Just-in-time Information Interfaces: A new Paradigm for Information Discovery and Exploration

Sven Laqua

First Supervisor: M. Angela Sasse
Second Supervisor: Robert Elliot Smith

A dissertation submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy of University College London

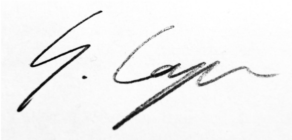
Revised thesis submitted in January 2015

Declaration

I, Sven Laqua, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources,

I confirm that this has been indicated in the thesis.



London, 29th January 2015.

Abstract

We live in a time of increasing information overload. Described as *"a byproduct of the lack of maturity of the information age"* (Spira & Goldes, 2007), information overload can be painful, and harm our concentration - the resulting choice overload impacts out decision-making abilities. Given the problem of information overload, and the unsatisfying nature of human-information interaction using traditional browsing or keyword-based search, this research investigates how the design of just-in-time information services can improve the user experience of goal-driven interactions with information.

This thesis explores the design of just-in-time information services through the iterative development of two strands of high-level prototypes (FMI and KnowDis). I custom-built both prototype systems for the respective evaluations, which have then been conducted as part of a series of lab-based eye-tracking studies (FMI) as well as two field studies (KnowDis). The lab-based eye-tracking studies were conducted with 100 participants measuring task performance, user satisfaction, and gaze behaviour. The lab studies found that the FMI prototype design did improve the performance aspect of the user experience for all participants and improved the usability aspect of the user experience for novice participants. However, the FMI prototype design seemed to be less effective and usable for expert participants. Two field studies were conducted as part of a two-year research collaboration, which lasted a total of 10 weeks and involved approximately 70 knowledge workers overall from across the globe. As part of those field studies, 46 semi-structured interviews were also conducted. The field studies found that the KnowDis prototype design did improve the user experience for participants overall by making work-related information search more efficient. However, while the KnowDis prototype design was useful for some knowledge workers and even integrated seamlessly into their day-to-day work, it appeared to be less useful and even distracting to others.

This thesis is wholeheartedly dedicated to my wife, Emily, and to my sons, Leonardo and Luca. Without their continuous energy, support and patience I would never have completed this work.

This thesis is also in memory of my grandparents, Paul and Ilse Laqua, who both passed away within days, just before and shortly after my submission of the original thesis. I knew they were ill and was meant to visit them back home in Germany shortly after submitting this thesis – may they rest in peace.

Foreword

“A scientist is supposed to have a complete and thorough knowledge, at first hand, of some subjects and, therefore, is usually expected not to write on any topic of which he is not a master. This is regarded as a matter of *noblesse oblige*. For the present purpose I beg to renounce the *noblesse*, if any, and to be the freed of the ensuing obligation.

My excuse is as follows: We have inherited from our forefathers the keen longing for unified, all-embracing knowledge. The very name given to the highest institutions of learning reminds us, that from antiquity to and throughout many centuries the *universal aspect* has been the only one to be given full credit.

But the spread, both in and width and depth, of the multifarious branches of knowledge by during the last hundred odd years has confronted us with a queer dilemma. We feel clearly that we are only now beginning to acquire reliable material for welding together the sum total of all that is known into a whole; but, on the other hand, it has become next to impossible for a single mind fully to command more than a small specialized portion of it.

I can see no other escape from this dilemma (lest our true aim be lost for ever) than that some of us should venture to embark on a synthesis of facts and theories, albeit with second-hand and incomplete knowledge of some of them - and at the risk of making fools of ourselves.”

Erwin Schrödinger, “What is Life?" (1944)

Table of Contents

List of Figures

List of Tables

1. Introduction

Life is a continuum of events and decisions. Each event and every decision is triggered by exposure to information, processing of information, and acting upon information. Sometimes we strive to plan for as many eventualities as we can, but life typically unfolds in a much more random, or serendipitous manner.

"Life is the sum of all your choices."
(Albert Camus)

Accepting this paradigm means that - besides all the planning one can handle - serendipity still shapes our lives, and impacts what we do to a large degree. The challenge in life is to manage serendipity to our advantage. Making good decisions can be seen as a way to maximise positive serendipitous encounters and minimise negative serendipitous encounters. But decision-making is neither easy nor straightforward. And for the most part of our lives – for the uncountable number of choices we face on a daily basis – decision-making occurs subconsciously (Zaltman, 2003) or with little reflection.

The key to good choices is making informed decisions, and the increasingly digital realm of human existence provides ever-increasing amounts of information. Over the past two decades, technology has transformed the way we create, access and manage information. The Internet does not just place unimaginable amounts of information at our fingertips; it also empowers each of us to contribute information through comments, tweets, blogs or a plethora of sharing and social tools. This shift has created an imbalance between the ease of information contribution and meaningful information seeking. Much of our daily exposure to information is noise that needs to be micro-managed. Yet technology should empower individuals to manage serendipity more effectively and facilitate positive serendipitous encounters.

Most recently, Garcia-Molina et al. (Garcia-Molina, Koutrika & Parameswaran, 2011) have argued that *"satisfying a user's information need"* is one of the most fundamental problems in computer science today. The authors suggest that *"the goal is to present to the user only information that is of interest and relevance, at the right place and time"* (Garcia-Molina, Koutrika & Parameswaran, 2011).

1.1. Personal Motivation

This thesis is a continuation of work on the Focus-Metaphor approach, a novel user interface concept, which evolved from my BSc and MSc theses.

The original idea for the Focus-Metaphor Interface (FMI) can be traced back to my BSc thesis on the *"Concept and User Interface Design for a Computer Supported Collaborative Learning Platform for Intercultural Communication"* (Laqua 2003). A first mock-up user interface (UI) was developed as part of this thesis. Subsequently, the idea substantiated in my MSc thesis on the *"Creation of Virtual Social Networks in Distanced, Informal Learning Settings through collaborative Story Writing – Implementation and Testing of a New Metaphor Prototype"* (Laqua 2004). Here, concept and UI were adjusted and enhanced and a more sophisticated high-level prototype was developed. Said prototype was evaluated online through participants from around the world, and furthermore tested in an eye tracking experiment. The results of this study were published in *"The Focus-Metaphor Approach: A Novel Concept for the Design of Adaptive and User-Centric Interfaces"* (Laqua & Brna, 2005).

This early iterative work on FMI prototypes laid the foundation for the paradigm that evolved in the research reported in this thesis through continued iterative design, development and evaluation.

1.2. Scientific Motivation

> *"What is a knowledge worker to do in a world where the Sunday edition of the New York Times has more information than the amount of information an average person alive 400 years ago might have come across in his lifetime?"* Spira & Goldes (Spira & Goldes, 2007)

The unimaginable amount of information available at our fingertips today is fuelling the problem that is ***information overload*** (Mooers, 1959; Eppler & Mengis, 2004; Wilson, 2005; Schwartz, 2005; Spira & Goldes, 2007; Fear, 2008; Claburn, 2009; Shanker & Richtel, 2011; Spira, 2011). However, as we can only process so much information at a time, information overload is only one side of the problem. In contrast to the immense growth of the overall amount of available information, the level at which we effectively process information is not changing that much: watching a music video on YouTube today, or watching tribes men chanting a song in 5000 BC – the amount of cognitive processing is fairly constant.

What has changed fundamentally is the level of choice people in the developed world are exposed to today. On the web in particular, the amount of information, the number of information sources, and the range of information services all contribute to ***choice overload*** (see section 2.2.7). Choosing the right information, making an informed decision and feeling confident about that decision has become exceptionally hard.

The amount of information and choices generated by society will only increase further, thus to help manage information overload more effectively, it is choice overload that needs to be addressed. Choice overload can lead to **poor decision-making**, or – even worse - **decision aversion**, which in psychology describes an individuals' growing *"tendency to avoid decision making"* as making a decision gets harder (Gerrig & Zimbardo, 2007).

On the web, link-based navigation is the embodiment of decision-making processes. Before the invention of hyperlinks, one could only browse digital documents by going backwards or forwards, just like turning pages in a book – it is the invention of hyperlinks that fundamentally changed how we interact with information.

While the original concept of a hyperlink – static author-created one-directional pointers – works reasonably well for well-organised information spaces of limited complexity, on a social web with 2 billion users, this concept is bound to fail. Today's ubiquitous reliance on search for much of the most mundane information retrieval tasks highlights the failure of static hyperlinks. In some way, much of today's use of keyword-based search - where a user manually specifies a navigation target - can be interpreted as a first step towards increasingly dynamic concepts of navigation.

In summary, this thesis investigates the problem of information overload, and its impact on information retrieval processes such as information discovery and exploration. The premise of this thesis is that neither traditional link-based browsing, nor keyword-based search are providing a satisfying enough user experience to navigate the ever-growing realm of available information.

1.3. Goal-driven Interactions with Information

In his book on *Information Foraging Theory* (IFT), Pirolli (Pirolli, 2007) argues that human interaction with information can be described through rational analysis. This assumes human behaviour to be intrinsically rational. It certainly is an admirable goal to strive for a rational model of human behaviour and worth pursuing, but when looking at the elements of analysis Pirolli lists for the various layers of explanation for human interaction with information, some doubts should be raised. According to Pirolli, IFT captures elements of rational analysis such as *states, resources, state dynamics, constraints, affordances, feasible strategies,* and *optimisation criteria* but lacks other analysis elements, such as *environment, goals, preferences, knowledge, perception,* and *action* for knowledge-level analysis, as well as *cognitive states* and *cognitive processes* for cognitive-level analysis (Pirolli, 2007).

In contrast to rational analysis, a knowledge-level analysis of interacting with information describes a system in terms of user intentions, which are shaped by user preferences, background knowledge and a motivating purpose (Pirolli, 2007). The critical aspect of knowledge-level analysis is the appreciation of the user as an individual, who acts within an environment and applies prior knowledge and preferences, to goal-driven behaviour such as information search tasks.

1.3.1. Iterative Support for Information Discovery and Exploration

This section outlines the iterative process of human interaction with information through a knowledge level analysis. It results in the formulation of the information goal continuum – a model of iterative support for information discovery and exploration.

In its essence, human-information interaction can be described simply as a person with an information goal retrieving some information from a (much) larger pool of available information (see Figure 1). In the digital realm, that pool of available information is more often than not the web – a cloud of data.

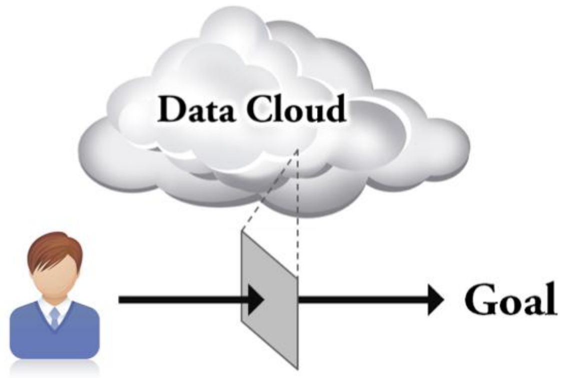


Figure 1: Simplified Model of Goal-driven Human Information Interaction

While the abstract representation of human information interaction in Figure 1 captures the general concept, it lacks the real-life complexity of most interactions with information. It is this simplified concept, which informs traditional keyword-based search. But embedded in this simple representation are two fundamental challenges that this thesis addresses.

1.3.1.1. The First Challenge

The process required to satisfy an information goal is typically (and increasingly) more complex than accessing a single entity of information. More often than not, information needs are open-ended or multi-faceted (White, Marchionini & Muresan, 2008), or people simply prefer orienteering strategies over keyword-based search (Teevan, Alvarado, Ackerman, *et al.*, 2004). Considering that the process of navigating is also a form of active learning (Schulmeister, 2007), cutting it short may not even be desirable.

Whether people engage in more complex information exploration processes due a preference for orienteering or an innate desire to learn, ultimately, as information goal grow increasingly complex, the number of entities of information required increases (see Figure 2). Bates describes the process of navigating or orienteering along a path of information entities motivated by an information goal as 'Berrypicking' (Bates, 1989).

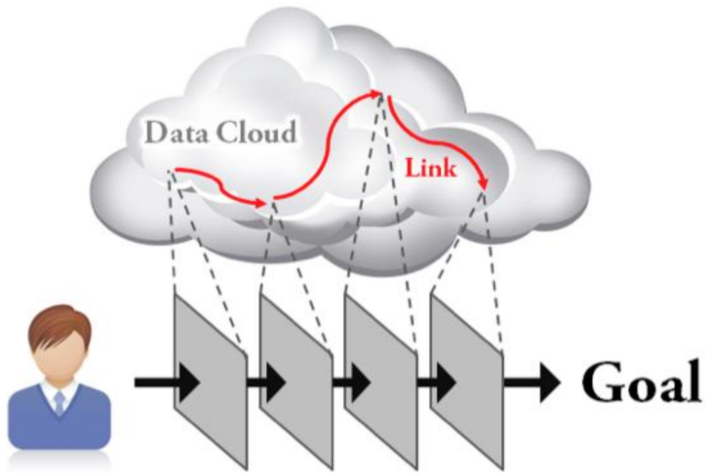


Figure 2: Model of Goal-driven Human Information Interaction

This process of navigating a cloud of data does not follow a simple model of information retrieval (Bates, 1989) or necessarily rational and thus predictable decision-making processes as implied by information foraging theory (Pirolli, 2007). Instead, people may encounter information (Erdelez, 1995) by muddling through (Hollnagel, 1992), or simply be driven by serendipity. As a result, their information needs may evolve over time and shift in direction, either altering existing goals, leaving them unfulfilled or re-prioritised (Hearst, 2009). Underlying these information experience processes is a fundamental problem with navigation on the web (see section 2.3.1) that frequently interferes with effective information exploration.

While navigating the web using hyperlinks has become one of the most innate activities for most people, the devil is in the detail, and its impact is experienced on a daily basis. More often than not a user starts engaging in one activity, to realise that a few minutes on she has drifted off to a completely different activity. This problem is rooted in a combination of external and internal aspects. Externally, information overload (see section 2.1.3) impacts our ability to concentrate (Wilson, 2005) or form decisions (Schwartz, 2005) by emerging from factors such as the innate complexity of the information, the tools used to process the information, or the tasks and processes needed to carry out with it (Eppler & Mengis, 2004). Internally, the problem of information overload is facilitated by our limited perceptual bandwidth (Reeves & Nass, 2000) to attend and process information (Wood, Cox & Cheng, 2006) and our inability to realistically attend to more than one task at a time (Norman, 1992).

While information foraging theory (Pirolli, 2007) and its usage of the concept of information scent (Chi, Pirolli, Chen, *et al.*, 2001) is highly related to this problem, IFT focuses on the process of (active) information search, and thus is more concerned with the more elementary model as described in Figure 1. Information scent is used to describe how well a user will be able to judge the value of information they will encounter when choosing a particular navigational option or path. But within the school of thought surrounding PARC there is growing acknowledgement that a more detailed understanding of interacting with information requires thinking in broader dimensions – how people make sense of information (Russell, Furnas, Stefik, *et al.*, 2008; Russell, Pirolli, Furnas, *et al.*, 2009; Stefik, 2004).

1.3.1.2. The Second Challenge

The essence of information relevant to an information goal is not simply presented in all clarity. Although search engines such as Google or Bing increasingly attempt to satisfy basic information needs immediately within a search result snippet, more complex information goals require a process of exploration and of gaining insight. A typical web page today represents a large chunk of data, which may contain useful information but surrounds it with a plethora of information that is mostly irrelevant in the context of people's immediate information goals. This dilemma of human-information interaction will be discussed in more detail in section 1.3.4, arguing that the structure of information encountered, the linking of that information, and its visual presentation are all potential sources of confusion. To retrieve the facets of information relevant to the user's goal, which may or may not be present within the chunk of data that makes up a web page, the user needs to first make sense of that data (see Figure 3). This cognitive process involves selective visual attention (Wood, Cox & Cheng, 2006), cognitive processing (Hollnagel, 1992) within a very limited working memory (Cowan, 2010), and decision-making (Simon, 1972). While the failure to make sense of information may be routed in an individual's inability to comprehend relevant information, I would argue that the primary cause for ineffective or failed information discovery and exploration processes is choice overload (Iyengar & Kamenica, 2007).

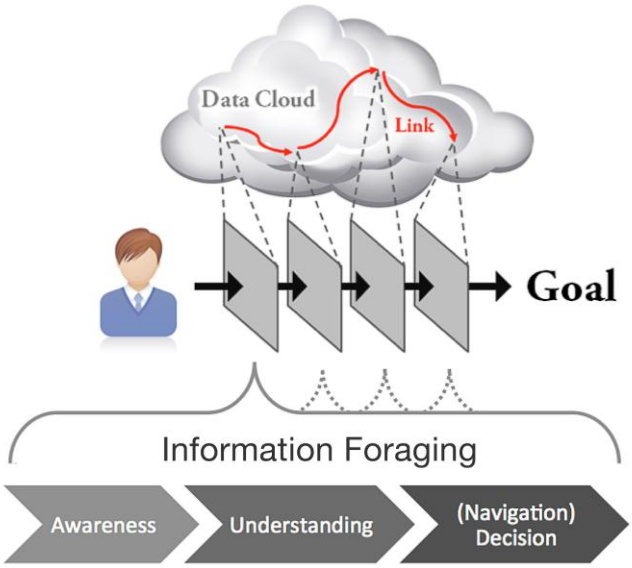


Figure 3: Information Foraging activities innate to information interactions

Information foraging activities require the user to become aware of a some data, understand it, and then make a decision on whether it is useful (e.g. and follow a link to further content) or shift attention to another bit of data (e.g. on the same page). As web pages typically present users with too many choices, too many bits of data to parse, effective foraging for information becomes hard, if not impossible. Research on choice overload (Iyengar & Kamenica, 2007; Fear, 2008; Schwartz, 2005) discusses this problem and why sometimes - less is more. Within IFT, the discussion of *"the ecology of information foraging"* (Pirolli, 2007) points to the same problem, suggesting that information is easier to make sense of if it *"resides within the same patch"* - information patches being used to describe groupings of information, e.g. on a web page.

Given humanity's increasing reliance on the web to support both ordinary and largely complex information problems, understanding and facilitating the effective exploration of relevant information may be one of the biggest challenges the web is facing today.

1.3.2. The Information Goal Continuum

On closer examination of the process of satisfying an information goal, one can describe the user's path towards that goal as a continuum of information discovery and exploration (see Figure 4). The user can take many paths, some of which will satisfy the initial goal, but not necessarily with the same outcome, as different bits of insight gathered along the path of the information goal continuum may have brought up different goal-specific facets.

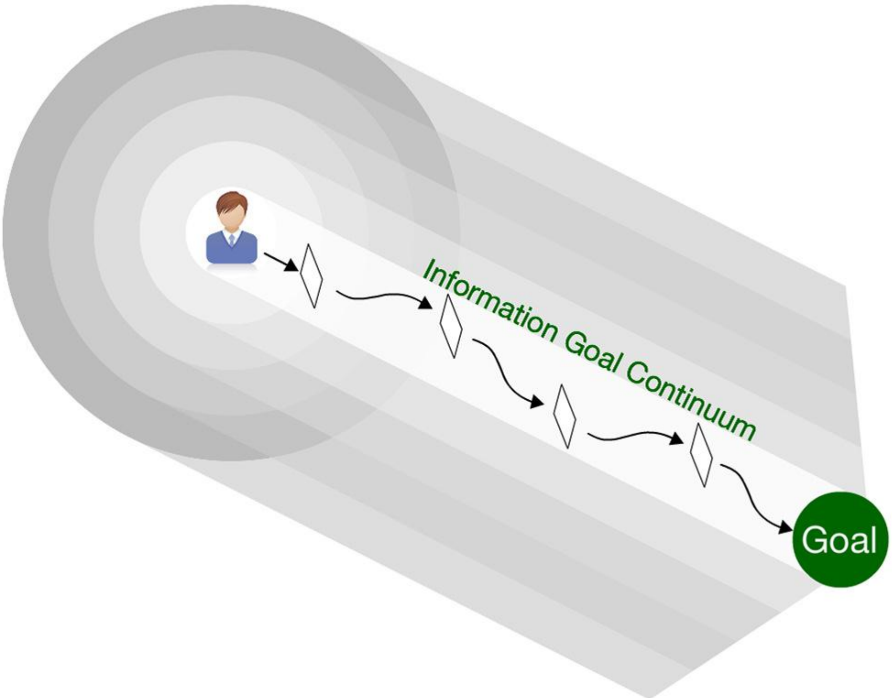


Figure 4: Information goal continuum for goal-driven information interactions

The user 'navigates' an information goal continuum when the iterative process of information foraging (awareness > understanding > decision) is sustained until an information goal is reached. When the user fails the reach their information goal by making wrong or ineffective navigation decisions – the user is leaving the information goal continuum. The failure to make sense of an information goal has been described earlier as a shift in attention (Hearst, 2009). The question is, whether not reaching the originally intended information goal has been intentional, or accidental. Within the continuum's model of iterative information discovery and exploration, the process of getting off track is very clearly defined as either one significant individual event of failed discovery or misguided exploration that makes the user leave the information goal continuum instantly (see Figure 5), or it could be the accumulation of multiple sub-optimal events of interactions with information that increasingly lead the user of track.

The main risk to information retrieval activities in general is that typically a large number of paths lead away from the initial goal, and thus out of the information goal continuum (see Figure 5). These 'bad paths' are represented by snippets of information, which elicit decision-making that is 'bad' in the context of a particular information goal at hand. The aspect of attempting to minimise the potential of 'bad paths' by generating links that are more relevant to a specific information goal is explored by research into adaptive navigation support (see section 2.4.2.2).

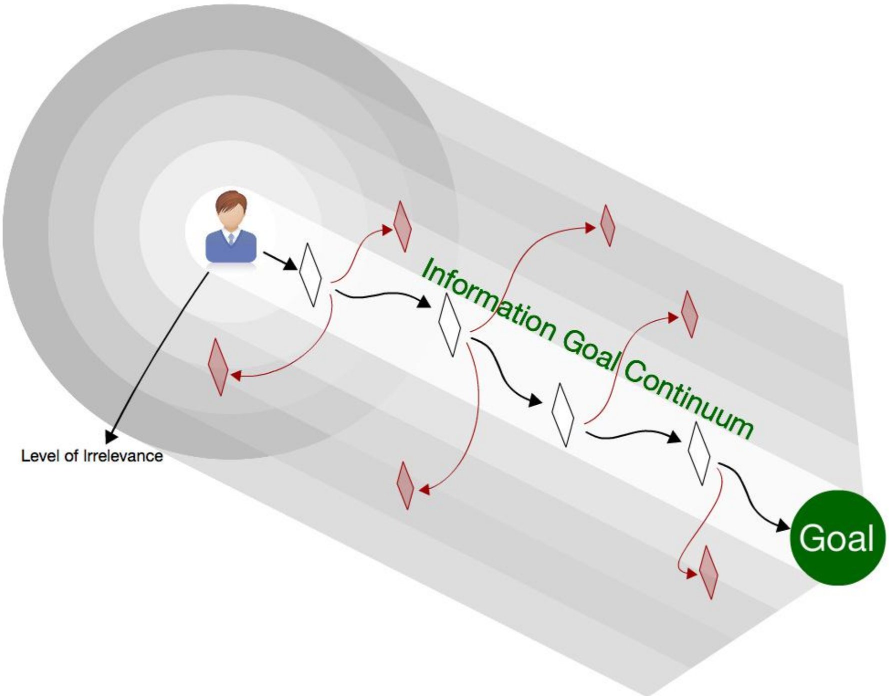


Figure 5: Impact of choice and decision making on navigating within a continuum of sense

On traditional web pages, each entity of relevant information is surrounded by a range of more or less irrelevant data. As the complexity of an information goal increases, the number of steps to make until this goal is satisfied typically increases as well. Consequently, the length of an information goal continuum grows, as does the likelihood to go astray before satisfying the information goal completely.

But the risk of leaving an information goal continuum is even higher when considering that users typically have a range of - at best complementary, at worst competing - information goals, which are vying for attention. For example a knowledge worker within an organisation may work on multiple similar projects concurrently. Those projects may require activities with overlapping information needs, such as conducting competitive analyses for multiple related products. As a result, the user may go astray towards side-goals that somewhat distract the user from the initial goal or worse lead the user completely off-goal. At best, such switches between related information goals incur task-switching costs.

In contrast, search engines such as Google do remove most extraneous information, and SERP-based interfaces work well for simple information goals that can be completed instantaneously in a single step (see Figure 6). However, the problem of the traditional search engine model is that it does not offer any direct support for complex - multi-step - information goals. The only options available to the user for a further exploration of information to reach a complex information goal are: (1) to go back and visit multiple search results from a single SERP, (2) to manually refine the keywords used in the hope to discover more related content through an alternative SERP, or (3) to fall-back into traditional browsing behaviour (see Figure 5).

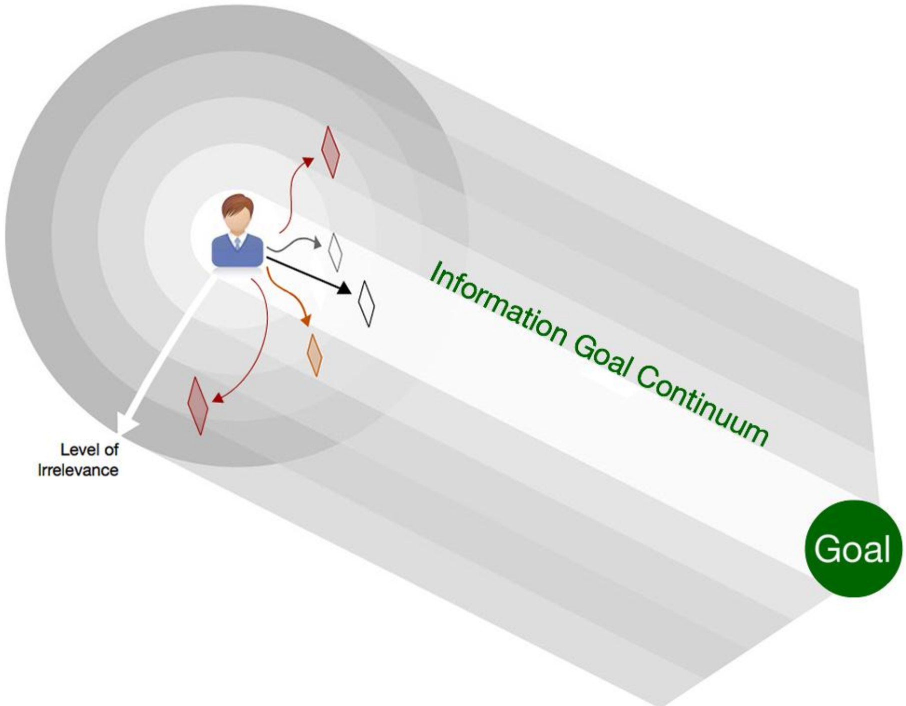


Figure 6: Model of an information goal continuum when using keyword-based search

A lot of existing research investigates the problems described above in isolation. Some research focuses on human aspects, investigating information overload (see section 2.1.2), human attention and its limitations (2.2.2), such as perceptual bandwidth (2.2.3), problems such as change blindness (2.2.4), and processes related to attention, such as decision making (2.2.5), the cost of task switching (2.2.6), choice overload (2.2.7), sensemaking (2.2.8).

