and Claire Warwick. Pret a rapporter: Evaluating digital libraries alone and in context. *Information Processing and Management*, 2008.

[24] Judd Charles M Brewer Marilynn B, Reis Harry T. *Research design and issues of validity. Handbook of research methods in social and personality psychology.* New York, NY, US: Cambridge University Press, 2000.

[25] D. Bruneau, M. A. Sasse, and J. D McCarthy. The eyes never lie: The use of eye tracking data in hci research. In *Proceedings of the CHI02 Workshop on Physiological Computing*, pages 441–480, 2002.

[26] Doug Bryan and Anatole Gershman. Opportunistic exploration of large consumer product spaces. In *Proceedings of the 1st ACM conference on Electronic commerce*, EC '99, pages 41–47, New York, NY, USA, 1999. ACM.

[27] George Buchanan, Sarah Farrant, Matt Jones, Harold Thimbleby, Gary Marsden, and Michael Pazzani. Improving mobile internet usability. In *WWW '01*, pages 673–680, New York, NY, USA, 2001. ACM.

[28] George Buchanan and Fernando Loizides. Investigating document triage on paper and electronic media. In *Procs. of the European Conf. on Reasearch and Advanced Technology for Digital Libraries*, number 35, pages 416–427, 2007.

[29] Chris Buckley, Gerard Salton, James Allan, and Amit Singhal. In *TREC*, 1994.

[30] Moira Burke, Anthony Hornof, Erik Nilsen, and Nicholas Gorman. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Trans. Comput.-Hum. Interact.*, 12:423–445, December 2005.

[31] Georg Buscher, Susan T. Dumais, and Edward Cutrell. The good, the bad, and the random: an eye-tracking study of ad quality in web search. In *SIGIR*, pages 42–49, 2010.

[32] Donald Byrd. A scrollbar-based visualization for document navigation. In *Proceedings of the fourth ACM conference on Digital libraries*, DL '99, pages 122–129, New York, NY, USA, 1999. ACM.

[33] Katriina Byström and Kalervo Järvelin. Task complexity affects information seeking and use. pages 191–213, 1995.

[34] Paul Cairns and Anna L. Cox. *Research Methods for Human-Computer Interaction*. Cambridge University Press, New York, NY, USA, 1st edition, 2008.

[35] Philip K. Chan. A non-invasive learning approach to building web user profiles, 1999.

[36] C. Chen. Visual spatial thinking in digital libraries - top ten problems. In *Joint Conference in Digital Libraries*, 2001.

[37] Eugene Jr Chewning and Adrian M. Harrell. The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task. In *Accounting, Organizations and Society.*, pages 527–542, 1990.

[38] Andy Cockburn, Carl Gutwin, and Jason Alexander. Faster document navigation with space-filling thumbnails. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, CHI '06, pages 1–10, New York, NY, USA, 2006. ACM.

[39] C. Cool, N. J. Belkin, O. Frieder, and P. Kantor. Characteristics of texts affecting relevance judgments. In *Proceedings of the 14th National Online Meeting*, pages 77–84, 1993.

[40] W. Cooper. Relevance and information behavior. In *Annual Review of Information Science and Technology (ARIST)*, pages 3–48, 1994.

[41] L. Cowen, L. J. Ball, and J Delin. An eye-movement analysis of web-page usability. In *People and Computers XVI Memorable yet Invisible: Proceedings of HCI 2002 SpringerVerlag Ltd.*, pages 317–335, 2002.

[42] Edward Cutrell and Zhiwei Guan. What are you looking for?: an eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 407–416, New York, NY, USA, 2007. ACM.

[43] Andrew Dillon. *Designing Usable Electronic Text, Second Edition*. Taylor & Francis, Inc., Bristol, PA, USA, 2004.

[44] Norman Donald. *The Psychology Of Everyday Things*. Basic Books, 1988.

[45] Denis Drieghe, Keith Rayner, and Alexander Pollatsek. The mind's eye: Cognitive and applied aspects of eye movement research. *Journal Of Experimental Psychology. Human Perception And Performance*, 573–605(5):954–959, 2003.