Another strand of existing research focuses on human interaction with information and information systems through traditional browsing or search-based interaction (see section 0), investigating *exploratory search systems* (2.3.5), *focus+context interfaces* (2.3.6), or *context-aware information-retrieval* (2.3.7) among other things.

Finally, a further strand of existing research focuses on primarily computational aspects of how we interact with information (see section 2.4) by investigating how to make search systems *contextual* (2.4.1.1), *personalised* (2.4.1.2), or *proactive* (2.4.1.3). This strand also investigates the design of systems that provide *adaptive content* (2.4.2.1), *adaptive navigation* (2.4.2.2) or *adaptive personalisation* (2.4.2.3).

The work in this thesis has aimed to look at this vast range of existing research, borrow from its methods and its insights and derive some understanding as to how these often independent strands of research may be connected. This process has largely happened implicitly over the years of conducting this research. It has informed the design, development and evaluation of concrete prototype systems that will be described in subsequent sections and in particular in the study chapters. However, this process has also informed the explicit attempt to verbalize the effort of connecting aspects of the range of research discussed in this research. This verbalisation – of the connection of the human side, the interaction side, and the computational side of how we interact with information – has led to the *just-in-time information paradigm* and is outlined below.

1.3.3. Just-in-time (JIT) Information Paradigm

Neither traditional link-based browsing, nor keyword-based search are providing a satisfying enough user experience to navigate the ever-growing realm of available information in a goal-driven manner. While keyword-based search does work increasingly well for simple fact-based information needs (e.g. *"How old is London?"*), it does not support more complex information needs nearly as well (e.g. *"How can I live a satisfying life in London?"*). The latter seems like an information need that is too complex for simple keyword-based search. More likely, we would use a complex mix of keyword-based search, browsing, back and forth between information sources, to very gradually improve our understanding around this information need. This process would be tedious. It is not something we usually question, as we take keyword-based search and link-based browsing as the given technologies to navigate the web. But this process is one of diminishing returns. Yes, there are great resources (like blogs, or lifestyle sections in online magazines, etc.) that accumulate a reasonable amount of useful information. And we thus commonly settle quite happily for a particular source of information in connection with a particular information need. But there are two problems with this approach:

1) The more specific and complex the information need, the harder it is to identify a particular information source that serves this need well enough.

2) The amount of information available on the web is sheer endless (thus the perceived information overload), yet it is not easily accessible in any way other than using keyword-based search.

There thus must be a way to more easily access larger amounts of relevant information, than to settle on one or a few promising information sources and to ignore the rest of potentially useful information as the information retrieval process to get to that information would be too tedious.

So could there be another method for information interactions that might support the mitigation of perceived information and choice overload in order to support effective decision-making as part of information retrieval processes?

1.3.3.1. A Definition of the JIT Information Paradigm

The just-in-time information paradigm describes human-information interfaces and services that meet the following description:

> *At the point of interaction with information through a user interface for the purposes of satisfying an information goal (1) the amount of information relevant to an information goal should be maximised, and (2) the amount of information extraneous to an information goal should be minimised. Further, the user should be provided with relevant information at the right time – in a just-in-time manner. At the right time refers to the pro-active provision of information in a contextual manner that does not require the user to manually articulate (e.g. through keywords) what they are looking for. Instead, the just-in-time information interface should facilitate serendipitous encounters with useful information relating to the user's original information goal in an anticipatory manner.*

If the above conditions per definition of the just-in-time information paradigm are fulfilled – maximising relevant information and minimizing extraneous information – the signal-to-noise ratio of information pertinent to the user's information goal should be optimal. If applied in a rigorous and continuous manner, the just-in-time information paradigm should enable the user to effectively navigate an information goal continuum

– that is foraging for information in a continuous manner from one relevant bit of information to the next without distractions, until the information goal – simple or complex – has been satisfied (see Figure 7).

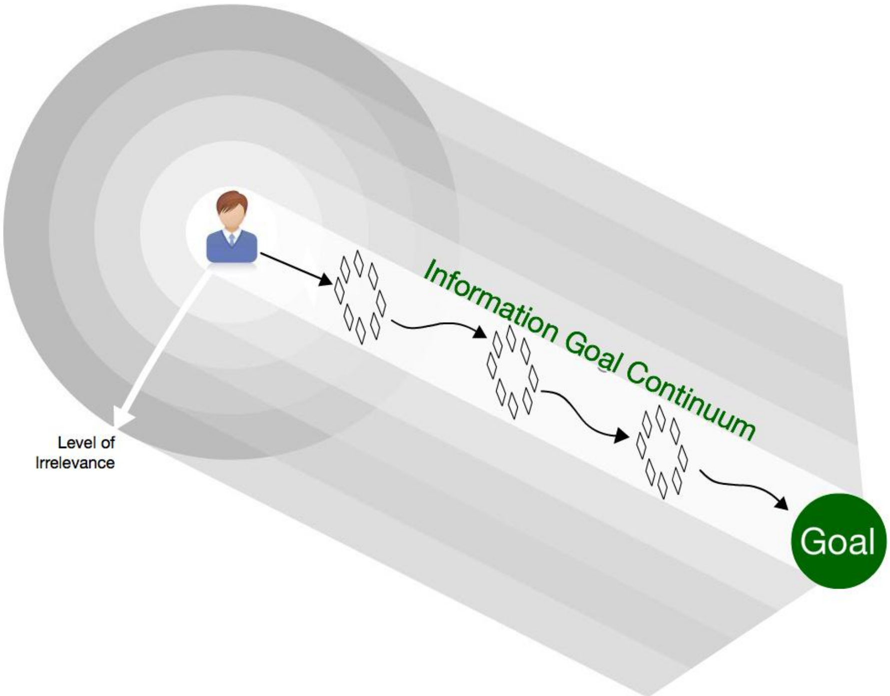


Figure 7: Ideal model of an information goal continuum when using a just-in-time information interface

It is worth noting that the just-in-time information paradigm does not intend to invent anything new really, but rather – in the spirit of a traditional design process – takes existing concepts, methods and technologies to combine them into something else, something that eventually becomes more than the sum of its parts.

1.3.4. The Human-Information Interaction Dilemma

From a user perspective, the Internet of today is a sheer endless universe of more or less useful information and services. Each user's unique perspective and usage context decides, which facets of this universe are interesting, and which are not. Without much guidance other than their own experience, users forage site after site (Pirolli, 2007), picking berries along the way (Bates, 1989), guided mostly by what catches their (immediate) attention. In many such episodes of digital information encountering as described by (Erdelez, 1995) users arrive at a point where they ask themselves what they actually wanted to look for in the beginning or at least have significantly deviated from their original goal.

Objectively, one could argue that information on the web is very extensively interconnected (via links). The problem is that from an individual's perspective - with the cognitive limitations as outlined in the human perspective (see section 2.2) and with a specific (complex) information need - the web appears very poorly connected. At any given moment in time whilst foraging (or berrypicking or encountering) the ratio of signal to noise in the range of available links - choices on how to progress in the information foraging process - is typically low.

A link is typically created with some intend by the link author, but the mental model of a user encountering said link may not necessarily be aligned with the mental model of the link author. In this context, one of the strengths of the content discovery paradigm of social networking sites such as Twitter or Facebook may well be the simple fact that the large amount of entropy an individual link is encoded in by its tweeters ensures that the likelihood of a match of the mental model of the user (who reads the tweet) and the mental model of the link author (who tweets a link) is maximised.

Information on the web is commonly wrapped into layers of an overall *information structure* (1), of a *navigational concept* (2) on how to link available information and of the *visual design* (3). Parsing these layers when trying to access relevant information during general purpose browsing requires time and cognitive effort at best, and at worst might mean that the user misses to identify relevant information. I refer to these layers as *layers of confusion* (Figure 8), as every user has to understand these layers to get to the relevant information (see .

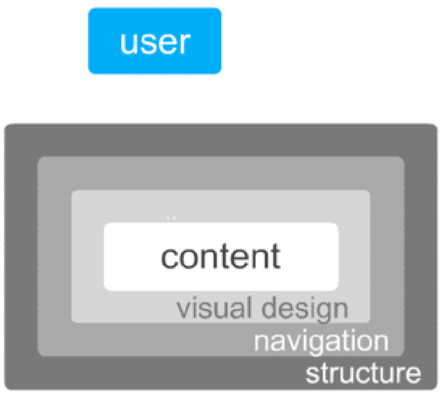


Figure 8: General purpose browsing

The problem users face when encountering these layers of confusion manifests in the universal use of keyword-based search. The reason why search engines are so popular and successful is that there is currently no other solution to provide easy access to any unknown content (coming from unfamiliar domains or content providers). Today, informational structures and navigational concepts have the scope a specific web site and very often they are not even consistent within that. The popularity of site searches highlights how even content provides increasingly struggle with the effective provision of transparent information structures.

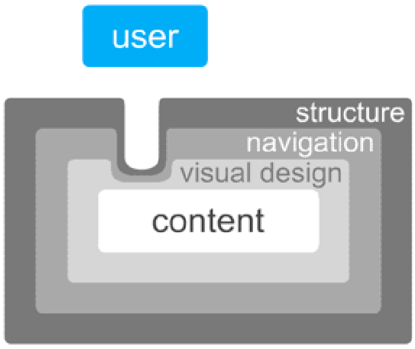


Figure 9: Directed search

When a user consults a search engine with the enquiry to help her find some specific information or service, what this search engine nowadays delivers is a list of web site links which ideally match what the user was looking for. In relation to the *layers of confusion* concept a search engine ideally allows the user to break through these layers (see Figure 9).

Depending on the accuracy of the search engine, the found pages contain the relevant information or they do not. The common ideal among today's search engines makers is a 'perfect match' - leading the user to exactly the right information. However, in reality perfect matches still consist of several layers of granularity, as highlighted by the model of *layers of search results quality* (see Figure 10).

Traditional search engines are focused on assisting users to break through the layer of structure at first. This reflects the general problem of bringing together different sites, or information across sites. Often, the search results lead to the homepage of a specific web site (Figure 10: search case 1). The more accurate the results of the search engine are, the further this allows the user to break through the next layer of navigation, as the user directly arrives at the page containing the desired information (Figure 10: search case 2). The visual layer then represents the last barrier, which the user has to pass in order to identify the desired information. Some search tools even can assist here, as they are able to highlight search terms in a page, thus minimising the need for the user to understand design and layout of a page (e.g. "in-page" search functionality embedded in modern browsers) and lead the user directly to the relevant information within a page (Figure 10: search case 3).

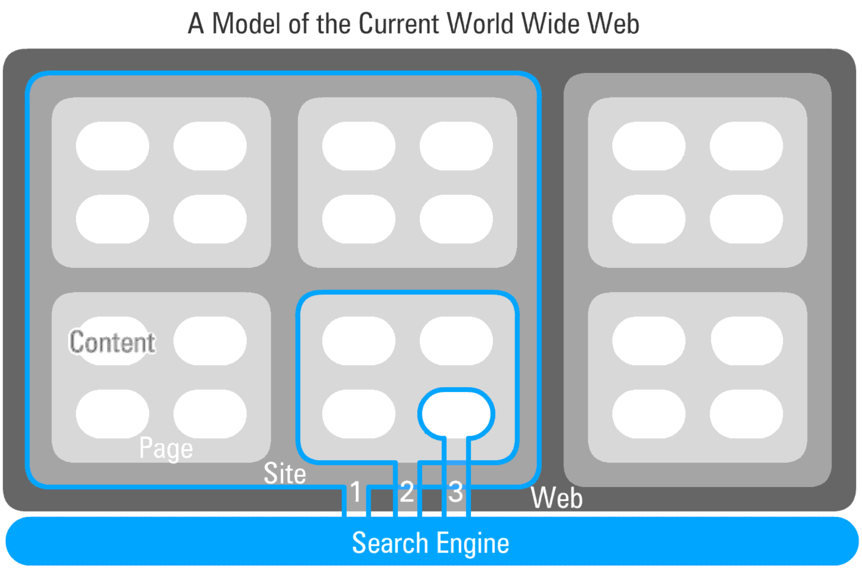


Figure 10: Layers of search result quality - a model of search result matching on the web (numbers 1,2 and 3 represent the different search cases)

This method of using search engines – undoubtedly the most prominent form of ‘exploring’ the Web as a whole – comes at a price: Mastering the layers of confusion through a search engine lets users *teleport* (Teevan, Alvarado, Ackerman, *et al.*, 2004) to individual nuggets of information. But more complex tasks require users to get a broader view. Often, users want to consult different sources and understand contexts. And depending on the information need, a single resource may not be sufficient. As a result, users ‘harvest’ one site after the other and crawl for information much like a search engine robot. Interestingly, the popularity of RSS feeds and other information aggregators underlines the fundamental problem with information interaction on the web. In essence, a feed reader acts as an agent to the user, harvesting potentially relevant information in an automated manner.

In the context of search engine use, the user provides some upfront cost by articulating search terms to be used. A search engine then provides search engine result pages (SERPs), enabled by automated search engine robots who have crawled the web and indexed all encountered information in place of the user. This enables users to crawl for an answer to their information needs on the SERPs themselves - a small subset of all the potentially relevant information on the web (see *search engine interaction model*, Figure 11).

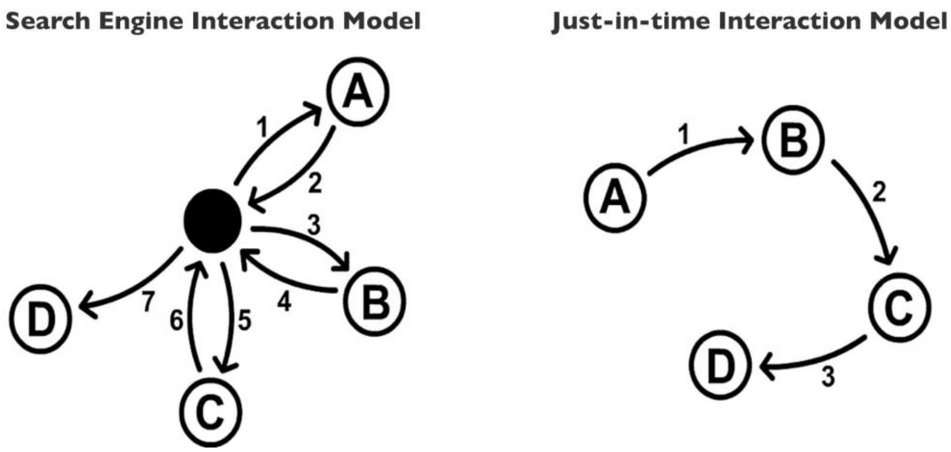


Figure 11: Comparison of the Search Engine interaction model with the FMI model

This method of interaction can be characterised as trial and error, run by every search engine user on the list of given results. The search engine interaction model visualises the number of steps a user has to go to arrive at a desired location "D" (see Figure 11, left). When assuming that users will in average find what they are looking for with the fourth link, this generates seven navigational steps - always from the search engine results to the listed page and back to the search results. This interaction is very distracting for users and often this activity is abandoned after only a few tried alternatives. Users might try other search terms, other search engines or even switch to a different (explorative) search strategy. Existing research confirms users prefer to browse explore information rather than hope for the right search results to turn up (Teevan, Alvarado, Ackerman, et al., 2004).

In contrast, when exploring information as described by the just-in-time information paradigm (see Figure 7), the same search scenario would follow a just-in-time interaction model (Figure 11, right), which les the user maintain other search result previews in context while interacting seamlessly with the currently selected content.

1.4. Research Question

Given the problem of information overload (outlined above and discussed in detail in section 2.1.2), and the unsatisfying nature of human-information interaction using traditional browsing or keyword-based search, this research investigates the following question: *How to design just-in-time information services to improve the user experience of goal-driven interactions with information*?

In order to tackle this question, this research explored the possibility of iteratively evolving the concept of Focus-Metaphor Interfaces (FMI) in a manner that supports the just-in-time information paradigm. The research thus started to investigate the following question:

(RQ 1) For a Focus-Metaphor Interface (FMI), what are the effects of dynamically updating contextual elements during an information exploration task?

RQ 1 investigates how extending the original FMI concept (Laqua & Brna, 2005) to accommodate larger information spaces (see Chapter 4 – Preliminary Study 1). As the original FMI only utilised static contextual elements and does not contain any other navigational elements, this study has been vital to understand users' experiences of information exploration with an FMI when dynamically updating contextual elements. This study could have revealed that such an approach is unusable, in which case an alternative design approach would have to be chosen, going forward in this thesis.

As over time, input modalities change, this thesis also aimed to understand how a just-in-time information paradigm might accommodate information interactions by novel means such as touch or eye-gaze. As this research was conducted years before the release of the first iPad, and eye-tracking was a readily available technology in our lab, the natural conclusion was to evaluate the FMI's usability for eye-gazed interaction[1]. This thesis thus explored the following research question next:

(RQ 2) For a Focus-Metaphor Interface (FMI), what are the effects of selection by gaze based on dynamic dwell times on user-preference? And specifically, are there any differences in user preference between the two implementations of dynamic dwell times – 'static interest accumulation' and 'dynamic interest decay'?

RQ 2 investigates whether the use dynamic dwell times provides a usable approach to eye-gaze interaction with an FMI (see Chapter 5 – Preliminary Study 2). In particular, it evaluates two specific algorithms that implement dynamic dwell times - 'static interest accumulation' and 'dynamic interest decay'. The prototype built for this study used the same original FMI concept, but adopted for gaze-based interaction – it was thus called GazeSpace. If the evaluation of GazeSpace would have suggested poor usability of the

[1] The FMI was later also evaluated for touch-based interaction in (Beeharee, Laqua & Sasse, 2011), however this work does not form part of this thesis and is discussed in the Future Work section.

system, this could have hinted at problems with the interface itself, the implemented input modality, or a combination of those factors. However, the study found the GazeSpace implementation of the FMI to be highly useable.

As the results of preliminary study 1 and preliminary study 2 were promising, another iteration of the FMI was built using full-text similarity matching (see section 0) and then evaluated in a large lab-based study to answer the following research question:

(RQ 3) For a Focus-Metaphor Interface (FMI), what are the effects of interaction-driven dynamic updating of contextual elements on task performance and user preference; and how does user interaction behaviour differ?

RQ 3 investigates the task performance and user preference of the new FMI prototype by comparing it to a traditional web-based UI using some specific hypotheses (see Chapter 6 – Main Study 1). This study uses a larger corpus of information than the previous studies (160 blog articles on healthy living) and mix of information search and information exploration tasks. The primary goal of this study was to understand how well the FMI would perform for information search tasks. If it performed poorly, this would suggest that the minimalist visualisation and interaction paradigm used in the FMI would not support the just-in-time information paradigm as hoped. However, as hypothesized, the FMI outperformed the traditional web-based UI. Somewhat surprisingly though, results for user preference were more mixed. In particular, the study suggests that the FMI is more beneficial to novice users than it is to experts. This could be related to more primed mental models in expert users on how to conduct information retrieval tasks and will always be a challenge when testing a new UI concept in the lab, in a somewhat abstract context. However, it consequently seemed desirable to evaluate the just-in-time concepts utilized in the FMI prototypes in a more real-world context and in a more longitudinal evaluation with a stronger focus on expert users. This approach was facilitated through an 18-month research collaboration with a large IT organisation. During this collaboration, a second just-in-time information prototype (KnowDis) was iteratively developed and evaluated through two field studies. For these field studies, KnowDis was deployed with (expert) knowledge workers of the IT organisation and used over the course of several weeks exploring the following research question:

(RQ 4) How do users respond to embedded proactive search in an email application?

As part of the investigation of RQ 4, a number of more specific research questions are also investigated to understand whether KnowDis is useful or distracting, and whether KnowDis has any perceived impact on knowledge workers' efficiency (see Chapter 7 – Main Study 2):

(RQ 4.1) Do users find having a proactive search tool embedded in their email application useful or not? If not, why not? If yes, how is it useful and how does it integrate with their day-to-day work?

(RQ 4.2) Do users find proactive search features distracting? If yes, how is it distracting? What can be done differently to make it less distracting?

(RQ 4.3) Do users think their work-related tasks that depend upon information search become more efficient and more effective when proactive search tools are available?

The primarily qualitative research work conducted as part of the investigation of RQ 4 was conducted in two phases (see section 7.4), involving 46 interviews with knowledge workers, the analysis of usage data logs gathered during the field studies as well as pre- and post-questionnaires. While it was initially intended to also integrate aspects of the FMI into the KnowDis prototype to support not just information discovery, but also information exploration, this work was ultimately out of scope given time and resource constraints[2].

1.4.1. How do FMI and KnowDis relate to the Just-in-time Information Paradigm?

Two complete prototype systems have been developed and evaluated – FMI and KnowDis - that strive to follow the design guidelines for just-in-time information services (see section 1.3.3) as much as possible. Some constraints, such as time,

[2] Some conceptual work was conducted on what a KnowDis prototype integrating aspects of the FMI could look like. Such an integration, and what it might look like is discussed in the Future Work section.

resources and the concrete research questions being investigated, had to be taken into account when designing the specific just-in-time prototypes.

The FMI prototype evolved over three years iteratively through design, development and evaluation in several lab-based studies – specifically preliminary study 1 (see chapter 4), preliminary study 2 (see chapter 0), and main study 1 (see chapter 6). The FMI prototype development focused on a quite literal interpretation of the assessing whether the second design guideline for just-in-time information services – *“minimizing the information extraneous to an information goal”* – is achievable, while maintaining users’ ability to effectively reach specific information goals. The FMI prototype also utilizes pro-active and contextual delivery of information to maximise the provision of information relevant to an information goal. It is important to note that the iterative design of the various FMI prototypes reflects an empirical exploration of the just-in-time information paradigm. As such, the iterations of the FMI represent proof-of-concept prototypes. By removing all unnecessary user interface elements, the FMI prototype represents the most essential interpretation of the just-in-time information paradigm. Its minimalist design may require augmentation for more widespread real-world use.

The KnowDis prototype evolved over two years iteratively through design, development and evaluation in two field studies within a large IT organisation – both field studies are reported in chapter 7. The KnowDis prototype development focused on assessing how a just-in-time information interface could be integrated into the work context of knowledge workers to actively support their information goals. Given the real-world constraint of embedding information discovery into an existing email work context (see section 7.6 on design decisions), the KnowDis implementation of the just-in-time information paradigm focused on design guideline one - maximising the information relevant to a knowledge worker’s information goals through proactive, contextualised recommendations of potentially relevant information. While it would have been desirable to also account for the second design guideline as well - minimizing the amount of information extraneous to an information goal – this would have required a level of development to customize the email client used within the organisation that was simply out of scope for this research project.

1.5. Thesis Contribution and Related Publications

The just-in-time information paradigm has been formulated as part of this thesis to help articulate and frame the theoretical and conceptual work that went into the design, development and evaluation of the FMI and KnowDis prototype systems. Both prototype systems are instantiations of the just-in-time information paradigm as discussed in section 1.4.1. Through the studies reported in this thesis, the prototype systems have demonstrated specific improvements to the user experience for the respective domains and contexts of use they were designed for and evaluated in. In extension, these studies provide early validations for how to design just-in-time information services to improve the user experience of goal-driven interactions with information.

The FMI prototype demonstrated that it can be mapped onto reasonably large information spaces, and through dynamic contextualisation of its navigational elements provide significantly better task performance for information search tasks than a traditional blog-based web interface (see section 0). The KnowDis prototype demonstrated that it can be successfully used by knowledge workers as alongside their day-to-day work and make work-related tasks that depend upon information search more efficient (see section 7.11.3).

A more detailed discussion of the contributions made in this thesis can be found in the conclusions (chapter 8). The final chapter provides a more detailed reflection on the specific research findings (see section 8.2), as well as a discussion of more general contributions to the HCI community (see section 8.3) – such as the just-in-time information paradigm, new algorithms for gaze-based interaction, and advice for future PhD students grounded in reflections on the research process followed in this thesis.

The research in this thesis has resulted in a number of publications at ACM CHI, BCS HCI, and the German Chapter of ACM (Mensch & Computer). The following table (see table 1) provides an overview of the publications that are part of this research, and which part of the thesis they relate to.

Table 1: List of publications relating to thesis (sorted in reverse chronological order)

Publication	Contribution	In Thesis
Laqua, S., Sasse, M.A., Gates, C., and Greenspan, S. (2011). *Do you KnowDis? A User Study of a Knowledge Discovery Tool for Organizations.* In Proc. of CHI 2011, Vancouver, BC, Canada.	Study 2	Chapter 7
Beeharee, A.K., Laqua, S., and Sasse, M.A. (2011). *Navigating Haystacks at 70mph: Intelligent Search for Intelligent In-Car Services.* In: Proc. of MIAA workshop at IUI 2011, Palo Alto, California, US.	Future Work (In-car systems)	Chapter 8
Laqua, S., Sasse, M.A., Gates, C., and Greenspan. S. (2009). *Making Sense of the Unknown: Knowledge Dissemination in Organizations*. In: Sensemaking Workshop, CHI 2009, Boston, MA.	Study 2	Chapter 7
Laqua, S., and Sasse, M.A. (2009). *Exploring Blog Spaces: A Study of Blog Reading Experiences using Dynamic Contextual Displays*. In Proc. of HCI 2009, 1-5 Sept. 2009, Cambridge, UK.	Study 1	Chapter 6
Laqua, S., Bandara, S.U., and Sasse, M.A. (2007). *GazeSpace: Eye Gaze Controlled Content Spaces.* In Proc. of HCI 2007, 3-7 Sept. 2007, Lancaster, UK.	Preliminary Study 2	Chapter 5
Laqua, S., Patel, G., and Sasse, M.A. (2006). *Personalised Focus-Metaphor Interfaces: An Eye Tracking Study on User Confusion.* In Proc. of Mensch und Computer 2006 , 3-6 September, Gelsenkirchen, Germany. (German Chapter of ACM)	Preliminary Study 1	Chapter 4

1.6. Structure of the Thesis

Chapter 1 provided an introduction to this thesis. It briefly explores the context of information overload to illustrate the motivation for the research conducted. This chapter then provides an introduction to the just-in-time information paradigm and the motivations behind it. The chapter then introduces the main research question investigated by this thesis, and discusses the more specific underlying research questions addressed by the individual studies. Chapter 1 concludes with a brief overview of the contributions made in this thesis as well as a list of publications related to this research.