[46] Didier Dubois, Henri Prade, and Ronald R Yager. *Readings in Fuzzy Sets for Intelligent Systems (The Morgan Kaufmann Series in Representation and Reasoning*. Morgan Kaufmann Publishers, 1993.

[47] G.B Duggan and S.J Payne. Skim reading by satisficing: Evidence from eye tracking. In *Journal of Experimental Psychology: Applied.*, pages 228–242, 2009.

[48] G.B Duggan and S.J Payne. Text skimming: the process and effectiveness of foraging through text under time pressure. In *CHI Conference on Human Factors in Computing Systems.*, pages 1141–1150, 2011.

[49] Susan T. Dumais, Georg Buscher, and Edward Cutrell. Individual differences in gaze patterns for web search. In *Proceeding of the third symposium on Information interaction in context*, IIiX '10, pages 185–194, New York, NY, USA, 2010. ACM.

[50] D Ellis. A behavioral approach to information retrieval system design. In *Journal of Documentation*, pages 171–212, 1989.

[51] David Ellis. The effectiveness of information retrieval systems: the need for improved explanatory frameworks. In *Social Science Information Studies*, pages 261–272, 1984.

[52] David Ellis, Deborah Cox, and Katherine Hall. A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, pages 356–369, 1993.

[53] Martin J. Eppler and Jeanne Mengis. The concept of information overload: A review of literature from organization science, accounting, marketing, mis, and related disciplines. *Inf. Soc.*, 20(5):325–344, 2004.

[54] Michael W. Eysenck and Mark T. Keane. *Cognitive Psychology: A Student's Handbook*. 2000.

[55] Thomas J. Froehlich. Relevance reconsideredtowards an agenda for the 21st century: introduction to special topic issue on relevance research. *J. Am. Soc. Inf. Sci.*, 45:124–134, April 1994.

[56] George W. Furnas and Samuel J. Rauch. Considerations for information environments and the navique workspace. In *Proceedings of the third ACM conference on Digital libraries*, pages 79–88, 1998.

[57] H. J. Goldberg and X. P Kotval. Computer interface evaluation using eye movements: Methods and constructs. In *International Journal of Industrial Ergonomics*, pages 631–645, 1999.

[58] Laura A. Granka, Thorsten Joachims, and Geri Gay. Eye-tracking analysis of user behavior in www search. In *SIGIR*, pages 478–479, 2004.

[59] Laura A. Granka, Thorsten Joachims, and Geri Gay. Eye-tracking analysis of user behavior in www search. In *Proceedings of the 27th annual international ACM SIGIR conference on*

Research and development in information retrieval, SIGIR '04, pages 478–479, New York, NY, USA, 2004. ACM.

[60] Zhiwei Guan and Edward Cutrell. An eye tracking study of the effect of target rank on web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 417–420, New York, NY, USA, 2007. ACM.

[61] David J. Harper, Sara Coulthard, and Sun Yixing. A language modelling approach to relevance profiling for document browsing. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, JCDL '02, pages 76–83, New York, NY, USA, 2002. ACM.

[62] David J. Harper, Ivan Koychev, and Yixing Sun. Query-based document skimming: a user-centred evaluation of relevance profiling. In *Proceedings of the 25th European conference on IR research*, ECIR'03, pages 377–392, Berlin, Heidelberg, 2003. Springer-Verlag.

[63] David J. Harper, Ivan Koychev, Yixing Sun, and Iain Pirie. Within-document retrieval: A user-centred evaluation of relevance profiling. *Inf. Retr.*, 7:265–290, September 2004.

[64] R Hastie. Problems for judgment and decision making. In *Annual Review of Psychology*, pages 653–683, 2001.

[65] Marti A. Hearst. Tilebars: visualization of term distribution information in full text information access. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 59–66, 1995.

[66] J.M. Henderson and F Ferreira. *The interface of language, vision, and action: Eye movements and the visual world.* New York, NY, US: Psychology Press., 2004.

[67] David G. Hendry and David J. Harper. An informal information-seeking environment. *J. Am. Soc. Inf. Sci.*, 48:1036–1048, November 1997.

[68] Dara Lee Howard. Pertinence as reflected in personal constructs. *J. Am. Soc. Inf. Sci.*, 45:172–185, April 1994.

[69] Anthony Hughes, Todd Wilkens, Barbara M. Wildemuth, and Gary Marchionini. Text or pictures? an eyetracking study of how people view digital video surrogates. In *Proceedings*

of the 2nd international conference on Image and video retrieval, CIVR'03, pages 271–280, Berlin, Heidelberg, 2003. Springer-Verlag.

[70] A. Hyrskykari, S. Ovaska, P. Majaranta, KJ. Rih, and M Lehtinen. Gaze path stimulation in retrospective think aloud. In *Journal of Eye Movement Research*, pages 1–18, 2008.

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

[72] Paul Janecek and Pearl Pu. Searching with semantics: An interactive visualization technique for exploring an annotated image collection. In *On The Move to Meaningful Internet Systems 2003: OTM 2003Workshops*, pages 185–196. 2003.

[73] Paul Janecek and Pearl Pu. Opportunistic search with semantic fisheye views. In *Web Information Systems WISE 2004*, volume 3306, pages 668–680. Springer Berlin / Heidelberg, 2004.

[74] Paul Janecek and Pearl Pu. An evaluation of semantic fisheye views for opportunistic search in an annotated image collection. *International Journal on Digital Libraries*, 5:42–56, 2005.

[75] Joseph W. Janes. Relevance judgments and the incremental presentation of document representations. In *Information Processing and Management*, pages 629–646, 1991.

[76] Kalervo Järvelin and T. D. Wilson. On conceptual models for information seeking and retrieval research. *Inf. Res.*, 9(1), 2003.

[77] M. Jones, G. Marsden, Mohd-Nasir, Boone K. N., and G Buchanan. Improving web interaction on small displays. In *International Journal of Computer and Telecommunications Networking*, pages 1129–1137, 1999.

[78] Steve Jones, Matt Jones, and Shaleen Deo. Using keyphrases as search result surrogates on small screen devices. *Personal Ubiquitous Comput.*, 8(1):55–68, 2004.

[79] David Jonker, David Schroh Bill Wright, Pascale Proulx, and Brian Cort. Information triage with trist. In *Conference on Intelligence Analysis*, 2005.

[80] M. A. Just and P. A Carpenter. Eye fixations and cognitive processes. In *Cognitive Psychology.*, pages 441–480, 1976.

[81] Jacob R. J. K and K. S. Karn. Eye tracking in human-computer interaction and usability research: Ready to deliver the promise. In *Cognitive Psychology.*, pages 441–480, 1976.

[82] Sprck Jones Karen. A statistical interpretation of term specificity and its application in retrieval. In *Journal of Documentation*, pages 11–21,, 1996.

[83] Peter R. Keller and Mary M. Keller. *Visual Cues: Practical Data Visualization.* IEEE Computer Society Press, Los Alamitos, CA, USA, 1994.

[84] Diane Kelly and Nicholas J. Belkin. Reading time, scrolling and interaction: exploring implicit sources of user preferences for relevance feedback. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '01, pages 408–409, New York, NY, USA, 2001. ACM.

[85] Diane Kelly and Colleen Cool. The effects of topic familiarity on information search behavior. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 74–75, 2002.

[86] S Thomas Khun. *The Structure of scientific revolutions 2nd Edition.* University of Chicago Press, 1970.

[87] Robert R. Korfhage. To see, or not to see, is that the query? In *Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '91, pages 134–141, New York, NY, USA, 1991. ACM.

[88] Carol Collier Kuhlthau. A process approach to library skills instruction. In *School Library Media Quarterly*, pages 35–40, 1985.

[89] Carol Collier Kuhlthau. Developing a model of the library search process: Cognitive and affective aspects. In *Reference Quarterly*, pages 232–242, 1988.

[90] Carol Collier Kuhlthau. Inside the search process: Information seeking from the user's perspective. *JASIS*, 42(5):361–371, 1991.

[91] Carol Collier Kuhlthau. *Seeking Meaning. A Process Approach to Library and Information Services.* Ablex Publishing, 1996.

[92] K.L. Kwok, P. N. Deng Dinstl, H.L. Sun, W. Xu, P. Peng, and J. Doyon. Chine t: a chinese name finder system for document triage. In *Conf. on Intelligence Analysis*, 2005.