Chapter 2 provides an extensive critical review of relevant literature**,** starting with a discussion of the information age (see section 2.1), in which aspects of human-information interaction are outlined and the problem of information overload is analysed. This chapter then analyses the human perspective of the information overload problem (see section 2.2), focussing on aspects of information processing, decision-making and sensemaking. Next, chapter 2 discusses the interaction perspective of the information overload problem (see section 2.3), specifically analysing the main categories of web-based navigation - general purpose browsing and directed search. Section 2.3 also reviews focus + context interfaces, and concludes with a critical analysis of the fundamental problems with web-based navigation. This chapter concludes with a discussion of the computational perspective of the information overload problem (see section 2.4). In this section, relevant technologies used for search systems, adaptive systems and personalisation are reviewed and an educational reasoning for personalisation is outlined.

Chapter 3 provides a discussion of relevant research methods. It briefly discusses the relationship of qualitative and quantitative approaches, outlines the range of prototyping methods available. This chapter then provides an overview of eye tracking technologies and relevant measures used in HCI research. After a discussion of evaluation approaches used within similar domains, this chapter concludes with brief overview of the methods employed in this thesis. The details of the particular method employed in each study are discussed in the respective study chapters.

Chapter 4 reports on the first preliminary design validation study investigating the effects of dynamically updating contextual elements during an information exploration task for a Focus-Metaphor interface (FMI). Following up early work on the Focus-Metaphor Approach (Laqua & Brna, 2005), this study explores the impact of maintaining a minimalistic visualization style - a central content element, seven context items, and no further navigational elements - for information spaces with larger amounts of content than the original static prototype that was used in (Laqua & Brna, 2005).

Chapter 5 reports on a second preliminary study in which the effects of selection by gaze based on dynamic dwell times on user-preference for a Focus-Metaphor Interface are investigated. For this study, I developed new algorithms aimed at improving the user experience when selecting information rich user interface elements using eye gaze (see section 5.4.3).

Chapter 6 reports on Main Study 1 - an eye tracking study conducted with 60 participants - to gain an understanding of how people interact with blog environments. This study investigated the effects of interaction-driven dynamic updating of contextual elements on task performance and user preference for a Focus-Metaphor Interface. Main Study 1 also explored how user interaction behaviour differs when using an FMI.

Chapter 7 reports on Main Study 2 – a two-year collaborative research project conducted as part of this thesis with CA Labs of Computer Associates. As part of this collaboration a knowledge discovery system (KnowDis) for organisations was iteratively designed, developed and evaluated through 2 field studies to explore how users respond to embedded proactive search in an email application. This study found that KnowDis can improve the user experience for participants overall by making work-related information search more efficient. The study also explores in detail different types of knowledge workers and how their perceptions of KnowDis differs.

Chapter 8 summarises the findings of this research, discusses its contributions to the field of human-computer interaction (HCI) and explores potential future work.

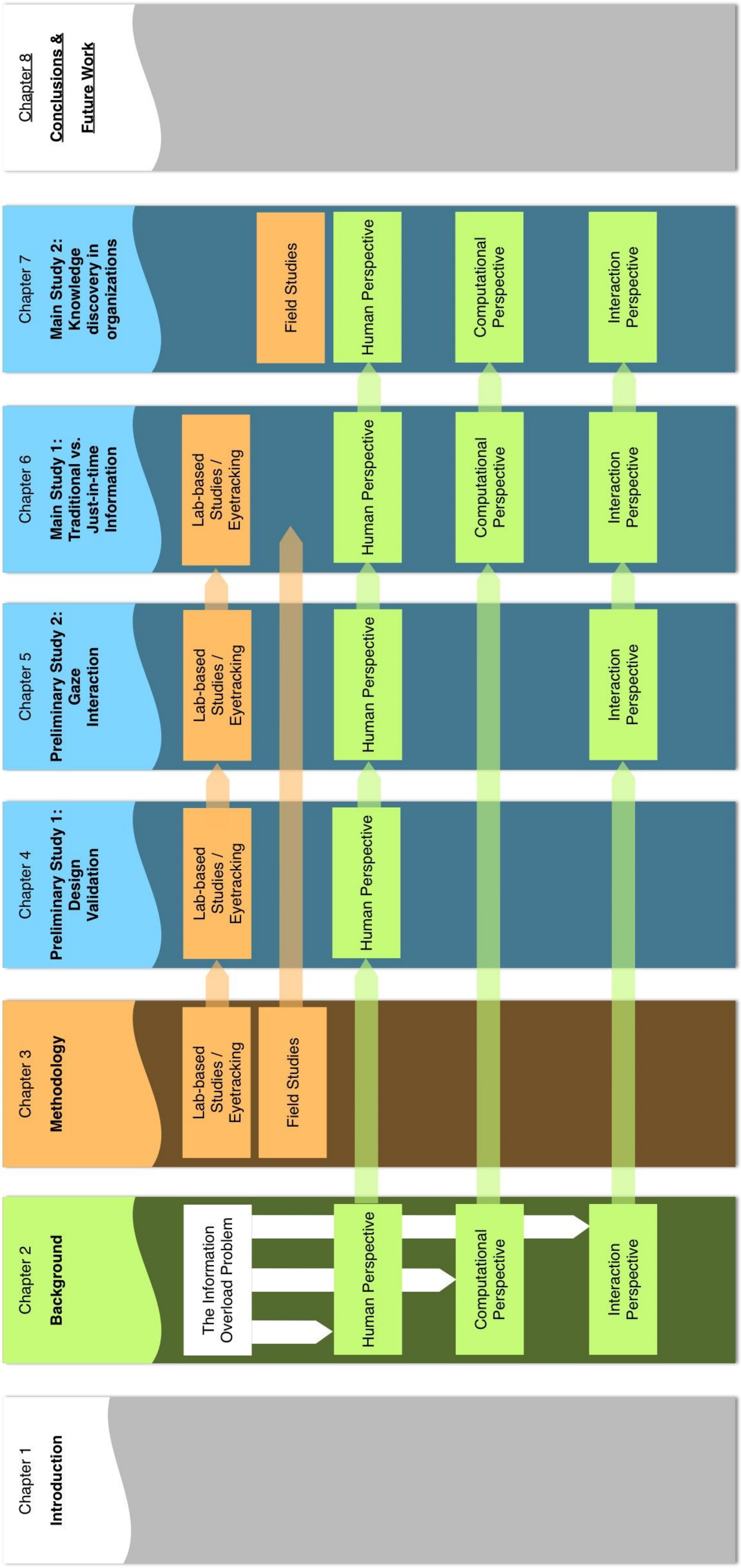


Figure 12. Structure of the thesis

2. Background

The growing problem of information overload is the main motivation in this thesis. In part, this problem can be attributed to an explosion of available information. But this thesis takes a different perspective on the information overload problem: it suggests that the way we have grown accustomed to accessing digital information through general purpose browsing and keyword-based search is increasingly ineffective as our information needs grow more complex.

As such, the research reported in this thesis investigates how to design just-in-time information services that can improve the user experience of goal-driven interactions with information. Consequently, the main goal of the review and analysis of relevant background literature is to provide a theoretical grounding for the just-in-time information paradigm as well as the prototype systems that have been developed and evaluated to validate the paradigm. As the just-in-time information paradigm draws from a range of research disciplines such as human-computer interaction (HCI), cognitive psychology, educational theory, information science, as well as aspects of machine learning (ML), the following review and analysis of the relevant literature has to be broad. In the attempt - to use Schroedinger's words - *not to make a fool of myself,* the relevant literature is also reviewed and analysed with considerable depth, where it is deemed relevant.

Over the years it took to carry out and write up this research, facets of the just-in-time information paradigm have started to emerge on the Web. These facets are linked to concepts such as personalised search, recommendation systems, etc. whose origins predate this thesis. This thesis does not attempt to improve upon any of those concepts in particular; rather, it takes a higher-level view and explores the possibility of weaving together various relevant concepts and techniques to address the problem of information overload at a more fundamental level through the proposed paradigm, and in particular through the design, development and evaluation of specific prototype implementations of the just-in-time information paradigm.

In Section 2.1 - "The Information Age" – I discuss the shift in information creation, management, and retrieval and outlines the complexity of the problem that is information overload.

Sections 2.2, 2.3 and 2.4 explore relevant literature from three different perspectives to contribute to a more holistic understanding of information overload.

In section 2.2 - "The Human Perspective" of information overload - concepts such as perception, awareness, cognition and decision making are reviewed and analysed to understand how people try to make sense of information when dealing with information tasks.

In section 2.3 - "The Interaction Perspective" of information overload - techniques for human information interaction with a focus on contextualised information visualization techniques and search interaction are reviewed and analysed.

In section 2.4 - "The Computational Perspective" of information overload - technologies such as search systems, tag-based systems and the semantic web and links them to techniques such as adaptation, personalisation, and recommendation are reviewed and analysed.

2.1. The Information Age

A number of landmark papers shaped the information age. From Shannon's Mathematical Theory of Communication in 1948 (Shannon, 1948), via Simon's work on Rational Decision Making (Simon, 1979), to Weick's *"Sensemaking in Organizations"* (Weick, 1995) our understanding of information, and how to use it effectively, has evolved significantly. However, on-going technological progress is re-defining our relationship with information. In ever-shorter cycles, new ways to create, and to share information surface, while the amount of information available at our fingertips increases exponentially. Yet, our cognitive bandwidth to process information stays limited by natural physiological constraints (Cowan, 2001).

2.1.1. Defining Information

"The problem is that we still have to agree about what information is exactly" -
Luciano Floridi

A multitude of definitions of information exist, such as (Shannon, 1948; Stonier, 1991; Floridi, 2004). These theories evolved alongside technological and scientific progress, but are often pragmatically simplified to fit the context in which information as an entity is being discussed. An interpretation of information in the context of biology is different from an interpretation of information in the context of computer networks, or philosophy - yet different perspectives may be complimentary.

The following discourse of information strives to find a balance between breadth, detail and pragmatism to set the scene for the subsequent discussions of applied information use, information processing, and information management. As no general agreement on a universally accepted definition exists, it is deemed more vital to convey the perspective on information applied in this thesis.

2.1.1.1. A Mathematical View of Information

Shannon's "Mathematical Theory of Communication" (Shannon, 1948) is the foundation for the field of information theory. Shannon introduces *"the effect of noise in a channel"* to extend the general theory of communication (see Figure 13). While a discussion of the technical details of the theory is outside the scope and focus of this thesis, Shannon's discussion of *choice*, *uncertainty* and *entropy* is nonetheless of remarkable relevance (see page 10 in Shannon, 1948). Following his description of an information source as a stochastic process (or discrete Markov process) - *"a physical system, or a mathematical model of a system which produces … a sequence of symbols governed by a set of probabilities"*, where the information or message is generated "symbol by symbol", Shannon discusses the impact of choice, uncertainty and entropy on the process of information production (or generation).

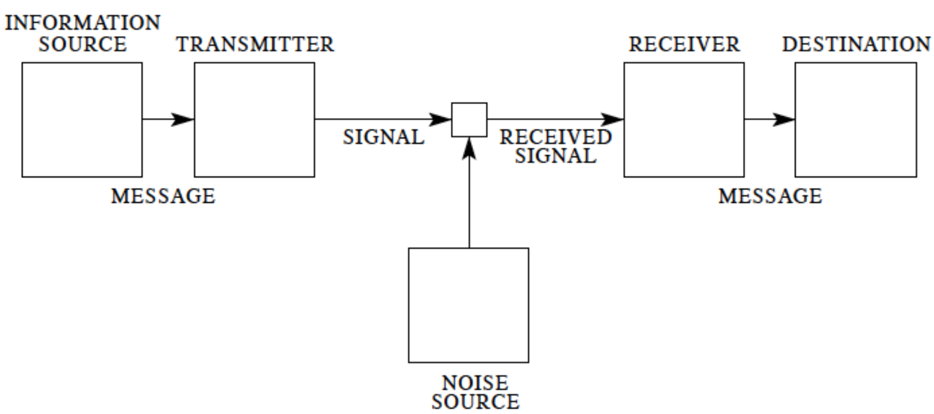


Figure 13: Schematic diagram of a general communication system (Shannon 1948)

2.1.1.2. A Philosophical View of Information

In contrast to the mathematical view of information as (well-formed) data communicated through a stochastic process, a philosophical view of information is concerned with the semantics of information - when data becomes meaningful (Floridi, 2011). Such *semantic content* can be instructional or factual (see Figure 14).

This distinction between information as well-formed data and semantic content is necessary to illustrate that information without any context - without human interpretation, based on some sort of background knowledge - is meaningless. This may seem like a trivial conclusion - as soon as a person looks at data, it will have to apply some basic or complex interpretation of it, thus giving at least some meaning to it. Yet it also highlights that the vast amount of information that exists somewhere (e.g. within a returned 1,000,000 search result hits for a keyword search query), but which a person has not or will not look at, s as a result meaningless, and thus irrelevant data.

Referring to the extensive work by (Floridi, 2011), the Stanford Encyclopaedia of Philosophy offers a very broad classification of information (see Figure 14). In his work, Floridi introduces a theory of strongly semantic information, arguing that *"true semantic content is the most common sense in which information seems to be understood"* (Floridi, 2004).

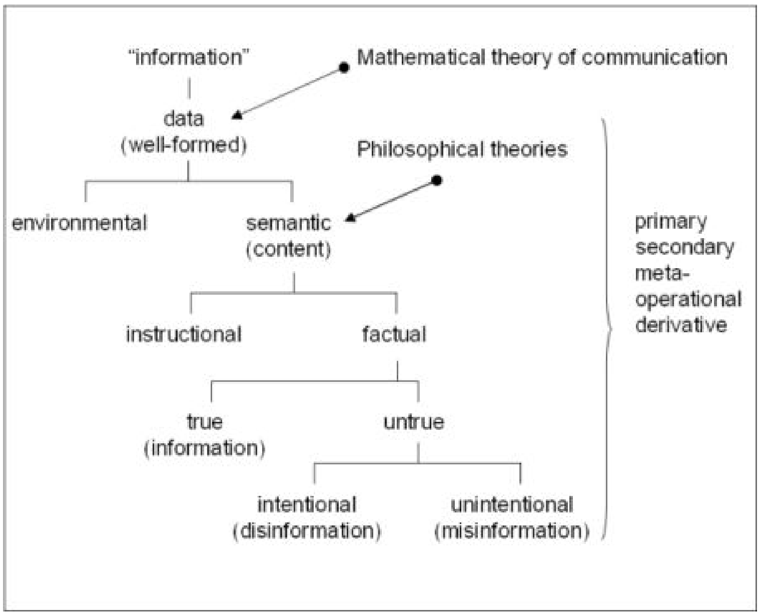


Figure 14: A Classification of Information (from Stanford Encyclopedia of Philosophy) (Floridi, 2011)

2.1.1.3. A Pragmatic View of Information

In their elaborations on “Working Knowledge”, Davenport & Prusak (Davenport & Prusak, 2000) pragmatically describe *“information as data that makes a difference”*. While staying close to Shannon in describing information as *“a message [that] has a sender and a receiver”* (Davenport & Prusak, 2000), the authors argue that information is meant to make an impact on its recipient by shaping insight or more generally making an impact in some way. Davenport & Prusak further conclude that it is *“the receiver, not the sender, [who] decides whether the message he gets is really information”*. To illustrate this point, (Davenport & Prusak, 2000) provide a scenario most of us encounter on a regular basis: *“a memo full of unconnected ramblings [which] may be considered ‘information’ by the writer but judged to be noise by the recipient”*.

Furthermore, Davenport & Prusak’s scenario is somewhat simplified in that it ignores the context in which information exchanges occur. Assuming the writer or sender does not intend to send irrelevant information, the recipient may still perceive said information as noise, simply due to bad timing. Not all information is relevant at any point in time - rather most information is only relevant at very specific points in time. If the recipient is focused on a specific task - say preparing breakfast for a crying baby - even the most useful information that does not fit the current task context might be regarded as noise. Understanding context is absolutely vital for the design of effective information systems and when discussing information overload in the digital realm. It will be discussed separately in the relevant sections of this thesis on focus + context interfaces (see section 2.3.6) and contextual search (see section 2.4.1.1).

2.1.2. Information Overload

Having too much information can be painful and troublesome (Mooers, 1959), harm our concentration (Wilson, 2005), and have a negative impact on our decision-making abilities (Schwartz, 2005). One could argue that until recently, bad decisions were made due to a lack of information. But today, bad decisions are being made due to information overload.

> *“What is a knowledge worker to do in a world where the Sunday edition of the New York Times has more information than the amount of information an average person alive 400 years ago might have come across in his lifetime?”* (Spira & Goldes, 2007)

Spira & Goldes describe information overload as *"a byproduct of the lack of maturity of the information age"* (Spira & Goldes, 2007) and argue that solving the information overload problem may be one of the most significant challenges for the 21st century.

In the context of knowledge work, Spira & Goldes link information overload primarily to the explosion of communication - namely email - which causes frequent interruptions leading to lower productivity and hampered innovation. According to a survey conducted in 2005, Spira & Goldes estimated said cost of interruptions for *"companies in the United States alone"* to be $588 billion in lost man-hours (Spira & Goldes, 2007). Following new research, they since revised their estimates to $900 billion in 2008 (Spira, 2008) and to $997 billion in 2011 (Spira, 2011).

While such an approach to sizing the problem of information overload may ignore the undeniable benefits that increased communication can lead to, the key question is not whether more communication or less communication is better, it is how to maximise 'good' and minimise 'bad' communication. Or put differently, it is about increasing the signal and reducing the noise.

Information overload is also not simply a problem related to work efficiency - it impacts society on every level - be it personal financial decision making (Fear, 2008), or killing innocent civilians in Afghanistan (Claburn, 2009; Shanker & Richtel, 2011).

Eppler and Mengis' analysis of *"the concept of information overload"* (Eppler & Mengis, 2004) illustrates the pervasiveness of the problem, as they consolidate *"literature from the domains of organisation science, marketing, accounting, and management information systems"* into a conceptual framework of information overload (see Figure 15).

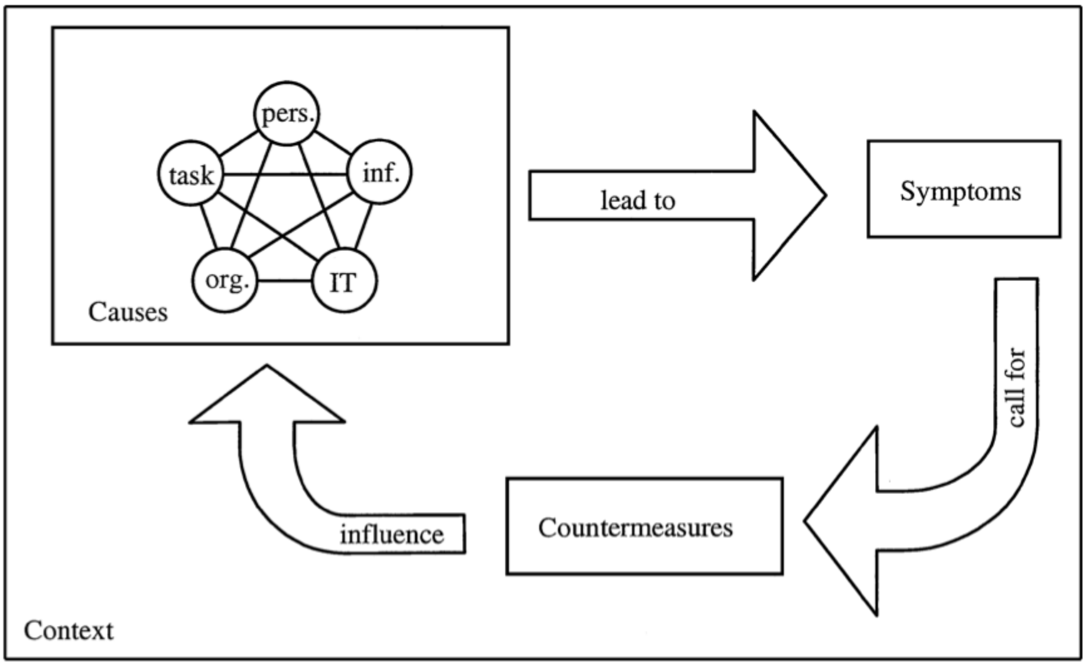


Figure 15: Conceptual Framework to structure research on information overload according to (Eppler & Mengis, 2004)

Eppler and Mengis identify five constructs that cause information overload (Eppler & Mengis, 2004), arguing that *"information overload emerges not because of one of these factors but because of a mix of all five causes"*:

1. the information itself (its quantity, frequency, intensity, and quality)
2. the person receiving, processing, or communicating information
3. the tasks or processes that need to be completed (by a person, team or organisation)
4. the organisational design (i.e., the formal and informal work structures)
5. the information technology that is used (and how it is used)

In addition to the five constructs suggested by Eppler and Mengis, an effective discussion of information overload should consider the three perspectives on information as discussed in the previous section - a mathematical perspective, a philosophical perspective and a pragmatic perspective. These perspectives on information provide three key insights:

1. Whenever information is communicated, the potential of noise needs to be considered
2. Information becomes meaningful upon human interpretation
3. The context in which information is being perceived is vital to its perceived usefulness

Accepting this premise, **reducing information overload requires minimising noise** (by limiting choice), thus improving the effectiveness of human interpretation. Moreover, effectively minimising noise by reducing choice requires the availability of information in the right (task) context.

As a result of the above three perspectives on information, ***this thesis stipulates a just-in-time information paradigm as mechanism to cope with information overload.*** This just-in-time information paradigm is inspired by the just-in-time (JIT) production strategy as popularised by Toyota as key aspect of lean manufacturing. Just as JIT aims to minimise or eliminate the use of (manufacturing) inventory for production, JIT in the context of interacting with information aims to minimise or eliminate the information noise, which has a negative impact on human ability to make sense of relevant information.

The **just-in-time information paradigm also relates to the concept of just-in-time (JIT) mental models** (Klein, Phillips, Raill, *et al.*, 2007; Sieck, Klein, Peluso, *et al.*, 2007). Klein et al. (Klein, Phillips, Raill, *et al.*, 2007) suggest that a clear distinction between JIT mental models and comprehensive mental models exists, pointing to an expert's domain knowledge as an example of the latter. In contrast, a JIT mental model is representative of people who may attempt to tackle a complex information problem in a largely unknown domain. The utilisation of the JIT mental model thus describes succinctly the type of information experiences that the JIT information paradigm has been proposed to address.

2.1.3. Summary

This section reviewed essential facets of the information age, by defining different perspectives of information, discussing relevant aspects of human-information interaction and analysing the problem of information overload. It provides some framing for how our understanding of information evolved from a purely mathematical view (Shannon, 1948) to more philosophical (Floridi, 2011) and more pragmatic views of information (Davenport & Prusak, 2000). This section concludes with a detailed analysis of the information overload problem in general, and how it affects knowledge workers in organisations in particular. By illustrating the range of perspectives used to define information as well as the range of problems encountered by people when faced with information overload, this section is meant to make a case for the value of investigating alternative ways to improve the user experience of human-information interactions.

2.2. The Human Perspective

As early as 1988, Marchionini & Shneiderman suggested that effective design of hypertext systems - the precursor to the Web - requires research into information retrieval, interface design, and cognitive science (Marchionini & Shneiderman, 1988). Over 20 years ago, Marchionini & Shneiderman argued that *"the linchpin of an information-seeking theory is the human user"*, and that a better understanding of how cognitive processes guide information seeking in general is important to better understand information seeking behaviour in electronic environments (Marchionini & Shneiderman, 1988).

Given the knowledge we have today on how web-based systems evolved since that time, and how these systems have changed our attitudes to and relationship with information, one could read the following more as a warning, rather than a mere observation: Marchionini & Shneiderman argued that *"any system that supports information-seeking must structure knowledge to make it accessible",* because using a system and interacting with the information it provides in the manner that the system supports will impact the way we think.

Marchionini & Shneiderman also highlight the aspect of learnability of a system suggesting that "*a system that is easy to learn may not be easy to apply in full"* because the mental model we develop of a system is shaped by our initial experiences with it (Marchionini & Shneiderman, 1988). The flexibility/complexity trade-off, Marchionini & Shneiderman elude to when discussing "*the tension between the learnability and applicability of a system"* has further implications. Any use of a system applied to tackle future (information) problems will be sub-optimal, if the initial impression (of the system) has shaped *"an incomplete and simple conceptual model"*.

Marchionini & Shneiderman's argumentation results in two key insights:

1. The way we access information impacts the way we think.
2. Using an (information) system in a simplistic way may impact our ability to solve future (more complex) problems.

It seems thus vital to thoroughly understand the process of human information processing, and to rigorously follow its implications when designing information-centric systems, as has been attempted in this research.

2.2.1. Stages of Human Information Processing

In his model of human information processing, Wickens & Hollands effectively capture the interplay of human perception and human cognition processes as an interdependence of bottom-up and top-down processing (see Figure 16) (Wickens & Hollands, 2000). The key stages in Wickens' model are *sensory processing*, *perception*, *memory* and *cognition*. Wickens defines these stages as follows (Wickens & Hollands, 2000):

Awareness (or sensory processing) is defined as the stage, in which *"information and events in the environment [...] gain access to the brain"*.

The **perception** stage *"is driven both by sensory input (which we call bottom-up processing) and by inputs from long-term memory about what events are expected (which we call top-down processing)"*.

The **cognition** stage is made up of working memory, where *"cognitive operations [such] as rehearsal, reasoning, or image transformation are carried out"* and long-term memory, a *"less vulnerable, and hence more permanent"* part of cognitive processings.

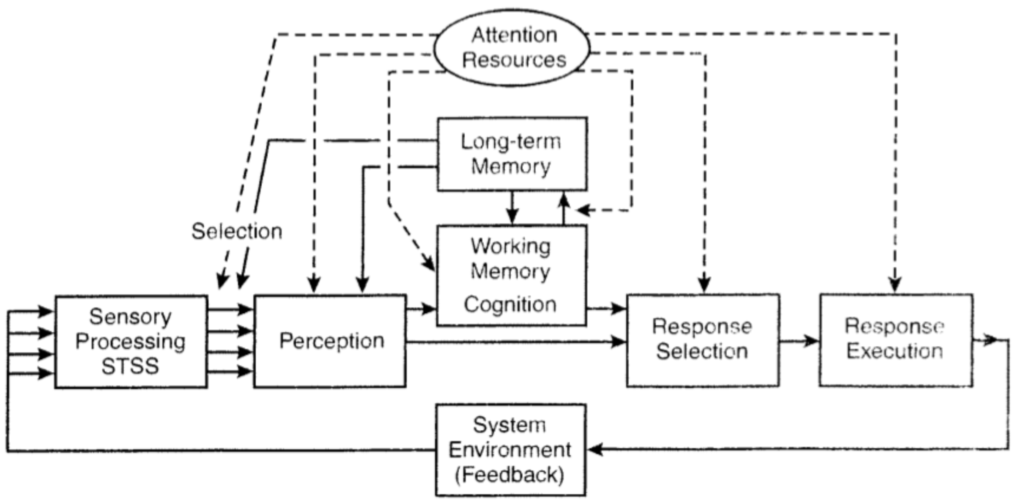


Figure 16: A model of human information processing stages

2.2.2. Designing for Attention

Wood et al. (Wood, Cox & Cheng, 2006) define (visual) attention as *"a process of selection and selective processing"* due to the brain's limited bandwidth for processing information. The authors refer to (Pylyshyn, 2003) to point out that our "*visual attention system seems […] to be prone to inattentional blindness".* Wood et al. discuss a number of metaphors used to describe visual attention, and conclude that no theoretical consensus exists due to strong differences in some and similarity in other metaphors.