[93] David M. Levy. I read the news today, oh boy: reading and attention in digital libraries. In *Proceedings of the second ACM international conference on Digital libraries*, DL '97, pages 202–211, New York, NY, USA, 1997. ACM.

[94] Peter Liebscher and Gary Marchionini. Browse and analytical search strategies in a full-text cd-rom encyclopedia. In *School Library Media Quarterly*, pages 223–233, 1988.

[95] Ziming Liu. Reading behavior in the digital environment: Changes in reading behavior over the past ten years. In *Journal of Documentation*, pages 700–71, 2005.

[96] Fernando Loizides and George Buchanan. The myth of find: user behaviour and attitudes towards the basic search feature. In *Joint Conference on Digital Libraries*, pages 48–51, 2008.

[97] Fernando Loizides and George Buchanan. An empirical study of user navigation during document triage. In *Proceedings of the 13th European conference on Research and advanced technology for digital libraries*, ECDL'09, pages 138–149, Berlin, Heidelberg, 2009. Springer-Verlag.

[98] Fernando Loizides and George R. Buchanan. Performing document triage on small screen devices. part 1: structured documents. In *Proceeding of the third symposium on Information interaction in context*, IIiX '10, pages 341–346, New York, NY, USA, 2010. ACM.

[99] Miles Macleod. *Usability Evaluation in Industry: Performance Measurement and Ecological Validity.* London, Taylor and Francis, 1995.

[100] Gary Marchionini. *Information Seeking in Electronic Environments.* Cambridge University Press, 1995.

[101] Gary Marchionini and Ben Shneiderman. Finding facts vs. browsing knowledge in hypertext systems. In *IEEE Computer*, page 70, 1988.

[102] Kugel Peter Benenfeld Alan R. Marcus, Richard S. Catalog information and text as indicators of relevance. In *Journal of the American Society for Information Science*, pages 15–30, 1978.

[103] Catherine C. Marshall and Frank M. Shipman, III. Spatial hypertext and the practice of information triage. In *Proceedings of the eighth ACM conference on Hypertext*, pages 124–133, 1997.

[104] Catherine C. Marshall, Frank M. Shipman, III, and James H. Coombs. Viki: spatial hypertext supporting emergent structure. In *Proceedings of the 1994 ACM European conference on Hypermedia technology*, ECHT '94, pages 13–23, New York, NY, USA, 1994. ACM.

[105] Michael E. Masson. Cognitive processes in skimming stories. In *Journal of Experimental Psychology: Learning, Memory, and Cognition*, pages 400–417, 1982.

[106] J May and L A Tweedie. Modelling user performance in visually based interactions. In *People and computers VIII*, pages 95–110, 1994.

[107] Carol Bergfeld Mills and Linda J. Weldon. Reading text from computer screens. pages 329–357, 1987.

[108] Jolley Janina M. Mitchell Mark L. *Research Design Explained*. Harcourt Brace College Publishers, 1996.

[109] Stefano Mizzaro. Relevance: The whole history. *JASIS*, 48(9):810–832, 1997.

[110] Neema Moraveji. Improving video browsing with an eye-tracking evaluation of feature-based color bars. In *JCDL*, pages 49–50, 2004.

[111] Mauro Mosconi, Marco Porta, and Alice Ravarelli. On-line newspapers and multimedia content: an eye tracking study. In *Proceedings of the 26th annual ACM international conference on Design of communication*, SIGDOC '08, pages 55–64, New York, NY, USA, 2008. ACM.

[112] Mauro Mosconi, Marco Porta, and Alice Ravarelli. On-line newspapers and multimedia content: an eye tracking study. In *SIGDOC '08: Procs. 26th ACM international conference on Design of communication*, pages 55–64, New York, NY, USA, 2008. ACM.

[113] David Nicholas, Paul Huntington, Hamid R. Jamali, and Anthony Watkinson. The information seeking behaviour of the users of digital scholarly journals. pages 1345–1365, 2006.

[114] Donald Norman. *The Design of Everyday things*. Basic Books, 2002.

[115] Kenton O'Hara and Abigail Sellen. A comparison of reading paper and on-line documents. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 335–342, 1997.

[116] Kenton O'Hara, Fiona Smith, William Newman, and Abigail Sellen. Student readers' use of library documents: implications for library technologies. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 233–240, 1998.