In keeping with the interpretation of attention for this thesis, Wood et al. settle on defining visual attention as a mechanism similar to "*a camera's zoom lens".* Just as a zoom lens allows zooming out to get the bigger picture or zooming in to attend to some detail, our visual attention system can utilise its *"fixed amount of processing resources"* to distribute attention across a large part of our visual field with little intensity, or to bundle attention onto a much smaller area more intensely (Wood, Cox & Cheng, 2006). In other words, the level of cognitive processing, invariably linked to the complexity of an (information processing) task (e.g. Hollnagel, 1992, Ortoleva, 2008), impacts the way we attend to visual cues in the periphery of our locus of attention. The more (cognitively) focused we are on a task, the less focused we are on the information around us.

2.2.3. Perceptual Bandwidth

In "perceptual bandwidth", Reeves & Nass (Reeves & Nass, 2000) illustrate the range of definitions of perception that exist in Human-Computer Interaction (HCI) (see Figure 17). The authors stress the need for applying existing knowledge about human perception to how humans interact with each other and their environment through computers, particularly in the context of computer presentations that are "*rich in perceptual cues"* (Reeves & Nass, 2000).

P1, P3, and P4 revolve around perceptions by humans (see Figure 17), whereas P2 describes how a machine perceives the real world. Reeves & Nass include the computer's perception and recognition of objects, people and their emotions, personalities or gender in P2, but note that as a computer's perception is virtual, it is "*not regulated by human psychology"* and thus *"need not follow the rules of human perception at all"* (Reeves & Nass, 2000).

Given recent advances in mobile computing, and the ability to use smart phones as augmented reality displays (see Figure 18), one could argue that a firth definition should be added: "Human perception of the real world mediated by a computer". Such augmentation of reality (as in the example below) specifically aims to address the problem of perceptual bandwidth limitations by facilitating the processing of potentially relevant information.

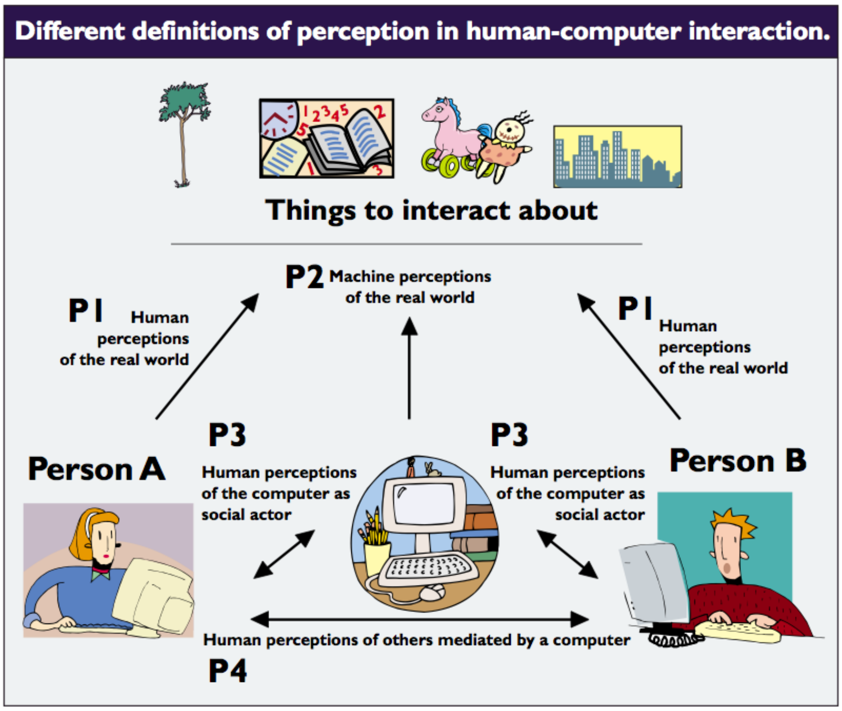


Figure 17: Different definitions of perception in human-computer interaction (Reeves & Nass, 2000)

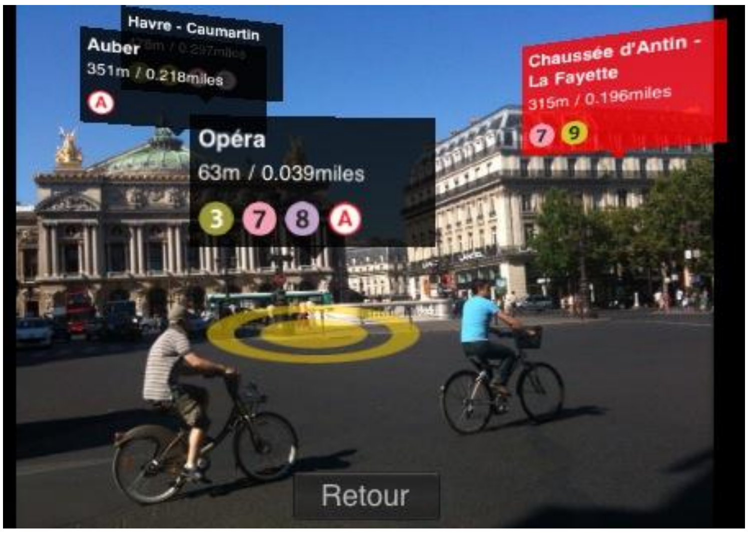


Figure 18: Example of an augmented reality display for smartphones

Novelty (of information) is another aspect relevant to perceptual bandwidth limitations. Novel information - whether novel people and places in real life, or a novel media presentation on a visual display - are *"perceptually more interesting"* and thus *"get attention"*. (GEIGER & Reeves, 1993) However, as Geiger and Reeves note, it is crucial to understand that when such novel information is unrelated to a primary activity or task, the perceptual impact (required attention) is stronger, than when novel information is related to a primary activity or task (GEIGER & Reeves, 1993). As a result, perceiving information unrelated to a task will make the processing of relevant information more challenging - something particularly undesirable with increasing task complexity. We encounter such unrelated information on an on-going basis when browsing the web - in the form of adverts, irrelevant content, or simply navigational structures that have to be made sense of.

2.2.4. Change Blindness

In their review of recent literature on change detection, Simons and Levin (Simons & Levin, 1997) discuss *change blindness* in the context of cognitive psychology. They point out people's inability to recognise changes within a scene if no localisable change has occurred. In the reviewed experiments, this 'localisability' is being undermined mainly through global transitions in the scene or distractions of attention which places the change in the scene to be between saccades. Simons and Levin (Simons & Levin, 1997) argue that people do not notice "*changes to scenes when they do not produce motion on our retina that attracts attention"*. The authors further note that *"change blindness supports the phenomenal experience of continuity"* as the amount of information being preserved between views is not large enough to notice less significant changes (Simons & Levin, 1997).

In essence, what Simons and Levin (Simons & Levin, 1997) are describing is that (human) vision has evolved over time to work as efficiently and as effectively as possible in the continuous world that surrounds us. Visual perception has been tailored (through evolution) to the physical world and its natural constraints, which seems to be increasingly orthogonal to the digital world we are creating.

Reeves and Nass' (Reeves & Nass, 2000) assessment of perceptual bandwidth can be linked to Simons and Levin's (Simons & Levin, 1997) discussion of change blindness. As Reeves and Nass argue that "*visual changes cause disruption",* visual design for change should ensure that connections are made between related material, thus facilitating *"the impression of stability rather than chaos"* (Simons & Levin, 1997).

The direction followed in this research strives to minimize change blindness by designing interfaces that minimize irrelevant information and overall provide a UI that is as minimalist as possible.

2.2.5. Decision-making

> *"True genius resides in the capacity for evaluation of uncertain, hazardous, and conflicting information."* (Winston Churchill)

In his work on theories of bounded rationality Simon illustrates how the *"constraints of the information-processing capacities of the actor"* (Simon, 1972) impact rational decision making. While Simon's focus is on organisational decision-making, he argues that rationality of an individual actor or an organisation are not that different. His *"limits of rationality"* in decision making are *risk and uncertainty*, *incomplete information about alternatives*, and *complexity* (Simon, 1972). Those limits of rationality motivate *satisficing behaviour* in that problems to be solved and decisions to be made are dealt with until a satisfactory alternative, which matches a pre-set "*aspiration level criterion"* (Simon, 1956) has been found.

More recently, Simon argued that the *inner environment* of our minds influences our *"rational behavior in the real world"* as much as the *outer environment* in the world that surrounds us (Simon, 2000). In this late work, Simon singles out the increasing problem of uncertainty, suggesting that the *"processes for generating alternatives for choice"* have not been researched sufficiently. He further suggests that existing cognitive psychology research on scientific discovery provides a "*very useful starting point"* for future work (Simon, 2000).

Crosby et al. provide a less economics-driven, but more HCI-centric definition of decision making as a *"cognitive process of deciding upon a particular course of action based on available data"* (Crosby, Iding & Chin, 2003). Their work being focused on augmented cognition, Crosby et al. (Crosby, Iding & Chin, 2003) argue that both *"task characteristics"* (bottom-up information), as well as *"the task performer's cognitive processes"* (top-down information) impact the perceived complexity of a task.

2.2.6. The Cost of Task Switching

In an environment, which promotes a mentality of always-on, instant access to information (as externalised knowledge), people are constantly reminded of - and enticed to tackle - the multitude of tasks and goals on their mental to-do list. On the Web, people increasingly use a variety of web sites in parallel, often leading to 'tab overload' in the browser. And the ability to easily switch between tasks is subjectively perceived as a way to increase efficiency, and minimise dead time.

However, working on more than one task at a time is generally less effective than focusing on a single one. Switching tasks requires people to adjust mental control settings leading, to considerable switching cost (Rogers & Monsell, 1995; Rubinstein, Meyer & Evans, 2001).

Rogers & Monsell investigated the *"costs of a predictable switch between simple cognitive tasks"* (Rogers & Monsell, 1995), arguing that control processes required to organise cognitive processes are not well understood. Striving to progress understanding in this area, the authors studied the costs to performance involved to *"switch between two tasks afforded by the same input"* (Rogers & Monsell, 1995), utilising the process of *task-set reconfiguration* to analyse their findings. Rogers & Monsell found that *"switching predictably between two simple cognitive tasks"* causes substantially increased reaction times and error rates (Rogers & Monsell, 1995), that extended practice with the utilised characterisation tasks only mildly reduced switch costs.

Rubinstein et al. (Rubinstein, Meyer & Evans, 2001) found that switch cost between tasks increased for increased task complexity, but decreased with task cueing. They further found switching from unfamiliar tasks to familiar tasks to be less costly than switching from familiar to unfamiliar tasks. Rubinstein et al. suggest that task switching is typically *"mediated by a rule-activiation stage of executive control"* which disables rules related to prior tasks and enables rules related to current tasks (Rubinstein, Meyer & Evans, 2001). The authors further argue, that for very dominant tasks, rules may be permanently enabled in procedural long-term memory, because of extensive prior practice.

2.2.7. Choice Overload

In their research on *choice overload,* Iyengar & Kamenica found that exposing users to large sets of choices increases their preference for a "*simple, easy-to-understand option"* (Iyengar & Kamenica, 2007). The authors argue that research in the disciplines of psychology and economics suggests that users may benefit from *"a strictly smaller choice set"*. As such, the pure availability of too many choices (i.E. choice overload) is a task characteristic as discussed by Crosby et al. (Crosby, Iding & Chin, 2003), which contributes to increasing task complexity. In related work, (Fear, 2008) investigated the impact of choice overload on financial decision making. While Fear's findings are quite varied, many specific to the financial domain, one more generalizable conclusion is that people should be able to *"choose not to choose"*.

2.2.8. Sensemaking

Effective choice making can be linked directly to *"the cost structure of sensemaking"*, in which Russell, Stefik, Pirolli and Card (Russell, Stefik & Pirolli, 1993) introduced *sensemaking* to the HCI community. Russell et al. define sensemaking as the *"process of searching for a representation and encoding data in that representation to answer task-specific questions"*. The authors acknowledge that during sensemaking activities, different operations require different cognitive and external resources and that appropriate representations are crucial to deal with information tasks effectively. In this seminal work, Russell and colleagues further stress that when facing complex information tasks, a focus on the information retrieval process itself *"may help very little"*, as a better understanding of subtasks and task structures are also important.

Attfield & Blandford (Attfield & Blandford, 2009) integrate Russell et al.'s work (Russell, Stefik & Pirolli, 1993) with the data-frame theory by Klein et al. (Klein, Phillips, Raill, *et al.*, 2007) to improve the understanding of the cost structure of sensemaking and to support *"reasoning about system requirements"*. The authors propose a two-step approach which aims to improve understanding of domain-specific sensemaking activities and helps to identify *"the most costly elements of the sensemaking process"*. The authors' approach suggests to first analyse the most time consuming and effort inducing elements of the sensemaking process and to utilise a sensemaking process model to identify opportunities for improvements. In a second step, the approach suggests understanding the key frames that require cognitive effort in aligning the user's understanding and the system's representation of a domain and *"optimise the conceptual fit between user and system"* (Attfield & Blandford, 2009).

While the late 90s and early 00s were dominated by research into information retrieval and visualization, papers such as Teevan's *"The perfect search engine is not enough"* (Teevan, Alvarado, Ackerman, *et al.*, 2004) have contributed to a shift in how the HCI community is looking at information problems and information-centric processes:

> *"We set out to understand and develop new ways to visualise information. To a large extent we have succeeded. It's time to declare victory and move on to something else. The next big thing is sensemaking."* (Stu Card, as quoted in (Stefik, 2004)).

In (Russell, Jeffries & Irani, 2008), Russell et al. reflect on the importance of simple tools to support *"sensemaking for the rest of us"*. Russell et al. acknowledge that the early focus of the *"sensemaking tool space"* overtly focussed on *"extremely high-end visualizations and tools"* that may be useful to experts, but which for the majority of sensemaking activities *"don't seem to help, but [rather] hinder the process of sensemaking"*. Interestingly, in this 2008 paper - 15 years after (Russell, Stefik & Pirolli, 1993), Russell's definition of sensemaking has not changed much, describing it as *"the way people go about their process of collecting, organising and creating representations of complex information sets"*. Given this definition, and the focus on the aspect of creating representations of information, Russell argues that mere process of *"data collection, […] while foraging through resource materials"* does not represent sensemaking per se, but rather is precursor to it: *"the data was collected, examined and read, but showed little evidence of any significant manipulation or restructuring"*.

Pirolli's model of how intelligence analysts process data and ultimately make sense of it (see Figure 19) illustrates the point made by Russell in that it distinguishes between an information foraging loop and a sensemaking loop (Pirolli, 2007).

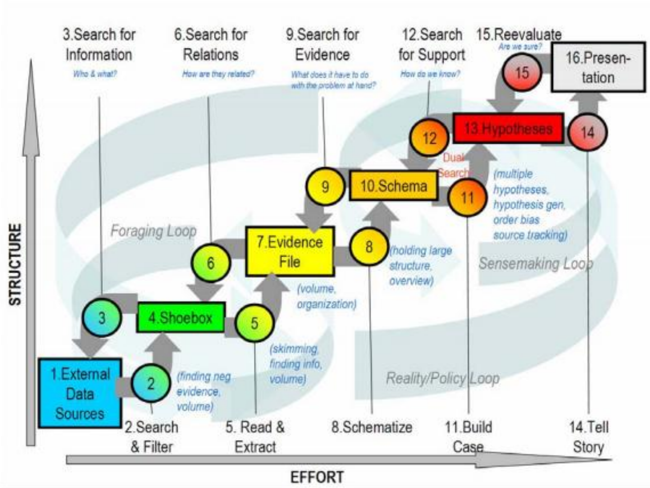


Figure 19: A notional model of intelligence analysis based on cognitive task analysis of experts (Pirolli, 2007)

Whereas the foraging loop involves searching for information, searching for relations and evidence, the sensemaking loop focuses on how schemas are formed based on that evidence that subsequently lead to the formulation of hypotheses, re-evaluation and finally presentation of findings of the conducted analysis work.

In their description of *"the nature of organized sensemaking"*, Weick et al. (Weick, Sutcliffe & Obstfeld, 2005) argue that *"sensemaking starts with noticing and bracketing"* - a person becoming aware of relevant information (in Weick's example a nurse noticing vital signs of a baby) and making a decision *"guided by mental models […] acquired during her work, training, and life experience"*.

2.2.8.1. Corporate Aspects of Sensemaking

In large organisations, the majority of information work consists of trivial, yet extremely diverse activities. As Russell (Russell, Jeffries & Irani, 2008) points out, complex and very specialized tools do not help or hinder the process of sensemaking in these everyday activities. What is needed are simple and unobtrusive tools, which seamlessly integrate into the general workflow of most users. The typical workflow interleaves multiple tasks (Bannon, Cypher, Greenspan, *et al.*, 1983), and sensemaking often requires moving from one application to another. Minimising the impact of these interruptions is important for user interface designers (McFarlane & Latorella, 2002). However, in some cases, one task is undertaken to make sense of another task, as when a knowledge worker searches among an archive of prior emails or a knowledge base in order to make sense of a current email.

> *"Sensemaking in our context is the process of connecting a series of individual bits of evidence to construct a larger, broader story or narrative. Sensemaking is about understanding how the individual events and entities referred to in the documents relate to each other and, when composed together, reveal a larger plot*." (Stasko, Goerg & Liu, 2008)

Many tools for sensemaking target the structuring and representation of familiar information. Documents and other organisational artefacts represent a form of group memory (Weldon, 2000). The difficulty for many knowledge workers is that they often do not know that relevant information exists within the organisation, and if they do suspect its existence, they often do not know where to find it. As a result, knowledge workers spend inordinate time searching and sensemaking. In (Millen & Fontaine, 2003), users reported spending approximately 15% of their work day on accessing or acquiring information. Other surveys report higher figures. According to one market research survey, in 2005 over 25% of a knowledge worker's time was involved in searching for relevant information, up from approximately 16% in 2001 (Outsell, 2005). An internal study conducted by Greenspan of CA Labs (the research collaborator involved in Main Study 2) in 2007 revealed that staff members in a technical services organisation of a large enterprise spent about 28% of their time searching for information in documents, emails and other sources. As cited in (Dubie, 2006), the Butler Group reports employees are both overwhelmed by the overload and lack of information; as much as 10% of staffing costs are lost *"because employees can't find the right information to do their jobs"*.

One of the major means of organisational communication and sensemaking is email. Emails are exchanged on various topics creating many overlapping threads, which the user interface tends to separate. Connections between one email thread and another often go unnoticed (see below for discussion on Xobni). But as noted above, group memory is manifested in many other sources of information, e.g. documents, wikis and blogs.

Sensemaking is not just a matter of understanding a single document or email, event or social interaction. It requires situating that experience in the context of what else is happening within a social group or larger context. However, knowledge workers may often be unaware of what they do not know. Considering Weick's (Weick, 1995) argument that organisational sensemaking is continuously challenged by the requirement for "interchangeability of people", not knowing what knowledge resides in the organisation daunts every new employee, and re-appears with every new project.

2.2.9. Summary

Section 2.2 discussed the human perspective of the information overload problem by reviewing relevant literature on how we become aware of information, actively perceive and process it, and try to make decisions based on that information. This section particularly focused on specific aspects of information processing that are relevant to information overload, such as perceptual bandwidth, change blindness, task switching costs and choice overload.

The limitations in human processing of information and their effects on our ability to cope with large amounts of information directly informed the formulation of the just-in-time information paradigm in this thesis and the decisions made in the design of the FMI and KnowDis prototypes. The discussion of sensemaking is used to illustrate how just-in-time information paradigm fits into the larger concept of cognitive task analysis by striving to optimize the information discovery and exploration aspect of the information foraging loop.

2.3. The Interaction Perspective

The following section explores the interaction perspective of the information overload problem by discussing relevant concepts of interacting with visual information displays. In doing so, this chapter explores how navigation strategies in general, and search in particular relate to the information overload problem.

This chapter further provides a detailed discussion of a wide range of focus + context style interfaces, and outlines how the approach of contextualisation of information can contribute to a potential solution of the information overload problem as outlined in the just-in-time information paradigm of this thesis.

A discussion of alternative interaction techniques such as (multi)touch, or gesture interfaces is out of scope of this thesis, but will be explored to some degree in the chapter on future work. A brief discussion on eye-gaze interaction techniques is provided in the context of Preliminary Study 2 (chapter 5) in which a gaze interaction prototype for just-in-time information interfaces has been developed and evaluated.

2.3.1. Human-Information Interaction (HII)

The shift towards an over-exposure to information has led to discussions about whether a *"separate field of human-information interaction"* (Jones, Pirolli, Card, *et al.*, 2006) is needed. At a CHI panel called *"It's About the Information Stupid"* (Jones, Pirolli, Card, *et al.*, 2006), one of the panelists, Stuart Card, argued that *"with global networks [...] information has been liberated to its own pure sphere"*, and that a *"new body of knowledge around human information interaction"* should be developed.

Coming from the same school of thinking at PARC as Card, in his book on *information foraging theory*, Pirolli (Pirolli, 2007) stresses the need to differentiate the field of human-information interaction (HII) from the field of human-computer interaction (HCI). Pirolli describes HII as a natural next step forward from traditional HCI, as information services become more pervasive, the interfaces to access information more transparent, and technologies to deliver information converge (Pirolli, 2007).

2.3.1.1. Information Foraging Theory (IFT)

According to Pirolli, his information foraging theory (IFT) (Pirolli, 2007) *"emerges from a serious consideration of Miller's notion of informavores"* (p13). IFT describes human information behaviour, modelled on *"hunting/food searching behavior"* as described by optimal foraging theory in biology by Stephens & Krebs (1986). Pirolli understands his work to be *"in spirit [...] to Allen Newell's (1990) program of Unified Theories of Cognition"* (p3) and argues that it is vital to understand people's ability to adapt when interacting with information in order to study real information foraging problems effectively. He describes the key adaptive force, which drives human interaction with information by referring to Simon, who argued that as information consumes "*the attention of its recipients [...] a wealth of information creates a poverty of attention"* (Simon, 1971). As a result, if large amounts of information are competing for attention, this attention needs to be allocated efficiently (as cited in (Pirolli, 2007), p40-41). This suggests that the design problem at the centre of "an information-rich world" is how to improve upon persons' ability to *"find and attend to information that is truly of value to them"* (Pirolli, 2007). This specific design problem elicited by Pirolli is at the heart of this research. Pirolli argues that the *"analysis of people interacting with information involves interrelated layers of explanation"* – namely: rational analysis, knowledge-level analysis, cognitive-level analysis, and biological-level analysis (see Table 2: Levels of Explanation (from Information Foraging Theory, p.18 in Pirolli, 2007).

Pirolli utilises this discussion of different levels of explanation to a) illustrate that *"people interacting with information involves interrelated layers [of analysis]"*, and to b) put information foraging theory - which applies rational analysis to the process of interacting with information - in a wider context. In unison with optimal foraging theory, IFT utilises elements such as *states, resources, state dynamics, constraints, affordances, feasible strategies,* and *optimisation criteria* to describe human information interactions and uses probabilistic models and Bayesian approaches to model human information behaviour in a rational manner.

While a philosophical discussion of whether humans are rational or irrational beings is out of the scope of this thesis, it appears at least necessary to mention the debate, as it questions the feasibility of any theory aiming to rationally describe human behaviour.

Table 2: Levels of Explanation (from Information Foraging Theory, p.18 in Pirolli, 2007)

Level	*Question*	*Stance*	*Analysis Elements*	*Examples*
Rational	What environmental problem is solved? Why is this solution a good one?	Design	• States, resources, state dynamics • Constraints, affordances • Feasible strategies • Optimization criteria	• Optimal foraging theory • Information Foraging Theory
Knowledge	What does the system know?	Intentional	• Environment • Goals, preferences • Knowledge • Perception, action	• Knowledge-level analysis
Cognitive	How does the system do it?	Information processing	• Cognitive states • Cognitive processes	• ACT-R • Soar
Biological	How does the system physically do it?	Biophysical	• Neural processes	• Neural models

2.3.2. Navigating the Web

In the early days of the Internet, web catalogues were a popular tool for the discovery of new web sites and the information they contain. With increasing scale, diversity and dynamics of information on the Web, the more or less static catalogues and the limited human resources behind them where replaced by more powerful gatekeepers - search engines. Over the last decade, search-based interaction has become the de facto standard for information retrieval on the web. Most people do not perceive any structure to the Web beyond what catches their attention on individual web sites. In the collective mind, search is the enabler and catalyst of information experiences. While information architects, user experience designers and graphic artists strive to optimise information experiences on individual sites, little search boxes have crept into their territory offering site searches and further aggravate the mantra of text-based search as one means to all ends.

How we navigate the web, is commonly categorised into *general purpose browsing* and *directed search* behaviour (Danielson, 2002) (Teevan, Alvarado, Ackerman, *et al.*, 2004). Links are the enabling concept for both behavioural categories. General purpose browsing describes the more exploratory behaviour of navigating page by page, where immediate *interestingness* - more or less subtle cues embedded in the link texts themselves or implied by the surrounding context of a link - is a strong driver of attention and action. But even when engaging in directed search, a goal-driven activity, where a concrete information need is to be satisfied, links play a vital part in the ranking of search results through their use in PageRank (Brin & Page, 1998) and related algorithms.

The popularity of the back button (Greenberg & Cockburn, 1999) and its widespread usage offers some indication of a conceptual problem in web navigation. The over-reliance on search engines not just for keyword-search but also as quasi-navigational aid, where search is used instead of bookmarking or remembering URLs, provide further evidence that web navigation as it has been intended may be - at a macro level - a flawed concept.

Current navigation and interaction behaviour with a web that is largely static in its visual structure of information and its perceived interconnectivity follows the cognitive reasoning that people develop mental maps of their experience of the web. Chun (Chun, 2000) points out that *"spatial context learning is ecologically significant"* as our perception of the physical environment that surrounds us is largely stable over time. As such, objects in the environment function as *landmarks* which provide "*useful navigation and orienting cues"* (Chun, 2000).

What Chun (Chun, 2000) refers to as *object cueing* can be related to the scenario of navigating web interfaces, where the human brain attempts to apply the same spatial cues. This is one of the reasons why good design practice recommends common locations for navigation, content, search boxes and other frequently used web elements.

Furthermore, (Chun, 2000) points out that the identities of objects are important, that background knowledge (schemas) allows an object to cue the presence of another object. He refers to this acquisition through experience as *co-variation* knowledge. This means that besides simple spatial cues, the semantics of a context are being interpreted by our brain. Translating that into a tangible scenario: *I know where the dishes are in my kitchen, but I do not necessarily know where they are in yours. However I might be able to limit the possible alternatives quickly once I see your kitchen.*

This example relates to web sites in the same way, but with one very significant problem - the Web is incomprehensible large, and users can hardly grasp a small subset of it. In order to manage all the input users are exposed to on the Web, their brains encode much of a *scene* as noise – as irrelevant information. Simons and Levin relate this phenomenon to a problem called *change blindness* - encoding *"the gist of a scene without explicitly coding the details"* (Simons & Levin, 1997) (see section 2.2.4 on change blindness).

2.3.2.1. The Connection between Navigation and Learning Processes

Using hypertext environments like the Internet for the development of learning applications requires an understanding of the importance of navigation, which is more than just the way of orientation and interaction (Schulmeister, 2007). It is especially an active form of learning, where the way people navigate from one information entity to the next directly influences the way they process this information and possibly acquire new knowledge - or as Fuster put it: *"learning and the acquisition of memory are based on the synaptic linkage of elementary cortical representations"* (Fuster, 1998).

Each individual utilises a unique network of those cortical representations - or personal knowledge - which he or she attempts to augment through the interaction with new information (the learning process). Each individual mental network consists of countless nodes of information entities, organised in topological, non-linear structures (Kuhlen, 1991). When interacting with hypertext environments like the web, the mental network of knowledge of an individual is confronted with the virtual network of information of the encountered environment.

Whereas novel educational research focuses on constructivism (Bruner, 1966) and related methods like *experimental learning* (Kolb, 1984) or *discovery learning* (Bruner, 1966), under the assumption that every individual has a different way to acquire knowledge, on the web, information is mostly linked in a static manner. While hyperlinks can provide some amount of flexibility if used in a sensible way, those potential advantages are frequently undermined by overloaded or badly structured web pages. In some ways, this leads to a virtual network of information - the web - which forces people to adapt their way of thinking to its own static information structures.

2.3.2.2. Navigation and Knowledge Building

With increasing complexity of available information (both in scale and diversity), interference between different bits of information increases. The promotion of links as a key means to value relevance of information has strongly contributed to these interferences. The static nature of links on the Web creates an environment that cannot adjust to the user's individual information needs (based on unique background knowledge and a specific information task context). Considering the way people process information, navigation is more than just orientation and interaction (Schulmeister, 2007), but much more an active form of learning. With regards to web-based information tasks, the way people navigate from one bit of information to another directly influences the way they process this information and potentially acquire new knowledge. Although most of today's web sites are dynamically generated, if the information contained does not match a user's actual information needs, at best it is being perceived as static and irrelevant bits of information. At worst, they will distract the user from her original information goal.

Current web-based user interfaces typically follow rigid and static visual designs, using grid and table-like layouts, where rows and columns blur the border between information, navigation and other visual noise. Such traditional approaches to web-design, which fill web-pages will large amounts of loose connected bits of information believe that these print-like UIs do not work for personalised information delivery. RSS feeds are an interesting example of new technology designed to cope with the earlier mentioned problems of information overload. RSS is particularly used in the blogging context to provide personalised information delivery. But, when looking at how phenomena like change blindness affect the perception of overloaded web pages it becomes clear that traditional interfaces (e.g. email-like news readers) represent a burden to the user (LIN, 2006).

2.3.2.3. Designing for Limitations of Human Attention and Human Cognition

Halverson et al. (Halverson, 2003) investigate effects of hierarchical structuring of information on eye movements and search performance and finding that structuring larger selections of choices into hierarchies using labels helps people to find a specific element. Their findings primarily support the idea of improved *findability*, by using visual structuring as an external memory aid. Halverson et al.'s findings directly relate to Norman's discussion of short-term memory limitations (Norman, 1992), as he distinguishes between *knowledge in the head* (internal information) and *knowledge in the worl*d (external information). Norman argues that good design practice should "*provide external aids to memory"* in order to compensate for the natural limitations of human memory and attention.

Work by Norman (Norman, 1992) and Oulasvirta & Saarilouma (Oulasvirta & Saariluoma, 2004) on short-term memory limitations and the (cognitive) mechanisms relevant for storing information in long-term memory provide vital insights into the design of contextual interfaces. Their findings imply that contextual information provided needs to be relevant to the current task, as providing information that is unrelated to the current task in a contextual manner would facilitate task switching and thus have negative effects on task performance or even task completion. Oulasvirta & Saariluoma (Oulasvirta & Saariluoma, 2004) suggest tasks to be *"organised into small and coherent chunks or episodes"* as a coping mechanism for interference caused by interrupting messages. This approach minimises the chance of *"an interruption occurring during elaboration"* - before the completion of a 'task episode'. The problems with this approach are a) the required overhead to break down complex tasks, and b) the challenge in defining an appropriate level of complexity for task episodes.

Another interesting perspective is provided by Wood et al., who propose the design of systems which are able to monitor users' attention levels and dynamically modify the interface such that user attention is maintained (Wood, Cox & Cheng, 2006). Wood et al. describe the suggested approach as a major challenge for the research disciplines of *HCI* and *cognitive engineering.* The authors approach aims at creating a *"new class of interfaces"* which are fundamentally different from existing systems and thus require more than just continuous iterations and incremental improvements of existing technologies (Wood, Cox & Cheng, 2006). Wood et al.'s approach is not unlike the one taking in this research – questioning the existing means to interact with information through traditional browsing or keyword-based search and instead investigating an alternative approach as described by the just-in-time information paradigm.

2.3.3. Search Interaction

Marchionini provides a detailed account of the range of activities people engage in when using search (see Figure 20) (Marchionini, 2006). While organising search activities in three distinct types - lookup, learn, and investigate - the crucial distinction is the one between lookup search (often also referred to as factual or navigational search), and exploratory search.

A lot of existing research on search primarily studies lookup-type search. This search type naturally lends itself to the measurement of search performance through the use of simple search tasks (Teevan, Cutrell, Fisher, *et al.*, 2009; Joshua Hailpern, 2011; Dumais & Cutrell, 2001; Patrick Baudisch, 2004) for two reasons:

1. Tasks are more homogenous, thus participants' interactions (anticipated or actual) are more comparable, and performance benchmarks can be more readily calculated.
2. Existing search systems work well for (more simple) lookup search tasks, than for more complex exploratory search tasks.

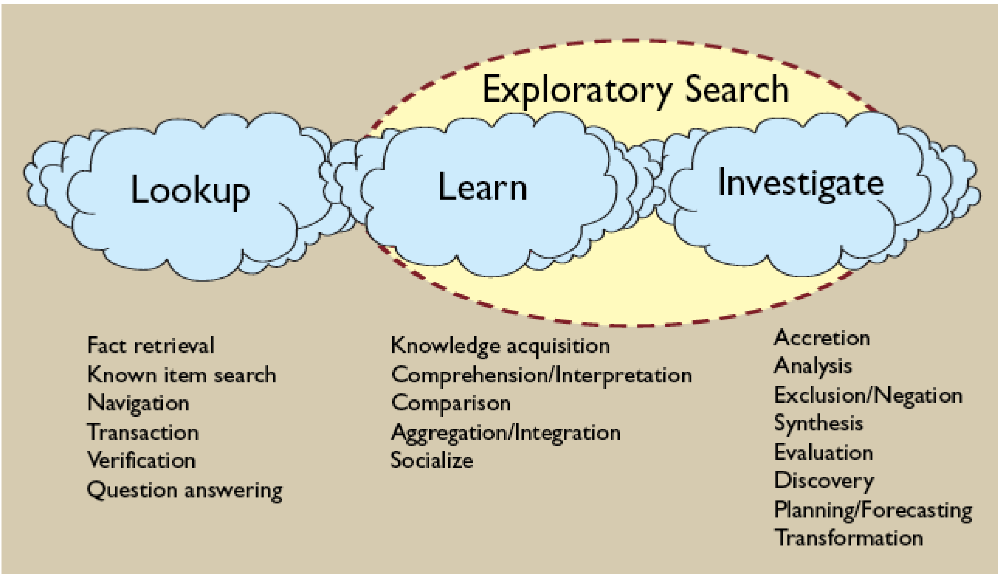


Figure 20: Categorisation of search activities by (Marchionini, 2006)

Given that early search systems were designed for simple lookup search tasks, and many researcher subsequently studied this type of interaction, one could argue that a self-enforcing cycle materialised in the form of continuing efforts to improve search performance (speed, error rates, etc.), and a persistent focus on simple search tasks in the design of search systems.

Heavily influenced by Schneiderman's eight golden Rules of Interface Design (Schneiderman & Plaisant, 2009), Hearst outlines the following design guidelines with a particular focus on search interfaces (Hearst, 2009):

1. Offer informative feedback
2. Support user control
3. Reduce short-term memory load
4. Provide shortcuts for skilled users
5. Reduce errors; offer simple error handling
6. Strive for consistency
7. Permit easy reversal of actions
8. Design for closure.

A few of those guidelines stand out in the context of this research - give users the ability to be in control of the (search) process (no. 2), reduce cognitive load (no. 3), and strive for a consistent experience (no. 6) when searching for information. The problem, and one aspect Hearst's book falls short on are interfaces for exploratory search and complex information problems. However, this may well just be a reflection of the state of research in this field, and Hearst argues accordingly: *"Most pressing is the need for better interfaces to support complex information seeking tasks"* (Hearst, 2009).

Hearst does acknowledge the need for exploratory search interfaces and exploratory browsing, discussing the process of search in the context of larger sensemaking activities. But sensemaking tools and sensemaking interfaces focus more on the process of representation building (Russell, Stefik & Pirolli, 1993; Russell, Jeffries & Irani, 2008), and less on the actual process of information discovery and gathering, which could be argued is at the core of exploratory search tasks.

More recently, research focusing on exploratory search has emerged (Qu & Furnas, 2008; White, Marchionini & Muresan, 2008; Hoeber & Yang, 2008; Marchionini, 2006; Ahn, Brusilovsky, He, *et al.*, 2008; Villa, Gildea & Jose, 2008; Indratmo & Gutwin,

2008) (Ruthven, Baillie, Azzopardi, *et al.*, 2008) and may accommodate the growing expectation of search systems to solve complex information problems as described by (Marchionini, 2006). The following section discusses relevant exploratory search systems in more detail.

2.3.4. Search Strategies

Directly related to the (technical) categories of search as described by (Marchionini, 2006) are the strategies people employ when engaging in search activities as described by (Teevan, Alvarado, Ackerman, *et al.*, 2004). Investigating users' search behaviour within email, personal files and on the Web, Teevan et al. (Teevan, Alvarado, Ackerman, *et al.*, 2004) studied the actual search strategies people exhibit when engaging in search activities. The authors found that users generally preferred orienteering strategies over keyword-based search (Teevan, Alvarado, Ackerman, *et al.*, 2004). Based on their findings, Teevan et al. propose distinguishing between 'teleporting' and 'orienteering' behaviour as distinct search strategies (see Table 3). The authors argue that the richness of contextual information provided through orienteering behaviour helps users make sense of where they are, where they want to go, and helps users to feel in control.

Arguably, orienteering as described by Teevan et al. can be likened to people's natural inclination for exploratory information behaviour. In contrast, teleporting or keyword-based search can be understood as the type of search behaviour primarily facilitated by keyword-based search systems designed for simple information tasks.

Table 3: Search strategies for the Web by (Teevan, Alvarado, Ackerman, et al., 2004)

Strategy	**Description**
orienteering	user tries to *"reach a particular information need through a series of small steps"*
teleporting	user tries to *"jump directly to their information target"*

Teevan et al. identify three main properties to reason users' preference for orienteering:

1. Decreased cognitive load
2. Maintained sense of location during task
3. Better understanding of results presented

The relevance of all three properties can be linked to the intrinsic complexity of the search task at hand. The more complex an information need, the more complex a search task, and the more preference a user should elicit with regards to orienteering-type search behaviour.

A common property of complex search tasks is that users may not be able to *"articulate exactly what they were looking for"* (Teevan, Alvarado, Ackerman, *et al.*, 2004), something catered for by orienteering behaviour, thus resulting in eased cognitive load. Complex search tasks are typically not satisfied by a simple factual snippet of information. Such tasks rather require users to explore potentially relevant information and slowly narrow down on the snippet(s) of information that may satisfy the information need.

White et al. (White, Marchionini & Muresan, 2008) describe the process of engagement in exploratory search as *opportunistic, iterative,* and *multi-tactical.* The authors argue that a user engaged in exploratory search typically provides a query that is not aimed to provide a perfect match, but rather just good enough to get the user close to relevant information. Once this has been achieved, the user then starts orienting within the environment to decide on which steps to take to get to that relevant information (White, Marchionini & Muresan, 2008).

2.3.5. Exploratory Search Systems

The following section discusses a range of exploratory search systems and concludes with a critical assessment of their relationship and relevance to the research presented in this thesis.

In *"Evaluating Exploratory Search Systems"*, White et al. describe exploratory search as an *"open-ended, persistent, and multi-faceted"* information need (White, Marchionini & Muresan, 2008). White et al. argue that exploratory search systems (ESSs) need to leverage *"new technological capabilities and interface paradigms"*, such as visualization systems, document clustering and browsing systems, and intelligent content summarisation systems, to support human-information interactions that are adequate for complex information needs. In order for an ESS to be effective, it needs to facilitate information explorations that aid decision making about where to go next, support comprehension of the information presented and result in higher information gains (White, Marchionini & Muresan, 2008).

Indratmo and Gutwin apply the use of interactive visualisations to the domain of blog archives to facilitate information exploration through a system called iBlogVis (see Figure 21). iBlogVis utilises tags and graphical visualisations to illustrate what a blog is about, and to surface a blog's *social interaction history*. Such functionality aims to facilitate giving users unfamiliar with a blog an inside view of a blog's *personality* and content as only a regular blog user would otherwise have. As such, iBlogVis is more targeted toward opportunistic information exploration, as opposed to supporting more targeted information exploration driven by complex information problems as investigated in this thesis. Whereas the work in this thesis is concerned with supporting a users' concrete information goal by making the most relevant information available to the user in a proactive manner, iBlogVis is meant to help a user understand what a given corpus of information is about, primarily through the visual exploration of metadata.

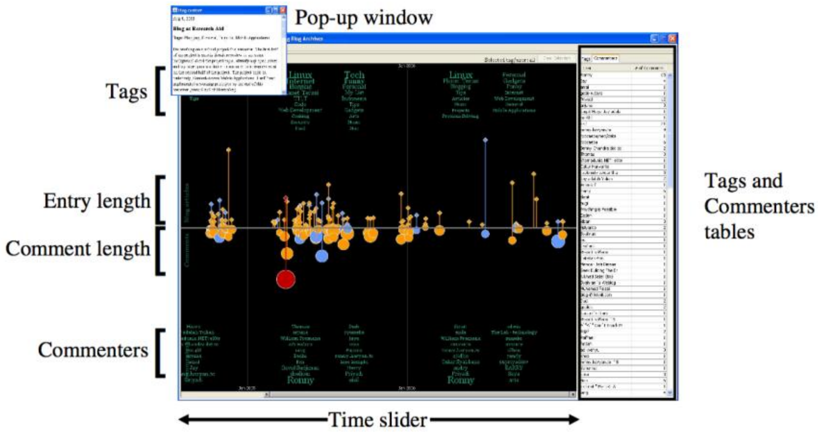


Figure 21: iBlogVis interface screenshot from (Indratmo & Gutwin, 2008)

Villa et al. (Villa, Gildea & Jose, 2008) propose a system called FacetBrowser to approach the problem of 'complex search tasks' (Figure 22). The authors acknowledge the need to explore different facets of a complex information problem concurrently, which FacetBrowser has been specifically designed for, allowing *"more than one search to be executed and viewed simultaneously"* (Villa, Gildea & Jose, 2008). The observation by Villa et al. that users need to explore different aspects or facets to solve a complex task is noteworthy, however their solution in form of the FacetBrowser interface is in effect a system that allows running three searches side-by-side, thus getting three sets of results on one screen. In their evaluation of the FacetBrowser interface, the authors compare it to a baseline interface, which emulates a more traditional search interface and requires the use of three tabs (instead of the three concurrently displayed searches in the FacetBrowser interface). While they found the FacetBrowser interface to be superior for complex tasks that required the use of multiple facets (or multiple searches), their findings simply imply that tabbed browsing is not ideal when dealing with complex multi-facet information problems, and a concurrent display of different tabs may be superior. This can be explained by the fact that switching between different views requires the user to remember results from a different, not currently visible view/tab. This increases cognitive load for the task and thus makes the combined view of the FacetBrowser interface more effective. The main difference between the FacetBrowser and the work in this thesis is conceptual. Where FacetBrowser aims to improve keyword-based search by letting the user run three keyword-based searches in one view, the just-in-time information paradigm aims to support the user proactively based on their existing (work) context without requiring the formulation of keywords. Furthermore, the concept of FacetBrowser seems more suitable to image or other forms of visual search, rather than search for textual information as is the focus in this research.

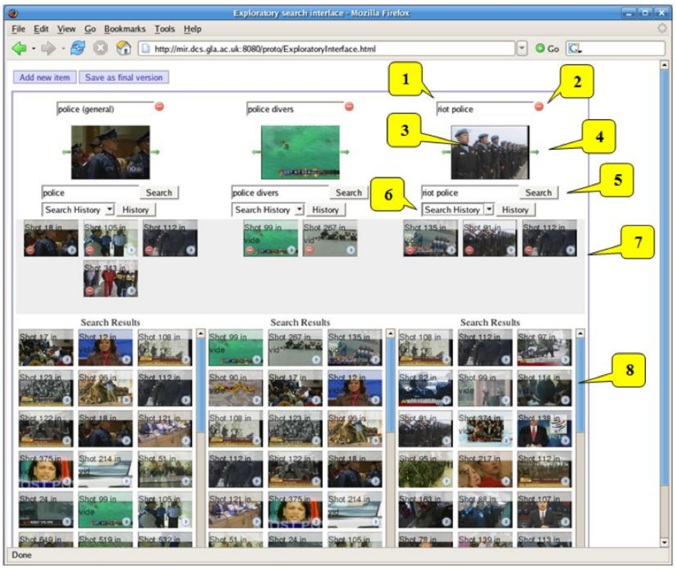


Figure 22: FacetBrowser interface screenshot from (Villa, Gildea & Jose, 2008)

Hoeber et al. (Hoeber & Yang, 2008) take a slightly different approach from Villa et al., identifying *"difficulties [in] crafting queries"* and *"the lack of support for constructing and refining queries"* as common problems among *"web searchers"*. The authors propose a system called WordBars (Figure 23), which provides a histogram of term frequencies as found in the first 100 search results to assist exploration tasks.

Exploration is supported by allowing users to refine the result display by adding or removing *“terms from the histogram … generating a new set of search results”*. Hoeber et al. (Hoeber & Yang, 2008) argue that one of the key problems with traditional search systems and the utilised search engine result pages (SERPs) is the presentation of results as *“a static ordered list”*. They suggest that a more *“interactive exploration, drawing upon the user’s understanding of their information need”* can surface more relevant documents which are somewhat hidden further down in the search results. Like FacetBrowser, WordBars is an augmentation of traditional search-based interaction. Its concept supports the manual re-formulation of a keyword-based query string in order to retrieve an alternate list of search results. In contrast, the work in this thesis focuses on providing relevant additional information proactively based on the content that the user is focusing on.

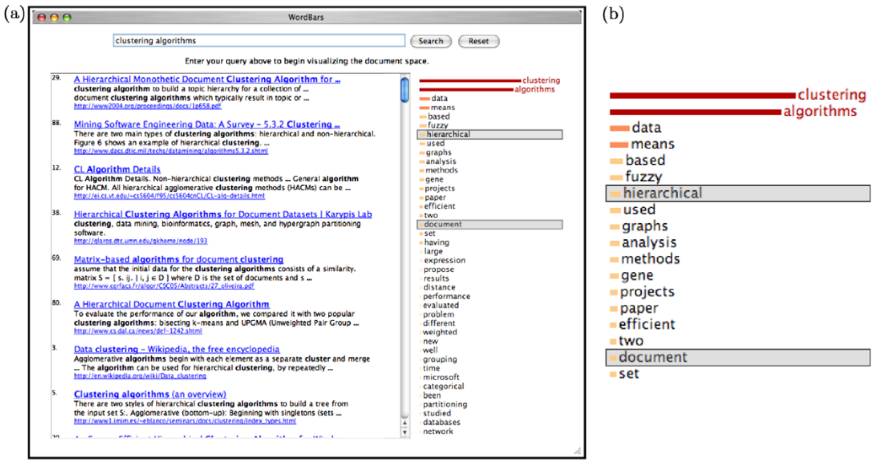


Figure 23: WordBars interface screenshot from (Hoeber & Yang, 2008)

Ahn et al. (Ahn, Brusilovsky, He, *et al.*, 2008) investigate *“personalization in task-based information exploration”* through TaskSieve, a *“web search system that utilizes a relevance feedback based profile, called a ‘task model’, for personalization”* (Figure 24). The authors argue that in contrast to most other personalised search systems, which model user interests in a more traditional manner, TaskSieve utilises a *“more focused task model”* to support exploratory search processes. The distinction made by Ahn et al. is between the more traditional long-term and stable user models that other systems generate over time, as compared to the more short-term task model as utilised by TaskSieve, which starts to model a task when a user starts searching with the tool.

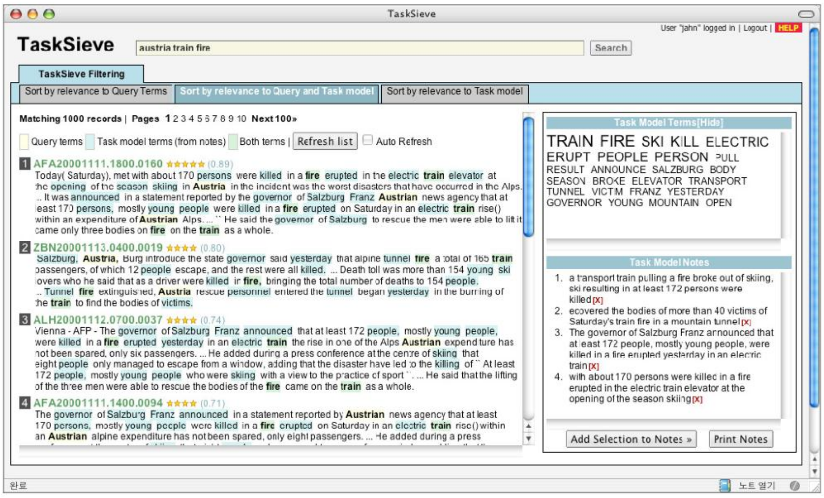


Figure 24: TaskSieve interface screenshot from (Ahn, Brusilovsky, He, *et al.*, 2008)

Out of the various tools compared in this section, the concept of TaskSieve seems to be the closest aligned with the goals of this research. However, where the approach taken with TaskSieve seems to focus on a rather advanced and complex tool aimed at an expert information analyst user, who is willing to manually curate a task model, the work in this thesis focuses on a more unobtrusive support of users information goals. In fact, the authors state that *"professional information analysts … appreciate more powerful and sophisticated information access tools [and] want to be in control of the system's work and highly value the transparency of the system mechanisms"*. (Ahn, Brusilovsky, He, *et al.*, 2008). The research on the just-in-time information paradigm focuses primarily on a 'typical' user who would be overwhelmed by a system like TaskSieve. Even the work reported in Main Study 2 (see chapter 7) investigating the support of knowledge workers information goals showcases that those professional knowledge workers – while likely highly skilled experts in their various fields – are by no means professional information analysts and would be highly unlikely to adopt a tool as complex as TaskSieve.

A detailed discussion of literature concerned with the evaluation of exploratory search systems can be found in section 3.2.9.

2.3.6. Focus+Context Interfaces

> *"Information Worlds were getting large, while our windows into those worlds were quite small"* (Furnas, 2006).

Focus + context interfaces aim to improve orientation within and exploration of large information spaces by embedding the visualisation of some focused information into additional contextual information. This approach preserves context while the user is focusing on a particular bit of information. This approach aims to benefits task performance and improve the general understanding of the informational structure underlying the information space.

Focus + context interfaces are a growing area of interdisciplinary research at the cross-section of research into HCI and information visualization: (Furnas, 1981; 1986; 2006), (Noik, 1994), (Lamping, Rao & Pirolli, 1995), (Brown, Marais, Najork, *et al.*, 1997), (Nation, Roberts & Card, 2000), (Card & Nation, 2002), (Gutwin, 2002), (Kosara, Miksch & Hauser, 2002), (Schrammel, Giller, Tscheligi, *et al.*, 2003), (Munzner, Guimbretière & Tasiran, 2003), (Patrick Baudisch, 2004), and (Laqua & Brna, 2005). Much of this research focuses on well-structured information, such as trees (Lamping, Rao & Pirolli, 1995; Nation, Roberts & Card, 2000; Card & Nation, 2002; Munzner, Guimbretière & Tasiran, 2003), or other more or less hierarchically organised information spaces. The following sections discuss some of the most relevant techniques.

2.3.6.1. Linear Context Visualizations

Mackinlay et al.'s 3-D *perspective wall* (Mackinlay, Robertson & Card, 1991) aims to *"support large scale cognition"* by applying a 3D metaphor to 2D representations of linearly structured information spaces and enable smoothly animated transitions between views. The perspective wall interface has been inspired by Resnikoff's observations of the human eye, which effectively integrates the eye's focus on a particular detail while maintaining a general view of the surrounding context. The perspective wall technique uses *"a physical metaphor of folding … to distort an arbitrary 2D layout into a 3D visualisation (the wall)"* (Mackinlay, Robertson & Card, 1991) yet retain the particular visual features of the original 2D layout, such as highlighted keywords, etc. The perspective wall has been applied to domains, such as document structures or individual text documents (see Figure 25).