[117] Tom Owen, George Buchanan, Parisa Eslambolchilar, and Fernando Loizides. Supporting early document navigation with semantic zooming. In *ICADL*, pages 168–178, 2010.

[118] Tom Owen, George Buchanan, Parisa Eslambolchilar, and Fernando Loizides. Supporting early document navigation with semantic zooming. In *ICADL*, pages 168–178, 2010.

[119] Bing Pan, Helene Hembrooke, Geri Gay, Laura A. Granka, Matthew K. Feusner, and Jill K. Newman. The determinants of web page viewing behavior: an eye-tracking study. In *ETRA*, pages 147–154, 2004.

[120] Taemin Kim Park. Toward a theory of user-based relevance: a call for a new paradigm of inquiry. *J. Am. Soc. Inf. Sci.*, 45:135–141, April 1994.

[121] P Pirolli and S Card. Information foraging. In *Psychological Review*, pages 643–675, 1999.

[122] J Pomerantz, S Nicholson, and R.D Lankes. Digital reference triage: An investigation using the delphi method into the factors influencing question routing and assignment. In *The Library Quarterly*, pages 103–120, 2006.

[123] H. L Poole. *Theories of the middle range*. Ablex, New Jersey, NJ, USA, 1985.

[124] William R. Reader and Stephen J. Payne. Allocating time across multiple texts: sampling and satisficing. *Hum.-Comput. Interact.*, 22:263–298, August 2007.

[125] Dillon A. Richardson, J. and C McKnight. The effect of display size on reading and manipulating electronic text. In *Contemporary Ergonomics*, pages 474–479, London, UK, 1989. Taylor and Francis.

[126] T.W Robbins, E.J Anderson, D.R Bradley A.C Barker, C Fearnyhough, R Henson, S.R Hudson, and A Baddeley. Working memory in chess. In *Memory and Cognition*, number 24, pages 83–89, 1996.

[127] George G. Robertson, Stuart K. Card, and Jack D. Mackinlay. Information visualization using 3d interactive animation. pages 57–71, 1993.

[128] 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.

[129] Tefko Saracevic. Comparative effect of titles, abstracts and full texts on relevance judgments. In *Procs. of the American Society for Information Science*, number 6, pages 293–299, 1969.

[130] Tefko Saracevic. Relevance: A review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, 26(6):321–343, November 1975.

[131] Tefko Saracevic. Modeling interaction in information retrieval (ir): A review and proposal. In *Proceedings of the ASIS Annual Meeting*, pages 3–9, 1996.

[132] Tefko Saracevic. Relevance reconsideredd 96. information science:integration in perspective. In *Proceedings of Second International Conference on Conceptions of Library and Information Science*, pages 201–218,, 1996.

[133] Linda Schamber. Relevance and information behavior. In *Annual Review of Information Science and Technology (ARIST)*, pages 3–48,, 1994.

[134] Linda Schamber, Michael Eisenberg, and Michael S. Nilan. A re-examination of relevance: toward a dynamic, situational definition. *Inf. Process. Manage.*, 26:755–776, November 1990.

[135] Moses Schwartz, Curtis Hash, and L. M. Liebrock. Term distribution visualizations with focus+context. In *Proceedings of the 2009 ACM symposium on Applied Computing*, pages 1792–1799, 2009.

[136] Frank Shipman, Robert Airhart, Haowei Hsieh, Preetam Maloor, J. Michael Moore, and Divya Shah. Visual and spatial communication and task organization using the visual knowledge builder. In *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '01, pages 260–269, New York, NY, USA, 2001. ACM.

[137] Frank Shipman, J. Michael Moore, Preetam Maloor, Haowei Hsieh, and Raghu Akkapeddi. Semantics happen: knowledge building in spatial hypertext. In *Proceedings of the thirteenth ACM conference on Hypertext and hypermedia*, HYPERTEXT '02, pages 25–34, New York, NY, USA, 2002. ACM.

[138] Frank Shipman, Morgan Price, Catherine C. Marshall, and Gene Golovchinsky. Identifying useful passages in documents based on annotation patterns. In *In Proceedings of ECDL03*, pages 101–112. Springer, 2003.