Both domains contain information that is linear and continuous, allowing the modelling of context in a simple linear manner. While the application of a linear contextual model is appropriate in those domains the accessibility of the information represented as context is limited. Primarily the use of coloured markers is used to identify certain elements within the context, as the readability is either severely impacted or non-existent. As a result, this type of focus + context visualisation primarily provides increased awareness of information structures. For example, in the display of text documents using a perspective wall, the highlighting of keywords offers some insight into how frequently those same keywords appear within the rest of the document.

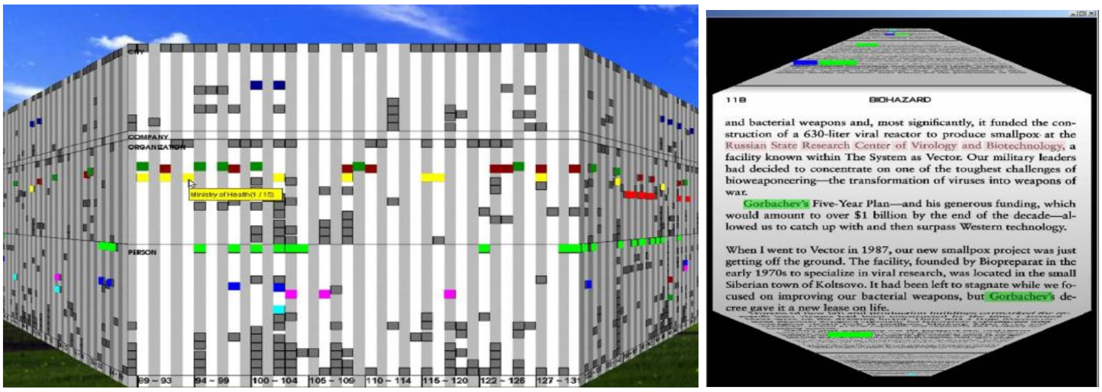


Figure 25: Focus + context visualisations from Xerox Parc

The focus + context interfaces applied to such structures are typically designed to surface more structural information, but not the actual content hidden within those structures. As a result, But most scenarios consist of more complex tasks, where the user wants to find what is behind the nodes. When using the example of DOI trees (Card & Nation, 2002), this sort of visualization works well if the task is to gain an overview of the structure, the network or how the entities relate to each other in a broad “macroscopic” view. But if the user actually wants to work with the content which is represented by “placeholders”, it is required to link from this visualization to an external source that contains this detailed information – the content. Now, when the idea of this type of visualization has been to visualise context and to understand the actual content in its context, then this is a bad solution as the content can only be accessed separated from the contextual visualization. Jumping between the detailed content and the entity-like visualization of context is cumbersome and mentally more challenging due to switching costs (Rogers & Monsell, 1995).

2.3.6.2. Fisheye Views

Furnas' work on "Generalized Fisheye views" 25 years ago was groundbreaking (Furnas, 1986), in that much of the later research on focus + context (F+C) visualisations has been based on or at least is related to his original concept. Furnas defined a degree of interest (DOI) function that calculates an a priori importance (API) for each point (or entity) in a structure and a distance (D) based on the point and its distance from the current focus of the user $DOI_{fisheye} = (x|.=y) = API(x) - D(x,y)$.

Furnas designed general fisheye views to provide a balance of information in focus and its surrounding context. Fisheye views can be used on a range of well-defined information structures, such as *lists, trees, acyclic directed graphs, general graphs* and *Euclidean spaces*.

Inspired by The New Yorker's 1976 cover of the 'View of the World from Ninth Avenue' by Saul Steinberg (Figure 26), his formalism has been inspired by a very basic two-dimensional "information space" – a map.

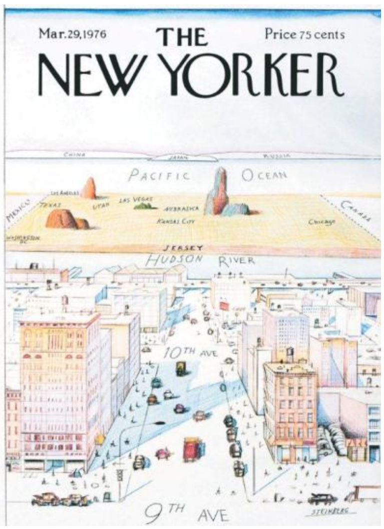


Figure 26: Cover of a 1976 issue of the New Yorker showing "View of the World from Ninth Avenue" by Saul Steinberg - it illustrates a fish-eye view of the map of the United States and represents any early focus + context style visualization - in this case to make a statement about what constitutes 'focus' and what constitutes 'context' from a Manhattan perspective.

Gutwin presents an **interactive fisheye view** to browse through web sites (Figure 27) (Gutwin, 2002). Moving the mouse cursor will shift the point of focus. This is a very dynamic and immediate solution to adapt the display of focused and contextual information. The contextual display of web pages linked to the current focus takes the form of scaled-down screenshot thumbnails. This solution results in visual "gestalt" information as the dominant information for contextual cues. Thus, for unfamiliar web pages contextual information is basically not existent. As Gutwin acknowledges, one of the main problems for users with this type of contextual display is to exactly point to a specific object with the cursor.

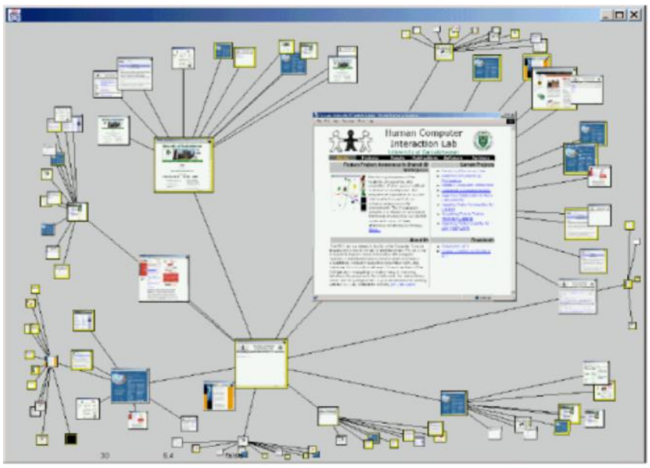


Figure 27: Gutwin's interactive fisheye view of web sites

In contrast to Gutwin's solution, the fisheye distortion in Baudisch et al.'s (Patrick Baudisch, 2004) **fishnet browser** is applied to an individual web page rather than an overall web site (Figure 28). However, the type of implementation used to enable a fish-eye style visualization of a web page still makes it harder to exactly point and select an page object by using a mouse. The same problem appears in fish-eye implementations, such as the **fisheye menus** by (Bederson, 2000). These examples show one core problem: The more a user would benefit from a contextual display by mapping larger information spaces to the display space, the less usable this visualisation gets. A user needs to somehow interact with objects within this information space but to accurately pointing to an object gets increasingly difficult.

Figure 28: Baudisch's fishnet browser

Generally, fish-eye view interfaces preserve *"the contextual relationship between a large number of objects"* (Demaine, 1996). But the various implementations address context on different levels. As pointed out by Furnas (Furnas, 2006), whose work on fish-eye views now spans three decades (Furnas, 1981; 1986), it is important to distinguish between what is presented and how it is presented when designing focus + context interfaces. While "users have tasks to do and need certain information to do them" it is crucial to understand what information is essential in the context those tasks (Furnas, 2006). Once such essential information is identified, the actual second step is to decide on how to best present that information. It could be argued that the interpretation of fish-eye views throughout the various implementations that have been created stayed very close to the original map-based concept. The primary goal seems to have always been to represent all context within a single view and then provide the fish-eye style focus to zoom in on a particular area for more detail. Such an interpretation seems to de-prioritise the 'what' against the 'how'. Of course, no focus + context interface will be without compromises, as physical restrictions such of display size, screen resolution or complexity of the information to be displayed all impact the design. However, for information tasks one might argue, whether displaying more context really results in more understanding and whether visualising the structure of all contextual information really benefits the user in most cases.

2.3.6.3. Degree of Interest Trees

At PARC (formerly Xerox Parc), Card and colleagues developed 'Degree of Interest (DOI) Trees' (Figure 29) as a means to represent and browse through very large hierarchies of information (Nation, Roberts & Card, 2000; Card & Nation, 2002). One of the limitations of DOI trees is its dependency on well-defined tree-like information structures. One could further argue that its attempt to display as much context as possible is an additional disadvantage. While the display of such rich contextual information can help users understand the relationships within such large information structures, displaying too many 'nodes' simultaneously can result in visual noise. The display of too many concurrently displayed nodes may result in information overload within the context itself. For example when faced with a particular information task, only a marginal subset of all nodes may be relevant. Thus, presenting less information in a more dynamic context that is relevant to the individual user could result in a more effective and more pleasant information experience.

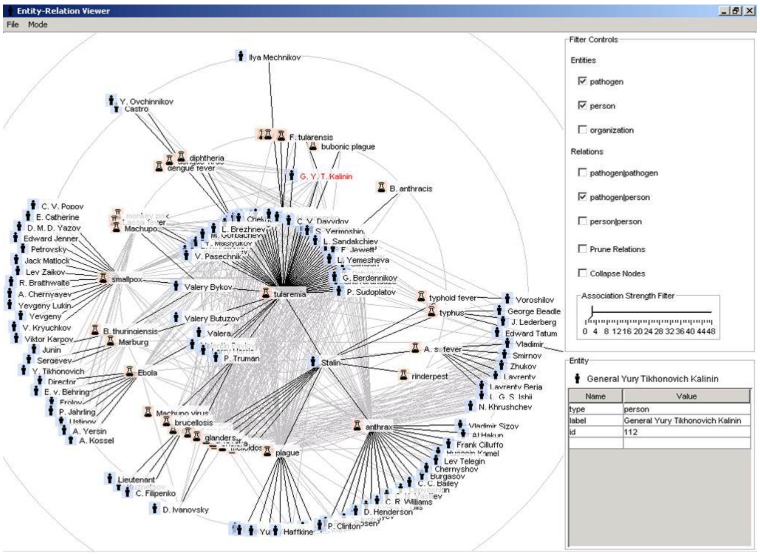


Figure 29: DOI tree interface

2.3.6.4. Treemap Interfaces

Newsmap (Weskamp & Albritton, 2004) is a visualization of news headlines utilising a treemap view to display and prioritise news stories based on Google News data (Figure 30). The newsmap interface provides contextual cues about news headlines through spatial organisation and colour contrast information. The size of the headline reflects importance, themes are colour-coded, and colour saturation degrades over time, meaning newer stories stand out much more than older ones. The newsmap web application offers a number of filter mechanisms to switch between or combine local news for various countries (US, UK, Germany, France, …), various topics or to search for particular keywords.

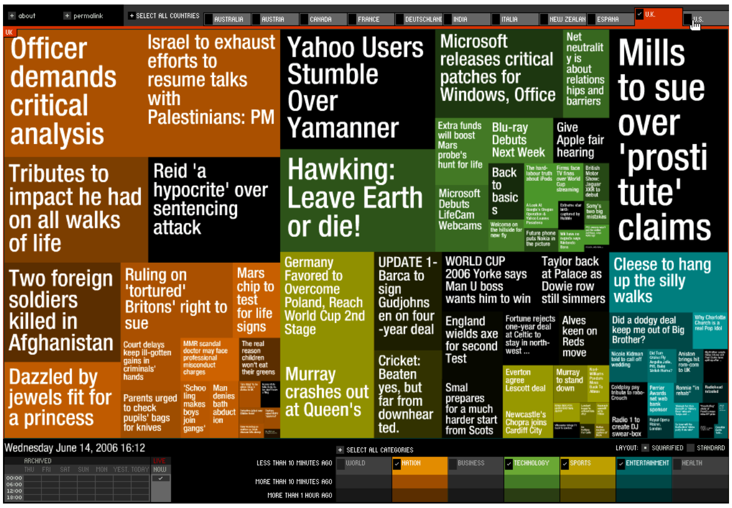


Figure 30: Newsmap - a treemap visualization based on news headlines from Google News

2.3.6.5. Alternative Focus+Context Visualizations

The Digg swarm interactive visualization by Design Studio Stamen (Stamen, 2006) represents another focus + context interface, in which the aggregation of social news is being visualised by circles which grow in size as they become more popular. Digg swarm also visualises relationships between stories and the actual users who 'dugg' the stories. Digg swarm's approach to revealing more detail about a news story while preserving the actual display of context effectively is seamless and not unlike the approach promoted as part of the original Focus-Metaphor approach (Laqua & Brna, 2005). With the sale of Digg to Betaworks in the summer of 2012, all labs projects have sadly been taken offline (Figure 31).

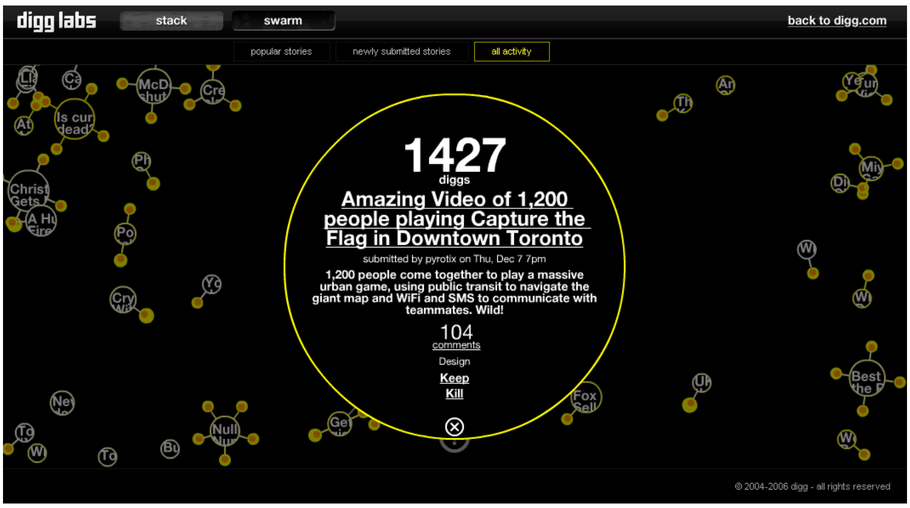


Figure 31: Digg's swarm visualization

2.3.7. Context-Aware Information-Retrieval

Somewhat orthogonally to how focus+context interfaces are designed to provide the user with some contextualization of the information they are attending to, context-aware information retrieval intends to contextualize the user using a system through attributes such as location, time of day, weather, likes or interests of the user, etc. (Brown & Jones, 2001). Context-aware information retrieval can be described as one facet of context-aware applications (Schilit, Adams & Want, 1994).

In Brown et al. (Brown, Burleson, Lamming, Rahlff et al., 2000), the authors identify six types of context-aware applications:

- Proactive triggering
- Streamlining interaction
- Memory for past events
- Reminders for future events
- Optimizing patterns of behaviour
- Sharing experiences

The research reported in this thesis is primarily concerned with aspects of proactive triggering, as well as streamlining interaction and optimizing patterns of behaviour.

Context-aware applications utilizing proactive triggering utilize *"the user's current context and trigger information to be presented"* (Brown, Burleson, Lamming, Rahlff et al., 2000). While this user's current context could be a location, time of day, etc., it could just as well refer to the information a user is currently attending to, or the application context she is in. The key concept is the fact that the system proactively adapts based on some triggering context. This aspect is further discussed in related research on proactive search (see section 2.4.1.3) and adaptive systems (see section 2.4.2).

The concept of streamlining interaction is intended to describe systems that optimize the process of human-to-human interaction, e.g. in the exchange of information. Interestingly, this can be interpreted as literally supporting users in the exchange of documents or other materials. However, in the context of increasing information overload (see section 2.1.2), or more specifically email overload for knowledge workers (see chapter 7), a specific case of streamlining interaction could be to refocus human-to-human interaction to the most essential communication only.

The other relevant type of context-aware applications is concerned with optimizing patterns of behaviour. Such applications typically aim to alter user behaviour through the provision of additional good-to-know information. This approach is at the heart of the KnowDis prototype work and its interpretation of the just-in-time information paradigm to support knowledge workers information goals (see chapter 7).

However, the majority of research into context-aware computing and applications is related to advances in mobile computing and the ability to build systems that adapt to contextual information collected from sensors that perceive the physical world. Back in the 90ies, Shilit et al.'s vision of a mobile distributed computing system was reportedly inspired by the vision of ubiquitous computing and focused on aspects of context that inform *"where you are, who you are with, and what resources are nearby"* (Schilit, Adams & Want, 1994). The work in this thesis is not concerned with location-based context-awareness, although future work based on this thesis (see section 0) has investigated this aspect as well (Beeharee, Laqua & Sasse, 2011).

2.3.8. Summary

Section 2.3 discussed the interaction perspective of the information overload problem by reviewing the different techniques used for human-information interaction. First, this section introduced the main categories of navigation on the web - *general purpose browsing* and *directed search* (Teevan, Alvarado, Ackerman, *et al.*, 2004). It then analysed some of inherent problems of web-based navigation by linking it to processes of learning and knowledge building and by discussing the relevance of designing interactions with information in a way that takes the limitations of human attention and cognition into account. Section 2.3 then reviewed relevant literature on search interaction and search strategies, followed by an analysis of exploratory search systems and focus + context interfaces. This goal of this section has been to review literature that provides further evidence for one of the key assumptions made in this research – that traditional browsing and specifically keyword-based search are not providing effective enough solutions to support users increasingly complex information needs. This section also intended to provide a reasonably broad overview of a range of alternative approaches taken by other researchers and the systems they built to tackles these problems – primarily exploratory search systems and focus+context interfaces –

as well as articulating how these systems and the concepts underlying them differ from the approach taking in this research.

2.4. The Computational Perspective

The following chapter discusses the computational perspective of the information overload problem. In doing so, the chapter explores a range of computational aspects related to how humans interact with computer systems and as relevant to this research.

Garcia-Molina et al. distinguish between three *information-providing mechanisms* which are able to satisfy information needs of users (Garcia-Molina, Koutrika & Parameswaran, 2011):

1. A search mechanism
2. A recommendation mechanism
3. An advertisement mechanism

Garcia-Molina et al. acknowledge the scale of the information overload problem by referring to it as a *"'deluge of data' in our workplaces and our homes"* (Garcia-Molina, Koutrika & Parameswaran, 2011). The authors propose a model to unify all three information-providing mechanisms (see Figure 32), as a way to only surface information to the user that is *"of interest and relevance, at the right place and time"* (Garcia-Molina, Koutrika & Parameswaran, 2011). On a technical level, each of the mechanisms aims to serve the same goal - utilising some context in the form of either an (explicit) query, the user's location or history to match against a *information objects* such as *ads, product descriptions,* or *web pages*. While Garcia-Molina et al. appreciate the potential differences among the mechanisms (in terms of delivery mode, beneficiaries, etc.), they stress the interchangebility of associating object type and information-providing mechanism (Garcia-Molina, Koutrika & Parameswaran, 2011).
This observation by Garcia-Molina et al. is significant in that it suggests that information should be treated in a generic manner, independent of its *object type*. One might even argue that the meaning or *character* of information is transient, depending on the unique interpretation attached to it by different users at a particular place, time and usage context. The same information can be provided as a targeted ad on Facebook, a recommendation on a site like Amazon, or appear as a search result in Google. In either case, it may or may not be relevant, depending on the accuracy of the user

context. Not infrequently, information that could be objectively characterised as useful to a user, may still be subjectively perceived as irrelevant as it reaches the user's attention at the wrong time and is consequently interpreted as noise - thus potentially ignored or even perceived as spam.

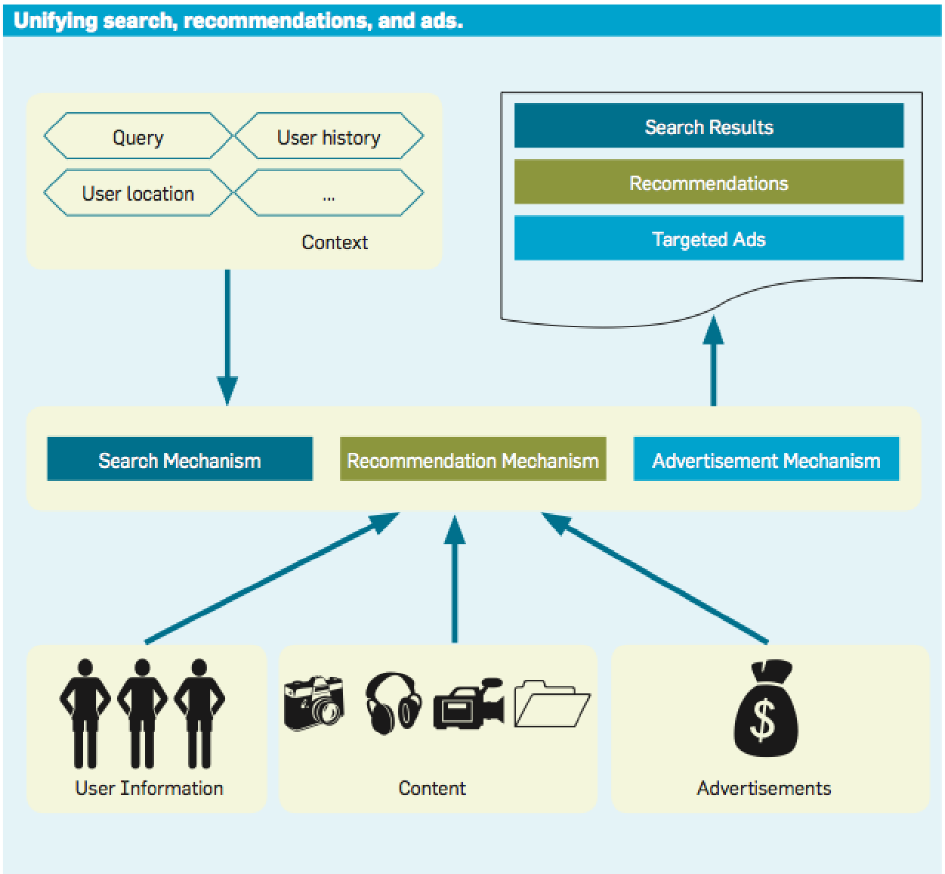


Figure 32: Model to unify search, recommendations and ads by (Garcia-Molina, Koutrika & Parameswaran, 2011)

In the context of this research and its assessment of information overload, the facets of the recommendation mechanism and the advertising mechanism are converging. The key differentiator between an advertisement and a recommendation - the financial gain of the ad provider - is out of scope of this thesis. Both recommendations and adverts utilise the same or very similar technologies that are also underpinning search. In addition, all three mechanisms are increasingly influenced by personalisation techniques.

The key differentiator relevant in this research for a discussion of search and recommendations (including adverts) is the delivery mode. A user typically 'pulls' search results using a search query, while recommendations are being 'pushed' to the user.

2.4.1. Search Systems

Human interaction with search engines has been discussed in section 2.3.2 and highlighted some of the problems around the paradigm of keyword-based search. The following sections review and analyse trends for augmentation of the original search interaction paradigm - *contextual search*, *personalised search* and *proactive search*.

2.4.1.1. Contextual Search

Kraft et al. define contextual search as a mechanism to proactively augment a user-generated search query by capturing a user's information need based on information they are currently browsing (Kraft, Chang, Maghoul, *et al.*, 2006). The authors' interpretation of context differs from the definition of context in information visualisations in that it is not applied when search results are presented but rather when a user queries the search engine. Kraft et al. present and evaluate three different algorithms to enable contextual search as per their definition (Kraft, Chang, Maghoul, *et al.*, 2006):

1. Query rewriting (QR)
2. Rank-biasing (RB)
3. Iterative filtering meta-search (IFM)

The algorithms presented aim to extract task context from *"the web page the user is currently browsing or a file the user is currently editing"* (Kraft, Chang, Maghoul, *et al.*, 2006). As such, their approach follows the philosophy of attempting to provide perfect search results to the user. While the authors' approach does not aim to give the user a choice and improve the user's understanding of the information retrieval process, the same algorithms could and should potentially be used as a tool to not just improve search results but also to enable a dynamically generated and more visual (search) context.

Another approach to contextualise search in a more user-controlled manner is search clustering. Vivisimo is one such search engine (and company), which emerged out of research at Carnegie Mellon University by Palmer et al. (Palmer, Pesenti, Valdes-Perez, *et al.*, 2001) on *hierarchical conceptual clustering.* The authors' initial implementation focused on the clustering of a large video library, and eventually grew into a dedicated search engine (see Figure 33). While the original vivisimo search engine is not available anymore, the company (also named Vivisimo) which developed it, built a newer version, called Clusty (see www.clusty.com), which it sold to Yippi Inc. in 2010 and again rebranded to Yippy (see search.yippy.com). The original Vivisimo company founded by three of the authors of (Palmer, Pesenti, Valdes-Perez, *et al.*, 2001) was recently sold to IBM. In contrast to the approach taken with Yahoo! Mindset, the search clustering offered by Vivisimo/Clusty/Yippy provides a richer characterisation of search facets which can be explored. Recently, Microsoft's search engine Bing started to offer a mechanism to emulate search clustering by recommending 'related searches' as a way to explore more specific facets of the original search query.

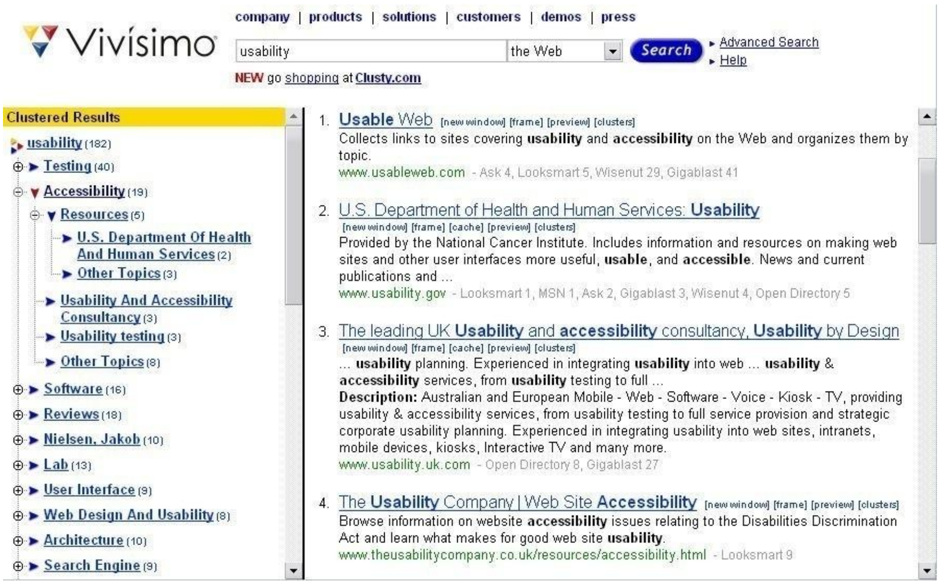


Figure 33: Screenshot of Vivisimo's Clustering Search Engine

2.4.1.2. Personalized Search

Micarelli et al. (Micarelli, Gasparetti, Sciarrone, *et al.*, 2007) provide an overview of personalised search on the Web, distinguishing between content-based and collaborative-based approaches (see Figure 34). The authors argue that personalised search is *"a potential solution to the information overload problem"* as such personalised search systems should be able to recognise specific user goals and *"predict aspects of [users'] future behavior"* (Micarelli, Gasparetti, Sciarrone, *et al.*, 2007).

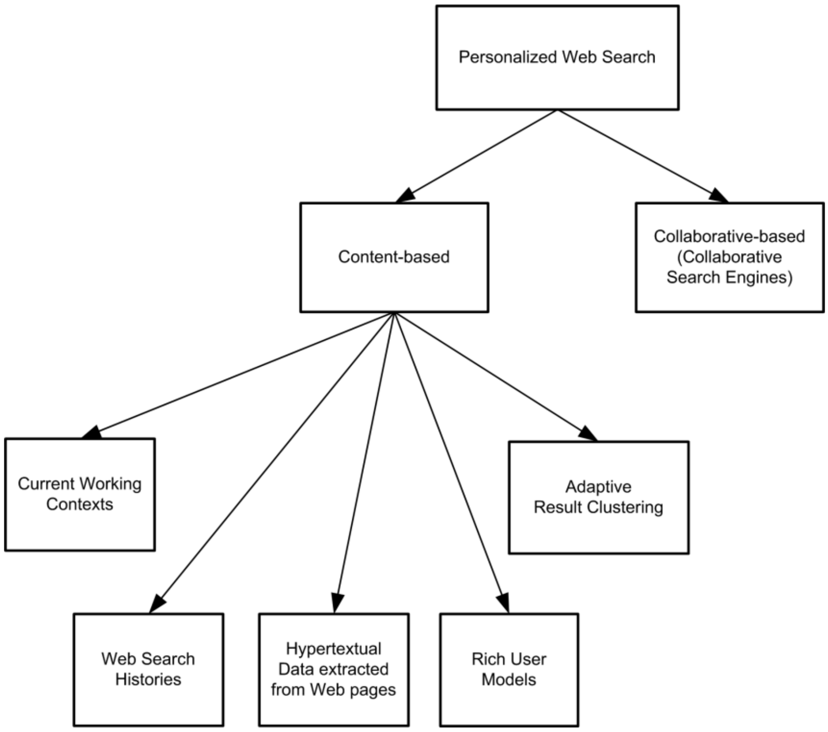


Figure 34: Breakdown of approaches for personalised web search (Micarelli, Gasparetti, Sciarrone, *et al.*, 2007)

Pretschner and Gauch (Pretschner & Gauch, 1999) propose ontology based personalised search to utilise the information interests of users for improved search results. The authors' approach represents early work in modelling complex (ontology based) user profiles and utilising them to personalise the display of search results through filtering or re-ranking - at a time, when Google was still running from a garage. Pretschner and Gauch found that the generated user profiles *"reflect actual user interests quite well"* (Pretschner & Gauch, 1999), and utilised those profiles to evaluate both search result re-ranking and search result filtering. In their assessment of both methods, re-ranking offered better results.

Pitkow et al. (Pitkow, Schütze, Cass, *et al.*, 2002) present a system called Outride, in which the mechanism for search personalisation is extended beyond Pretschner and Gauch's notion of user information interests by attempting to *"actively adapting the computational environment"* (Pitkow, Schütze, Cass, *et al.*, 2002). The Outride system takes the form of a sidebar within Microsoft's Internet Explorer offering a number of ways to access more targeted search results based on a user model which utilises contextualisation and individualization (Pitkow, Schütze, Cass, *et al.*, 2002) to personalise search. Pitkow et al. define contextualisation as *"interrelated conditions that occur within an activity"*. This includes factors such as the type of available information, the information currently looked at, and information about which applications are being used and the contexts in which they are used (Pitkow, Schütze, Cass, *et al.*, 2002). Individualization defines the set of characteristics that make the individual unique and thus distinguishable from others. This includes factors such as the goals of the individual, their prior and tacit knowledge, as well search histories (Pitkow, Schütze, Cass, *et al.*, 2002).

While the authors' system does sound promising, their evaluation does raise some questions in that it did not utilise actual user models of the participants, but rather *"a default user model"* (Pitkow, Schütze, Cass, *et al.*, 2002). Another aspect of concern is that no mention is made of the actual search tasks used to test the system, other than that they have been carefully chosen with an independent testing entity called eTesting Labs.

Ma et al. (Ma, Pant & Sheng, 2007) present a mapping framework, which maps interests of a user onto *"a group of categories in the Open Directory Project (ODP)"*. The authors' solution (PCAT) attempts to augment traditional search engines at the front-end by utilising user interests and taxonomy categories to augment the initial search results as received (see Figure 35). Similar to the system by (Pretschner & Gauch, 1999), Ma et al. made use of both search result filtering and re-ranking.

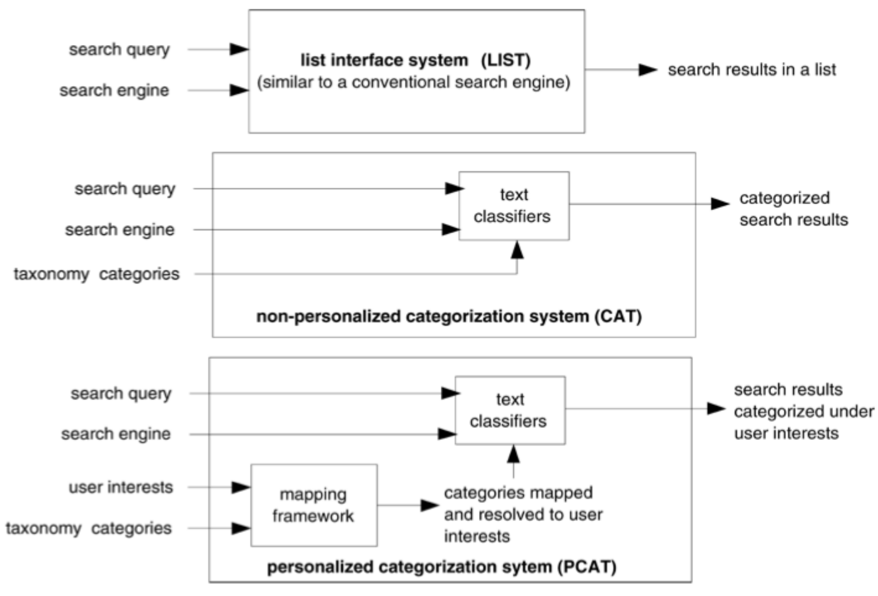


Figure 35: Overview of the three systems evaluated by (Ma, Pant & Sheng, 2007)

Another interesting approach to search personalisation comes from Speretta and Gauch (Speretta & Gauch, 2005), who explore the use of user search histories as a way to gather user information in a more unobtrusive manner. The authors evaluate the use of search queries and examined search result snippets to provide personalisation and found that both approaches were similarly effective.

While all of those search personalisation techniques and algorithms are quite promising in themselves, it is quite extraordinary that their application is commonly only tested in the context of re-ranking or filtering an original search result list. Given the domain, in which those systems are being developed, it is understandable that performance is of primary concern with regards to the evaluation of success or failure of a new personalisation technique. However, such evaluations tell little about the potential impact such a personalisation technique could have in the real world.

2.4.1.3. Proactive Search

> *"I ACTUALLY think most people don't want Google to answer their questions, they want Google to tell them what they should be doing next." Eric Schmidt (then CEO of Google, NYT interview)*

Recently, several researchers have explored the use of contextual search algorithms to develop proactive search capabilities (Billsus, Hilbert & Maynes-Aminzade, 2005; Kraft, Maghoul & Chang, 2005; Lawrence, 2000). In these studies, keywords are extracted from a current document in a browser and these keywords are used to find other related documents. Links to the related documents are presented in an information tool bar. These studies examine the usefulness of proactive search, in the context of search activities (i.e. while the user is browsing).

However, searching for information is often an interruption of other work activities. The typical purpose of search is to close a knowledge gap that was discovered during another activity, e.g., reading an email, composing a document, coding or upgrading software, designing a user interface, etc. - switching from one application to another takes time and attention, and often leads to further interruptions.

Proactive search techniques offer another approach to reducing the information overload associated with email. This approach is not focused on organising email or associating email with particular tasks. Instead the aim is to facilitate comprehension by providing additional context (related emails and documents). Billsus et al. refer to the unknown unknowns to argue that work *quality*, *efficiency* and *satisfaction* suffer *"when we are unaware of relevant information and human resources"* (Billsus, Hilbert & Maynes-Aminzade, 2005). Automatically generated recommendations are one way to assist people in discovering relevant information. Search engine providers (e.g., Google) and retailers (e.g., Amazon) use this technique frequently to display targeted advertisements that are related to the user's search query or the shopper's product selections.

2.4.2. Adaptive Systems

The problem of information overload has been outlined in section 2.1.3, and how it relates to us as an individual has been analysed in section 2.2 - the human perspective. How keyword-based search or traditional browsing fail to address the causes of information overload and only mitigate its effects has been discussed in sections 2.3.1 – 2.3.6. Research related to adaptive systems is one of the most promising areas to tackle the information overload problem - it involves research from communities such as *hypertext, user modelling, machine learning, natural language generation, information retrieval, intelligent tutoring systems, cognitive science,* and *web-based education* (Brusilovsky & Maybury, 2002).

In his work on *"Adaptive Hypermedia"* (Brusilovsky, 2001) and *"From Adaptive Hypermedia to the Adaptive Web"* (Brusilovsky & Maybury, 2002), Brusilovsky explains why a *one-size-fits-all* approach to web-based information systems does not work, and why systems are needed that can adapt to a user's information goals, tasks, personal interests or other characteristics by building an *explicit user model* around such user information. Brusilovsky differentiates between *adaptive systems* - which automatically adapt to the user - and *adaptable systems* - which allow the user to specify how a system should adjust. While such a distinction makes sense from a theoretical point of view, in practice, it is quite likely that a system will combine aspects of automatic adaptation and user-controllable adjustments. For example, personalised search (e.g. using location information, search history, etc.) is typically executed automatically, but - for privacy reasons among other things - can be controlled by the user. Within adaptive systems, a distinction is made between *adaptive content selection* and *adaptive navigation support* (Brusilovsky & Maybury, 2002).

2.4.2.1. Adaptive Content

Bunt et al. provide an overview of a range of computational techniques, which enable the provision of *"a tailored presentation of content"* (Andrea Bunt, 2007). The authors argue that adapting content to user needs effectively requires content adaptation and presentation. Bunt et al. define content adaptation as a mechanism to decide *"what content is most relevant to the current user and how to structure this content in a coherent way"* and content presentation as a mechanism to decide *"how to most effectively adapt the presentation of the selected content to the user"* (Andrea Bunt, 2007).

Bunt et al. differentiate three major strategies for content adaptation. First, the page-variant approach (Kobsa, Koenemann & Pohl, 2001) utilises pre-created versions of a page, from which a selection will be made from depending on the interaction context. Second, the fragment-variant approach is similar in nature to the page-variant approach but offers *"a finer level of granularity"* (Andrea Bunt, 2007), as it utilises pre-created fragments of information, which are selected and combined based on a) conditions met in an interaction context, or b) the background knowledge of the target user. Third, adaptation based on abstract information describes a set of approaches utilising techniques such as natural language generation to generate content-based abstract terms. While this body of work is potentially very relevant, its application is out of scope for this thesis.

In their discussion of content presentation, Bunt et al. (Andrea Bunt, 2007) discuss media adaptation techniques - choosing between text, graphics, spoken language, etc. to convey some information to the user. But their focus of discussion is on relevance-based techniques that allow maintaining focus and/or maintaining context within the content being presented. While not explicitly stating it, Bunt et al. are touching on the fairly large body of work on focus + context interfaces (which are being discussed in detail in section 2.3).

2.4.2.2. Adaptive Navigation

Brusilovsky defines adaptive navigation support as a set of technologies that process information needs, user preferences and background knowledge of a user to enable the adaptation of navigation mechanisms of a system (Brusilovsky, 2007). The author identifies five major technologies used to enable adaptive navigation (Brusilovsky, 2007):

1. *"Direct guidance"* - the use of a user model to recommend the 'next best' link for user to attend
2. *"Link ordering"* - the use of a user model to re-rank a list of links, such as a search result list

3. *"Link hiding"* - the use of a user model to hide, remove or disable irrelevant links
4. *"Link annotation"* - the use of a user model to "*augment the links with some form of annotation"*
5. *"Link generation"* - the use of a user model to dynamically generate either temporary or permanent new links as a means to recommend new content or to augment existing content with additional links

The first four technologies could be classified as more conservative adaption mechanisms in that they expect a traditional navigation mechanism as a source for adaptation, such as a default search engine result page (SERP). Conceptually, those four adaptation technologies are used to augment the default navigation mechanism, and thereby increase the information scent of that page based on e.g. a user model. The firth technology - *link generation* - is more innovative in that it does not imply the existence of a default navigation mechanism. Instead, link generation - as the name suggests - is generative in that it could be used to generate an entirely new navigation mechanism purely based on e.g. a user model.

2.4.2.3. Adaptive Personalisation

Billsus et al. advocate the use of adaptive personalisation, which describes systems that learn users' interests through the use of examples and then adapt when those interests change (Billsus, Brunk, Evans, *et al.*, 2002). In their work, the authors focus on adaptive interfaces for mobile use and propose a set of more general *"requirements for adaptive personalization and adaptive interfaces"* that are vital for successful user acceptance (Billsus, Brunk, Evans, *et al.*, 2002):

1. *"Provide a good initial experience and learn quickly for new users"* - The system should prevent a cold-start problem, and rather offer a non-personalised experience in the beginning, than waiting for some initial input by the new user.
2. *"Adapt quickly to changing interests"* - The system should adjust quickly to changing user tasks or interests.
3. *"Avoid tunnel vision"* - The system should support serendipity, making it easy for people to discover interesting new information, rather than just lock in users through the personalisation mechanism.

4. *"Do not require hand-tagging of content"* - The system should not require extra input by the user.
5. *"Avoid brittleness"* - The system should be robust in its personalisation, not changing too quickly on potentially irrelevant information, or at least offer a quick way of recovery.
6. *"Support multiple modes of information access"* - The system should provide different paths to information.
7. *"Respect individuals' privacy"* - The system should cater to individual users' privacy preferences and provide anonymity or a non-personalised mode of interaction.

While all of the above requirements are sensible, they are still only generic guidelines to validate the design of a system against. However, these requirements do not offer any insight into how adaptive systems could be designed to counter problems of choice overload, or support decision-making

2.4.3. Personalisation - An Educational Reasoning

In educational psychology, constructivism - first introduced by Immanuel Kant (1724-1804) as a developmental theory (Bagnoli, 2011) - is a theory of personalised learning experiences based on the self-construction of knowledge.

Von Glasersfeld describes a more recent variant of constructivism - the theory of **radical constructivism** - which suggests that knowledge is in persons' heads, who have "*no alternative but to construct what he or she knows on the basis of his or her own experience"* (Glasersfeld, 1995). The author further underlines the uniqueness of the human mind by arguing that *"all kinds of experience are essentially subjective"* and although we may think an experience to be similar to the one's we shared it with, we have *"no way of knowing that it is the same"*.

When exploring information on the web, it is crucial to understand such mechanisms of learning when designing information interfaces. Particularly when designing contextual information experiences, it is necessary to be aware of constructivist theories. As persons' subjective experiences of an information space are strongly influenced by their unique prior experiences, understanding context is a unique process.

Of course, aspects of task-focus play a role too, but it could be argued that they are more similar across individuals and thus easier to model (e.g. defining a set of tools required to write an email). In contrast, personalising contextual information to provide a subjective experience of the process/task may be much harder to achieve and requires more sophisticated strategies (e.g. defining the information required to respond to a specific email). In very restricted content domains, e.g. in map applications, which are a common early field of research in fisheye view interfaces, a generic approach and display of information may suffice as the map artefact is a shared mental model. However, as soon as domains gets more diverse and the information need more complex, a generic visualization will be decreasingly effective. The consensus between constructivist learning and a 'constructivist visualisation' of information may be the acknowledgement that making sense of information is an individual, unique process. Ultimately, this implies that contextual information interfaces need to be personalised to support the individual characteristics of each user. Finally, when considering personalisation in the design of a specific user interface, it is necessary to distinguish two very different types of personalisation: *personalisation of content* vs. *personalisation of structure.*

2.4.3.1. Personalisation of Content

The use of *personalisation of content* may be best illustrated by Amazon's use of personalisation, which uses a recommendation engine to adapt the display of products to the user's profile but by keeping the surrounding information structures and the visual layout of the site rigid. Nearly a decade ago, Amazon started using *"item-to-item collaborative filtering"* (Linden, Smith & York, 2003) to *"personalise its Web site to each customer's interests"*. Due to Amazon's vast scale even back then with *"tens of millions of customers and products"* (Linden, Smith & York, 2003), the authors found that no existing recommendation algorithm would be applicable. Consequently, Linden et al. developed their own solution which allowed scaling to *"massive data sets and produces high-quality recommendations in real-time"* (Linden, Smith & York, 2003). The authors further compare their solution with various recommendation algorithms including traditional collaborative filtering, cluster models and search-based methods.

Their main concern about traditional collaborative filtering is the inverse dependency of performance of the algorithm and recommendation quality. This might not apply to relatively small data sets (with thousands of entities – users and content items), but *"for* *very large data sets – such as 10 million or more customers and 1 million or more catalogue items – the algorithm encounters severe performance and scaling issues"* and improving the performance would *"reduce recommendation quality in several ways"* (Linden, Smith & York, 2003).

In their discussion of cluster models, Linden et al. point out the superior online scalability and performance when compared to collaborative filtering but at the price of classifying users into clusters of similarity, which leads to low recommendation quality and less relevant recommendations (Linden, Smith & York, 2003).

The third category analysed are search-based methods, which utilise search algorithms to construct search queries to *"find other popular items by the same author, artist, or director, or with similar keywords or subjects"* (Linden, Smith & York, 2003). However the authors identify two problems with search-based methods:

1. With growing data sets, search queries have to use a *"subset or summary of the data"* which results in reduced quality of the personalisation
2. Dependent on the nature of a query, recommendations are often too general or too narrow and thus do not deliver *"interesting items"*

2.4.3.2. Personalisation of Structure

The use of *personalisation of structure* is much less explored or even scientifically defined than the use of personalisation of content. While personalisation of structure implies the use of *link generation* to support *adaptive navigation* (Brusilovsky, 2007), it is meant to address more than just the navigation mechanism. Instead, personalisation of structure suggests that due to the individuality of each user, it is not just important to adapt what navigational choices to display, but also how to display them. Whereas one user may prefer a grid-like display of richly annotated navigational choices, another may prefer a basic list, and a third may prefer a tree-like structure. Such personalisation of structure may be bound to the type of task a user engages in, but it could also be envisioned that individual cognitive abilities impact the most effective visual arrangement as well as the number of elements that should be displayed.

A somewhat related technology, albeit without any dynamic ability of personalisation, are 'personalised home pages' such as MyMSN (http://uk.my.msn.com/), MyYahoo (http://uk.my.yahoo.com/), iGoogle (http://www.google.co.uk/ig) or Netvibes (http://www.netvibes.com/en). While personalised home pages are not adaptive in any way other than that new content is being surfaced automatically (such as weather updates, or news stories in a news widget), they allow the user to customize the presented information by adding, removing and arranging content widgets. It could be argued that through those content widgets, some type of manual personalisation is achieved that is based on custom subscriptions to information channels.

2.4.3.3. Awareness of Personalisation

The concept of just-in-time information interfaces aims to deliver only information relevant to a user's task in a minimalist manner, such that it aligns with users' cognitive information processing capabilities. Following this approach when displaying contextual information should increase users' awareness of relevant information. However, when discussing personalisation of content, some critical issues need to be considered. Lynch discusses new directions for personalisation and recommender systems, suggesting that users *"want to be aware of certain kinds of information, even if upon examination [they] determine that [they] don't like it, or disagree with it, or think that it's rubbish"* (Lynch, 2001). The often articulated danger with personalisation of content is that people are no longer being confronted with different points of view, or with contrasting opinions. When the provision of information is being adjusted to people's individual beliefs, processes of critical thinking are undermined.

2.4.4. Summary

Section 2.4. discussed the computational perspective of the information overload problem by reviewing relevant technologies used for search systems, such as contextual search, personalised search and proactive search. This section further discussed the use of adaptive systems to provide adaptive content, adaptive navigation and personalisation. This section finished with an educational reasoning for personalisation, by outlining the need for a distinction between personalisation of content and personalisation of structure.

This goal of this section has been to review literature that provides an overview of the range of computational concepts used to design contextual, proactive, adaptive or personalized systems. Firstly, this supports positioning the work in this thesis and its use of contextualisation, proactive information delivery and adaptive navigation in the wider research context. Secondly, while it has been out of scope to incorporate very advanced versions of these concepts into the design of the prototypes build and evaluated in this thesis, this section is also meant to illustrate the potential for future iterations of just-in-time information systems and services that make more effective use of such advanced computational concepts.

3. Methodology

This chapter provides an overview and discussion of the methods employed in this thesis to investigate how to design just-in-time information services to improve the user experience of goal-driven interactions with information. This chapter is not a detailed account for the specific application of the various methods in each study – those are included in a dedicated 'method' section for each study.

First, section 3.1 explains the motivation behind the specific research methods employed in this thesis. Second, section 3.2 provides an overview of empirical methods for HCI research that are relevant for this research. Finally, section 3.3 provides a brief overview of the specific methods applied in this thesis. This section includes a methodology matrix that maps all conducted studies onto the prototype system used, the context of use (lab-based or field study), the specific methods employed (eye tracking, etc.), the evaluation focus (summative or formative), the number of participants, and the usage context (task types, etc.) (see section 4.3.1). A detailed discussion of the particular research method used for each study can be found in the method section of each study chapter.

Section 4.3.2 elaborates on the usage contexts employed, specific measures taken, type of analysis conducted, and discusses specific challenges of ecological validity in each context. Section 4.3.3 concludes with an overview of the actual research processes employed in the lab-based and field studies.

3.1. Motivation / Context / Design Research

The traditional model of interacting with digital information - browsing by following links on a web page, or teleporting (Teevan, Alvarado, Ackerman, *et al.*, 2004) by using a search engine - has to date remained essentially unchallenged.

Carroll & Kellogg (Carroll & Kellogg, 1989) argued over two decades ago that innovations in the design of user interfaces typically lead research into human-computer interaction. Traditionally, theory would inform creation of new technologies or the design of new systems to application.

More recently, Stolterman (Stolterman, 2008) elaborated on the *"apparent paradox of HCI application leading HCI theory"* (Carroll & Kellogg, 1989) by suggesting that *design practice* is not sufficiently developed within the HCI research community and that a greater understanding of *design rigor* and *design discipline* need to be a priority of the research agenda of HCI research.

Two fundamental technological leaps, (1) the application of the concept of hyperlinks (as first coined by Ted Nelson) on the world wide web (as created by Tim Berners-Lee), and (2) the rise of keyword-based search through the conception of the PageRank algorithm by Larry Page and Sergey Brin have shaped people's experience of interacting with digital information – and their expectations. Those two leaps, while grounded in theory, are today utilised and the *de-facto* standard of interaction in very different usage contexts than originally envisioned.

The history of hyperlinks can be traced to Vannevar Bush (Bush, 1945), who describes a machine which allows linking different microfilms in a way that one could read related information from different microfilms in a continuous manner.

The history of PageRank (Brin & Page, 1998) can be attributed to work on classic citation analysis in academic literature. Page & Brin built PageRank to utilise the link structure on the web to model relationships between individual pages by interpreting every single link in a manner akin to an academic citation (Page, Brin, Motwani, *et al.*, 1998).

Following the argument that HCI design practice typically leads HCI research, it seems that those two technologies - the linking of documents and keyword-based search (embodied in HTML and PageRank) - are in a grander context concrete examples of interface innovations that have motivated subsequent research in HCI, rather than being continuously challenged by it.

The scientific method employed in this thesis, started from the working hypothesis that neither the traditional model of linking documents nor keyword-based search are sufficiently suited to support users to effectively deal with complex information needs and the mass of information they encounter today.

Carroll & Kellogg argue that traditional performance analyses, which only measure low-level isolable units, do so at the cost of ignoring the higher-level context and interaction effects (Carroll & Campbell, 1986). To overcome this problem, the authors recommend a mutual relationship between "*the articulation and rearticulation of a set of psychological claims and the iterations of design"* (Carroll & Kellogg, 1989).

The research presented in this thesis takes an exploratory approach, utilising a process of iterative development of research prototypes to explore the design of information interfaces for information discovery and exploration.

3.2. Overview of Relevant Empirical Methods

This section discusses the methods employed in empirical research, distinguishing between formative and summative evaluations, as well as qualitative and quantitative data collection and analysis techniques.

3.2.1. Formative and Summative Evaluation

Formative user-centred evaluations are *empirical*, *observational* methods that involve users in an iterative manner to assess the usability of a system (Gabbard, Hix & Swan, 1999). Formative evaluations focus on aspects of *usage context* and *ecological validity* in order to understand why a particular design decision is or is not working and to inform the design of future iterations. Such evaluations typically employ both qualitative and quantitative data collection and analysis methods (see section 3.2.2) to provide a richer picture of users' experiences with a system.

Weibelzahl et al. (Weibelzahl & Weibelzahl, 2007) also refer to formative evaluations as a *generative method*, which integrates the specification of a design and its evaluations *"into the same framework"*.

Carroll's work on scenario-based design is another example of a formative method of evaluation (Carroll, 2002). In it, the author argues that in design-driven projects, it is not possible to identify and control for all potential facets of an evaluation, that may lead to an improved understanding of *learning*, *human development*, or *workplace culture* (Carroll, 2002). Carroll further suggests that a thorough understanding of work practices is essential to facilitate *"the emergence of new designs* (Carroll, 2002).

Summative comparative evaluations are an *empirical assessment* of multiple - typically matured - system designs that are assessed using the same experimental tasks in order to determine user performance or user satisfaction differences (Gabbard, Hix & Swan, 1999). The focus in summative evaluations is on quantitative results that are primarily collected in lab-based studies.

Weibelzahl et al. argue that when summative evaluations are used as final validation, it is preferable to plan several small studies, rather than just one large study in order to minimise the impact of potentially flawed experimental designs (Weibelzahl & Weibelzahl, 2007).