[139] Frank M. Shipman, III, Haowei Hsieh, Preetam Maloor, and J. Michael Moore. The visual knowledge builder: a second generation spatial hypertext. In *Proceedings of the 12th ACM conference on Hypertext and Hypermedia*, HYPERTEXT '01, pages 113–122, New York, NY, USA, 2001. ACM.

[140] Ben Shneiderman. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 3rd edition, 1997.

[141] H. A. Simon. Rational choice and the structure of the environment. In *Psychological Review*, pages 129–138, 1956.

[142] L. K. Son and J. Metcalfe. Metacognitive and control strategies in study-time allocation. 26:204–221+, 2000.

[143] Robert Spence. *Information Visualization: Design for Interaction (2nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2007.

[144] Amanda Spink, Bernard J. Jansen, Dietmar Wolfram, and Tefko Saracevic. From e-sex to e-commerce: Web search changes. *Computer*, 35:107–109, March 2002.

[145] Anselm Spoerri. Infocrystal: a visual tool for information retrieval & management. In *Proceedings of the second international conference on Information and knowledge management*, CIKM '93, pages 11–20, New York, NY, USA, 1993. ACM.

[146] A. Strauss and J. Corbin. *Basics of Qualitative Research.* Sage Publications, 1998.

[147] Don R. Swanson. Subjective versus objective relevance in bibliographic retrieval systems. In *Library Quarterly*, pages 389–398, 1986.

[148] Van Gog T., F. Paas, J. J. G. van Merrienboer, and P Witte. Uncovering the problem solving process: Cued retrospective reporting versus concurrent and retrospective reporting. In *Journal of Experimental Psychology: Applied*, pages 237–244, 2005.

[149] Harold Thimbleby. *Press On: Principles of Interaction Programming.* The MIT Press, 2010.

[150] Anastasios Tombros, Ian Ruthven, and Joemon M. Jose. How users assess web pages for information seeking. pages 327–344, 2005.

[151] William M. K Trochim. *The research methods knowledge base .* Atomic Dog Publishing; 2nd edition, 1999.

[152] Edward Tufte. *Envisioning information.* Graphics Press, Cheshire, CT, USA, 1990.

[153] Tadeusz Tyszka. Information and evaluation processes in decision making. In *Social and moral values: Individual and societal perspectives*, pages 175–193, 1989.

[154] Herre van Oostendorp and Christof van Nimwege. Locating information in an online newspaper. In *Journal of Computer-Mediated Communication*, 1998.

[155] John von Neumann and Oskar Morgenstern. *Theory of games and economic behavior.* Princeton University Press, 1944.

[156] Ellen M. Voorhees and Donna K. Harman. *TREC: Experiment and Evaluation in Information Retrieval.* MIT Press, 2005.

[157] Nina Wacholder, Lu Liu, and Ying-Hsang Liu. Selecting books: a performance-based study. In *Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries*, pages 337–337, 2006.

[158] Peiling Wang and Dagobert Soergel. A cognitive model of document use during a research project. study i. document selection. pages 115–133, 1998.

[159] Peiling Wang and Marilyn Domas White. A cognitive model of document use during a research project. study ii. decisions at the reading and citing stages. *J. Am. Soc. Inf. Sci.*, 50:98–114, February 1999.

[160] T.D. Wilson. Models in information behaviour research. In *Journal of Documentation*, pages 249–270,, 1999.

[161] J. A. Wise, J. J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: spatial analysis and interaction with information from text documents. In *Proceedings of the 1995 IEEE Symposium on Information Visualization*, pages 51–58, Washington, DC, USA, 1995. IEEE Computer Society.

[162] William Wright, David Schroh, Pascale Proulx, Alex Skaburskis, and Brian Cort. Advances in nSpace The Sandbox for Analysis. McLean, VA, 2005.

[163] M. C Yovits and C. R Foulk. Experiments and analysis of information use and value in a decision-making. In *Journal of the American Society for Information Science*, pages 63–81, 1985.

[164] Fernando Loizides Zhao Geng, Robert S. Laramee and George Buchanan. Visual analysis of document triage data. In *International Conference on Information Visualization Theory and Applications (IVAPP)*, pages 151–163, 2011.