The research conducted in this thesis has - at a high level - a formative focus, as each iteration of a developed prototype informs the design of the next iteration of just-in-time information interfaces. However, one specific instance of a summative evaluation is present in Main Study 1 (chapter 6), where the developed prototypes are benchmarked against alternative interfaces using the same navigation tasks. In these studies objective data such as task times and error rates are measured to provide a direct comparison of the tested user interfaces. As part of this research, those summative evaluations were conducted to provide some validation for the general value of conducting the more high-level, iterative and ongoing formative exploration of just-in-time information interfaces. If for example, performance tests indicated a problem with the developed prototypes this could have hinted at more fundamental problems that would need to be addressed - for example by pivoting the design exploration in a somewhat different direction.

While Main Study 1 has both a summative and a formative evaluation focus, the remaining studies - Preliminary Study 1 (chapter 4), Preliminary Study 2 (chapter 0), and Main Study 2 (chapter 7) - have a primarily formative evaluation focus.

3.2.2. Qualitative and Quantitative Research Methods

The research questions determine the type of the data collected, and the methods used to collect and analyse that data (Punch, 2005). The most commonly used classification for research methodologies distinguishes between qualitative and quantitative approaches. Punch identifies a number of dichotomies to differentiate between the two approaches. According to Punch, qualitative research is *inductive*, investigates *natural settings* and aims to identify *cultural patterns* through the use of qualitative data (Punch, 2005). In contrast, quantitative research is *deductive*, investigates *artificial settings* and seeks to identify *scientific law* through the use of quantitative data (Punch, 2005).

In educational research, Ercikan & Roth have argued against a polarisation of qualitative and quantitative research, proposing an integrated approach instead (Ercikan, 2006). The authors suggest that such polarisation is neither useful nor effective for furthering scientific insight, as any observed phenomenon will provide insights from both quantitative and qualitative data collection. As an example, Ercikan & Roth refer to classroom observations, in which quantitative measures, such as type and frequency of interaction among students or student and teacher can be taken alongside the actual dialogues (Ercikan, 2006).

The inherent focus on the individual in educational research is an aspect that should receive greater attention within HCI research, particularly in the context of human-information interaction, where the satisfaction of information needs is evaluated. As we engage with information, we try to comprehend the encountered information, interpret it, and potentially utilise it to inform further action - thus engage in an episode of learning. The discussion of *learning loops* in the process of sensemaking (Russell, Stefik & Pirolli, 1993) illustrates the closeness of HCI and education research.

From a social research perspective, Olsen argues for mixing quantitative and qualitative methods to facilitate multi-disciplinary research, and to gain *"multi-perspective meta-interpretations"* (Olsen, 2004).

The research in this thesis touches a range of research disciplines with a strong focus on human-information interaction and information seeking. Following Olsen's and Ercikan & Roth's arguments as outlined above, this research thus combines qualitative and quantitative research methods.

3.2.3. Methods for Observing User Behaviour

Weibelzahl et al. (Weibelzahl & Weibelzahl, 2007) distinguish between direct and indirect methods for user observation. Direct methods, such as observation in the field, can take the form of *ethnographic* studies (Punch, 2005), which are typically more longitudinal in nature and strive for the observer's unbiased gathering of data through note-taking, surveys and/or interviews. Another direct observation method is *contextual enquiry* (Beyer & Holtzblatt, 1997), a more targeted field study method to observe user behaviour in a fairly short amount of time (e.g. a few hours) and gather feedback as a means to mitigate the risk of false interpretation of observed behaviour. In contextual enquiry, the observee's real-world tasks or activities (which are of interest to the observer) serve as a shared artefact for an observer to direct her enquiry. While user behaviour can be observed in lab-based experiments as well, the artificial setting often reduces the ecological validity of such behaviour.

Any method which requires some form of interaction between the observer and the observee causes some degree of interference with the observee's activities, which may 'contaminate' user behaviour.

Such contamination can be caused by the presence of an experimenter, leading to a change in the observee's behaviour. In more severe cases, an experimenter might interfere with the task or otherwise distract the observee. The change in behaviour may be different from what would happen in real-life settings without an observer.

Particularly in observations where performance of user behaviour is being measured, contamination is undesirable, as the observed outcomes may have been influenced. In these circumstances, or when observation is not feasible due to physical constraints (e.g. the observed may be spread across multiple countries), event or activity logging can represent a suitable method for indirect observation.

In this research, neither ethnographic studies or contextual interviews could reasonably be conducted. As Preliminary Study 1, Preliminary Study 2, and Main Study 1 all utilized very early stage prototypes, these were not mature enough to be deployed in the wild. In contrast, the prototype developed for Main Study 2 was built to a level that it could be deployed to the wild, however as participants were spread all across the globe, any on-site research would have required unreasonable amounts of travel.

3.2.4. Methods for Collecting User Feedback

The primary goal for collecting user feedback is to improve the researcher's understanding of users' opinions, preferences or motivations behind observed behaviour. Moreover, user feedback allows gathering demographic information, which can be particularly useful to identify individual differences and potentially group users/subjects for more detailed and nuanced analyses of other collected data.

3.2.4.1. Questionnaires

Questionnaires are the main quantitative method for *"the collection of [a] user's opinion"* (Weibelzahl & Weibelzahl, 2007), through the use of *multiple choice*, *ranked* or *scalar* questions. Ranked questions frequently utilise a Likert scale (Likert, 1932) to label alternatives using either an even or uneven number of options (Gamst, Liang & Der-Karabetian, 2011). If an uneven number of options is used, the middle option offers respondents the ability to pick a neutral alternative, such as 'neither agree or disagree'. Adams & Cox (Adams & Cox, 2008) point out that whilst there is some disagreement among researchers on whether or not to use a neutral option, in the end it is more important that the questionnaire itself is generally well-designed.

If a Likert scale with an even number of choices is used, a 'no opinion' option can be included. Scales which use an even number of options and omit a 'no opinion' option are referred to as *forced-choice* (Zavala, 1965).

The use of open-ended questions in questionnaires represents a simple way to collect user feedback in semi-structured form. In order to evaluate *"changes due to real or experimental user-system interaction"* (Peter Brusilovsky, 2007), a combination of pre- and post-test questionnaires can be used - the same questions being put to a participant before and after exposure to a lab experiment or a longitudinal field trial, and differences in responses measured.

In a recent review of user experience research, Bargas-Avila and Hornbaek found that across the 66 empirical studies they surveyed, questionnaires were the single most used method of data collection (53%) (Bargas-Avila, 2011). However, they raise concerns over the common use of *"self-developed questionnaires"* and *"ad-hoc scales"* (as opposed to validated ones), and suggest that more appropriate resources and instruments are readily available. The use of 'proven' questionnaires and scales, such as documented in (Lewis, 1995) is preferable to ad-hoc scales.

The problem with self-developed questionnaires is that results of such studies cannot be compared across similar studies, and thus findings cannot feed into the cumulative knowledge about the specific domain, user group, task or context. However, a validated data collection instrument can only be utilised if it fits the requirements of the specific research questions to be examined.

The research conducted in this thesis follows an exploratory approach, thus the verbatim use of such standardized questionnaires is either limited or not applicable. When appropriate, standardised questions and scales were re-used, such as the *Post Study System Usability Questionnaire* (PSSUQ), which is part of the IBM computer usability satisfaction questionnaires (Lewis, 1995), or the *Questionnaire for User Interface Satisfaction* (QUIS) (Chin, Diehl & Norman, 1988), which has been developed at the Human-Computer Interaction lab at the University of Maryland. The QUIS is currently in its 7^{th} version and has been validated by (Harper, Slaughter & Norman, 1997). The PSSUQ has also been thoroughly tested, and found to be well-suited for the effective measurement of user satisfaction in usability evaluations (Lewis, 2002).

Throughout the preliminary studies, as well as Main Study 1 and Main Study 2, questionnaires were used either as a primary means of collecting user feedback, or a secondary means to augment and contextualise other data collection methods, reusing either parts or all of QUIS or PSSUQ, when appropriate.

In particular, Main Study 1 stands out in this regard, as it collected a wealth of eye-gaze data, performance data, interaction data, as well as user feedback through questionnaires. In this case, detailed demographic questionnaires were used to profile participants (see section 6.4.2.1 for details) and enable a more refined analysis of performance, user satisfaction and gaze behaviour.

3.2.4.2. Interviews

Interviews allow a more flexible approach than questionnaires to gathering user feedback, such as *self-reported opinions and experiences, preferences,* and *behavioural motivations* (Weibelzahl & Weibelzahl, 2007), and are particularly useful in the context of formative exploratory evaluations.

Interviews range from structured, over semi-structured, to unstructured. In a structured interview, the researchers follows an interview script to elicit the interviewee's perceptions on a pre-defined set of topics, whereas in an unstructured interview, the researcher encourages the interviewee to talk about what they see as relevant and important.

Typically, a semi-structured interview may strike a useful balance in guiding the interviewer such that relevant aspects are covered and important issues that may not have been identified prior to the interview can emerge (Adams & Cox, 2008). Such aspects relate to how the environment, user goals, user preferences, prior knowledge, or any other characteristics related to an individual user impact on the use of a tool, or the engagement in an activity that is the focus of an investigation. Ideally, such interviews feel informal and more like a conversation, allowing the interviewee to elaborate in depth on aspects of particular interest to them (that are also related to the subject of the interview).

The research conducted in this thesis makes use of semi-structured interviews as part of Study 2, where the impact of embedding just-in-time information discovery (through the KnowDis prototypes) into knowledge-workers real-world activities is explored (see chapter 7). In this research, qualitative methods are used to explore the specific issues around the problem of information overload. Unstructured elicitation and data collection methods are used to identify emerging topics, e.g. to inform different types of users and generate persona profiles.

3.2.5. Methods for Data Analysis

Grounded Theory (Glaser & Strauss, 1967; Strauss & Corbin, 1990; Goulding, 1998; Furniss, Blandford & Curzon, 2011) and *Qualitative Content Analysis* (Mayring, 2000) are data analysis methods that utilise a systematic approach to support the generation of new theories.

Adams et al. (Adams, Lunt & Cairns, 2008) describe grounded theory as a *"method of qualitative research"*, which can make use of qualitative data as well as quantitative data to support the formulation of new theories. Grounded theory does not require the formulation of a hypothesis a priori, but instead can be used to study a particular area without specific expectations. Theories derived from the process of grounded theory can then be applied to formulate new hypotheses for further investigation (Adams, Lunt & Cairns, 2008). Grounded theory uses coding to analyse data from interviews or other forms of text through properties and dimensions. As Adams et al. point out, grounded theory is very time-consuming even for modest studies, thus it should be considered whether its use is appropriate and required to inform future iterations of a design.

Mayring (Mayring, 2000) describes qualitative content analysis as a *"bundle of techniques for systematic text analysis"*, which are grounded in content analysis techniques from communication science. Qualitative content analysis emphasises four aspects of qualitative text interpretation (Mayring, 2000):

1. Collected data needs to fit into a *model of communication*
2. Analysis of the data needs to follow step by step rules and procedures
3. Analysis focuses on the categorisation of text interpretations
4. Procedures used should ensure *reliability* and *validity* of the analysis

However, Mayring also notes that the procedural approach of qualitative content analysis may be less suited to *open-ended*, and *explorative* research questions, or *"if a more holistic [...] analysis"* is planned (Mayring, 2000) as the use of categories may be more restricting than it would be beneficial.

The research in this thesis conducted a substantial number of interviews as part of the field studies in Main Study 2 (see chapter 7). Applying a systematic approach of analysing this interview data through one of the methods above would have required significant amounts of time. Given the complexity of the conducted field studies, the implementation effort, the negotiation effort to actually deploy the developed prototype within a large IT organisation, the effort involved in conducting field studies over several weeks across time zones and continents, and the detailed analysis of other data, such as interaction and survey data, the use of grounded theory or qualitative content analysis could not be justified.

More importantly, it was felt that the very open-ended and iterative manner in which the feasibility of just-in-time information interfaces for information discovery and exploration was investigated, would not have been complemented appropriately by such an approach.

3.2.6. Prototyping

Prototypes are a vital aspect of any iterative design process. Prototypes can range from simple paper-based mock-ups, which may be appropriate for the validation of a (simple) visual design, to highly complex implementations that closely emulate (part of) a final product. But the actual prototyping approach taken depends on a number of factors, such as the type of design problem at hand, the design process used, and the skill set available as part of the design process. In general, prototypes are distinguished alongside three dimensions - fidelity, breadth, and re-usability.

Low fidelity prototypes are typically represented by rough sketches or wireframes that are cheap to produce and help communicate and evaluate visual design ideas in a basic manner.

High fidelity prototypes at the other end of the fidelity spectrum utilise high-level programming languages and typically implement parts of a system that closely resemble a potential first release candidate. Such prototypes are more effective to communicate and evaluate more complex concepts or novel interaction ideas that need to be experienced, rather than just looked at. High fidelity prototypes are also essential for any field-based evaluations.

Horizontal prototypes capture the breadth of the proposed system in terms of visually illustrating features and content. However, horizontal prototypes are shallow in that not all of the actual features are being implemented.

Vertical prototypes are utilised when only some features need to be represented. In a vertical prototype, these features will be modelled such that they resemble closely the potential feel of a first release candidate. A common prototyping process may combine high fidelity and vertical prototyping to explore the feasibility and potential user experience of a particular design concept.

Revolutionary (or throw-away) prototypes describe a category of prototypes purely created as a bi-product of the design process and to act as artefacts for evaluation purposes.

Evolutionary prototypes in contrast are meant to continuously improve upon each design iteration until a final prototype version may eventually transform into an actual release candidate version of the proposed system. As the process of evolutionary prototyping requires utilising the actual technologies needed for the release candidate, it is typically combined with vertical and high-level prototyping.

Dow et al. (Dow, Fortuna, Schwartz, *et al.*, 2011) suggest the use of multiple prototypes for improved exploration of the concepts and intentions underpinning said prototypes. While this seems to be a desirable goal, it needs to be noted that while Dow et al. speak of prototypes generically, specifically the objects of question they discuss are banner designs - or low-level prototypes which are reasonably quick to create. The time and effort involved in creating much more complex high-level prototypes most often prohibits the creation of multiple interpretations of a concept.

3.2.7. Contextual Design

Over two decades ago, Wixon and Holtzblatt (Wixon, Holtzblatt & Knox, 1990) introduced contextual design as an *"evolving collection of methods ... for building effective systems"*. They emphasised the importance of interviews - ideally conducted in a contextualised/situated manner - to build a better *"understanding of work in context"* (Wixon, Holtzblatt & Knox, 1990), and the involvement of users through iterative prototyping.

More recently, Holtzblatt et al. (Holtzblatt, Wendell & Wood, 2005) have adopted contextual design to fit well with the increasingly popular agile methods used by practitioners outside academia. This formulation of 'rapid contextual design' is promising, as it highlights the similarities between user-centred design techniques (as applied within academia) and agile development methods (as applied by practitioners), and thus encourages closer interconnection between system design and system development.

3.2.8. Eye Tracking

Eye tracking is a quantitative method utilising the measurement of eye movements to analyse a users' visual attention distributions on the stimulus provided, such as images, video, interactive systems, or even real-life scenes. Depending on the type of system used, the projection of eye-movements onto a target area is either relative to the position of the head, or absolute (for example a display in front of the user) also referred to as 'Point of Regard' (POR) (Duchowski, 2007). While the *"video-based corneal reflection eye tracker"* is the most commonly used system today (Duchowski, 2007), four categories of systems exist, which utilise different measurement techniques.

Electro-OculoGraphy (EOG) measures the electric potential of the skin using electrodes positioned around the eyes. This method is highly intrusive and while *"the most widely applied eye movement recording method some 40 years ago"* (Duchowski, 2007), it has been largely replaced by more unobtrusive methods.
Scleral contact lens/search coil uses a specially marked contract lens that has to be worn on the eye. While deemed the *"most precise eye movement measurement method"*, just as EOG, this method is highly intrusive and causes discomfort (Duchowski, 2007).
Photo-OculoGraphy (POG) or Video-OculoGraphy (VOG) measure *"distinguishable features of the eyes under rotation/translation ... and corneal reflections ... of a light source"* (Duchowski, 2007). POG or VOG combine a number of techniques which typically measure eye movements relative to a subject's head position, and thus require manual inspection of video recordings capturing eye movements and the subjects field of view. As this approach requires frame-by-frame assessment of the recorded materials, the frame rate is also the maximum sampling rate of the collected gaze data, and analysis is typically *"tedious and prone to error"* (Duchowski, 2007).
Video-based combined pupil and corneal reflection represent the de-facto standard eye tracking technology used in HCI research, but also in market research and increasingly for the development of systems utilising gaze-based interaction. This technique utilises one or several light sources (e.g. infra-red) to create corneal reflections either using dark or bright pupil reflections (Tobii, 2011). Some advanced eye trackers utilise a combination of dark and bright pupil reflection for more accurate measurements and an increased tolerance for extraneous factors such as *"experimental conditions and ethnicity"* (Tobii, 2011). Reference points are calculated from corneal reflections and the position of the pupil to account for eye rotation and head movements (Duchowski, 2007). This enables video-based pupil and corneal reflection eye tracking systems to measure *"a viewer's Point Of Regard (POR) on a suitably positioned (perpendicularly planar) surface on which calibration points are displayed"* (Duchowski, 2007).

Eye movements consist of positional and non-positional aspects (Duchowski, 2007). Non-positional aspects, such as pupil dilation and lens focusing are categorised as adaptation and accommodation movements and are of secondary importance with regards to positional movements. Duchowski identifies five such positional eye movements - saccadic, smooth pursuit, vergence, vestibular, and physiological nystagmus (Duchowski, 2007). However, only saccadic eye movements and their counterpart, physiological nystagmus, which are *miniature movements* of the eye, commonly referred to as fixations, are of practical relevance, as in the context of POR measurement in HCI research, the gaze path recorded and analysed consists of saccadic movements (or saccades in short) and fixations.

The key eye tracking measures analysed in HCI research are *time to first fixation*, *fixation count*, *fixation length*, and *observation length* (as measured in gaze time) for predefined *areas of interest* (AOIs) in the tested user interface or interface elements:

- *Time to first fixation* represents the time between the display of a stimulus (or specified AOI) and the first fixation within that stimulus (or AOI).
- *Fixation count* represents the number of fixations observed within a stimulus or AOI.
- *Fixation length* represents the sum of all individual fixation durations (e.g. in seconds), as measured under fixation count (and thus observed within a stimulus or AOI).
- *Gaze time* (or observation length) represents the combined time of fixation length and the duration of all saccadic movements between such fixations. While saccadic movements are very quick, ranging from 10 to 100 ms (Duchowski, 2007), their time is not negligible, and observation length or gaze time is typically distinctly longer than pure fixation length.
- *Observation count* represents the number of times a gaze path entered (and left) a stimulus or AOI.

Another relevant measure, derived from fixation length is *average fixation duration*, which represents the length of all fixations divided by the number of fixations (within an AOI). Average fixations are used to provide some indication of the level of visual processing required to parse a visual element.

3.2.9. Evaluating Exploratory Systems

White et al. (White, Marchionini & Muresan, 2008) argue that researchers have explored the development of exploratory search systems (ESSs), and interfaces which support exploratory search behaviour in recent years, but have paid little attention to unique requirements for effective evaluation of such systems. Instead, the focus of evaluation has rested with search systems which support simple lookup search, and which *"encourage minimal human-computer interaction"* (White, Marchionini & Muresan, 2008). White et al. acknowledge the need for experimental evaluations of retrieval algorithms or other aspects of search systems that are more closely related to pure information retrieval aspects, but suggest that with increasing complexity of search systems and increasing support for exploratory behaviour, more attention needs to be given to the human searcher: *"Search systems are not used in isolation from their surrounding context, i.e., they are used by real people who are influenced by environmental and situational constraints such as their current task"*.

Referring to (Ingwersen & Järvelin, 2005), White et al. argue that search systems need to be able to adapt to such contextual constraints to be used effectively, and that effective evaluation of such systems needs to make such contextual requirements part of its methodology (White, Marchionini & Muresan, 2008).

White et al. further suggest the measurement of learning effects as a means to evaluate exploratory search systems. Such evaluation of learning, while possible to measure based on efficiency through some sort of cost-benefit assessment (Pirolli, Schank, Hearst, *et al.*, 1996), should *"ultimately ... measure the depth and effectiveness of learning"* (White, Marchionini & Muresan, 2008). While White et al. liken exploratory search with sensemaking / information foraging activities as studied in depth by Pirolli and his colleagues at PARC, it seems that they differ in their interpretation of learning. Pirolli and colleagues may look at learning as something mathematically measurable and model-able, seeing learning as the process of making sense of information, whereas White et al.'s focus seems to be on the learning outcome.

As briefly touched on above, the effective measurement of learning is a challenge in itself, and increasingly so in the context of a highly dynamic system/environment. White et al. therefore suggest that *"the measurement of interaction behaviour [and] cognitive load"* helps getting *'a clear picture'* of the effectiveness of ESSs (White, Marchionini & Muresan, 2008). Other relevant aspects for the evaluation of ESSs are *"subjective measures such as user satisfaction, engagement, information novelty, and task outcomes"* (White, Marchionini & Muresan, 2008).

He et al. (He, Brusilovsky, Ahn, *et al.*, 2008) describe a framework for evaluating task-based information exploration in adaptive systems. While the authors link exploratory search primarily with *"professional users, such as intelligence analysts"*, activities such as information foraging and sensemaking are unquestionably relevant to 'normal users' as well. He et al. argue that their framework is tightly connected with the *complex, dynamic,* and *multi-facetted* tasks of intelligence analysts.

Qu and Furnas (Qu & Furnas, 2008) argue that effective assessment of exploratory search behaviour requires the examination of *"interwoven, interactive processes of both representation construction and information seeking"*. This argument is closely related to sensemaking research, where Russell et al. suggest that exploratory search systems need to support sensemaking activities (Russell, Furnas, Stefik, *et al.*, 2008; Russell, Pirolli, Furnas, *et al.*, 2009).

3.3. Methods Applied in this Thesis

This thesis utilises prototyping extensively in two separate strands. In the first strand, an information interface for just-in-time information discovery and exploration – the FMI prototype – is iteratively designed, developed and evaluated in a range of lab-based studies (see Preliminary Study 1, Preliminary Study 2, and Main Study 1). In the second strand, the evolved concept underlying the iterative design process from the first strand, has been applied to a secondary design process 'in the wild' which aimed to account for subjects real-world needs, and on developing a prototype that can be tested in an ecologically valid context (see Main Study 2).

3.3.1. Methodology Matrix Used in this Thesis

A combination of lab-based experiments and field studies accompanied the formative iterative design and implementation of two distinct prototype systems - the Focus-Metaphor prototypes (Prelim. Study 1, Prelim. Study 2, Main Study 1) and the KnowDis prototypes (Main Study 2). The research applied a range of quantitative and qualitative methods and evaluation/use contexts as outlined in the methodology matrix below (Figure 36).

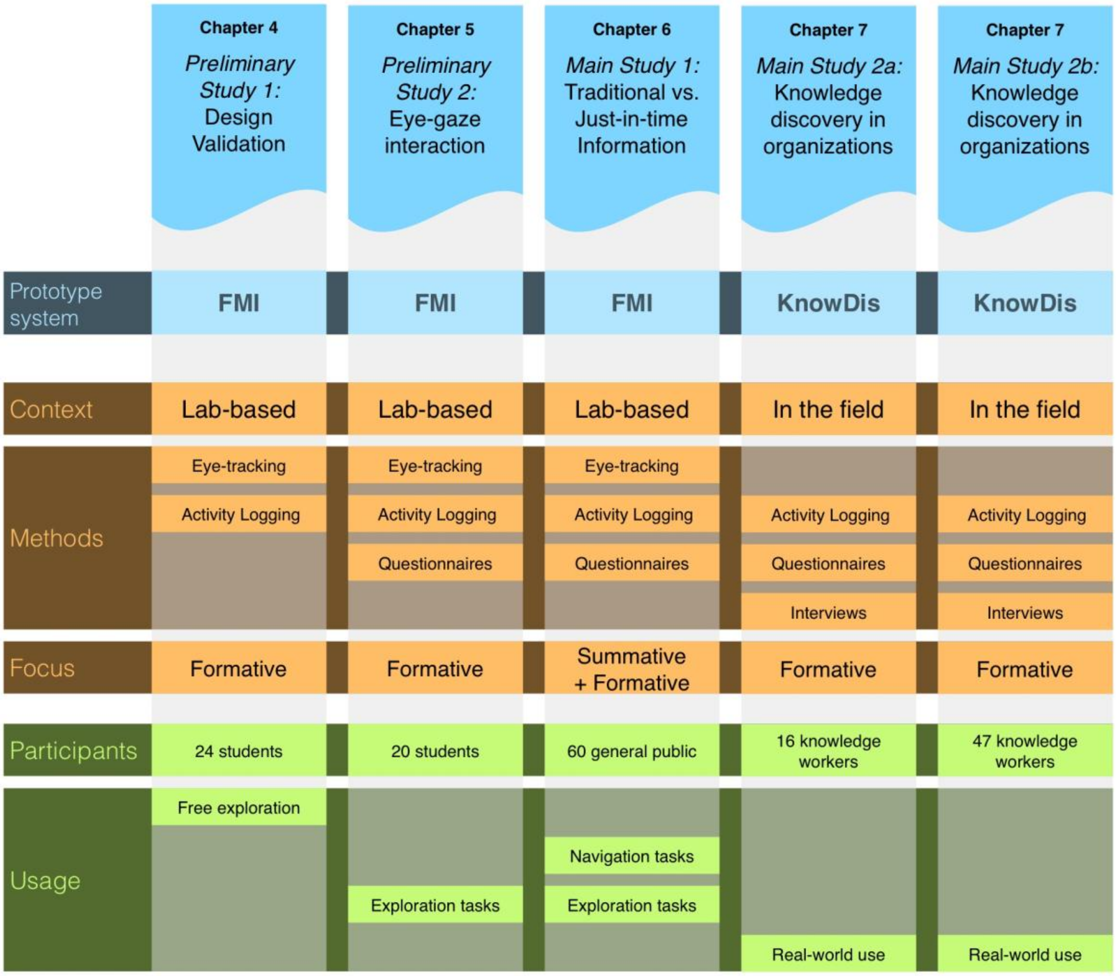


Figure 36: Methodology matrix for studies conducted in this research

3.3.2. Usage Contexts Employed in this Thesis

Evaluations of prototype systems employ a range of usage contexts that are more or less directed, depending on the desired level of ecological validity and type of data collection employed.

3.3.2.1. Goal-Directed Tasks

The majority of lab-based studies evaluates task performance as one of multiple facets. For such evaluations, measurements such as time taken, errors made, number of steps are utilised to support quantitative analyses.

Due to the interplay of top-down cognitive processes and bottom-up cognitive processes, human behaviour is intrinsically unique. Thus, performance related evaluations utilise simplified tasks that consist only of a few steps and are more predictable, more comparable, as a result.

3.3.2.2. Challenges of Ecological Validity

Tasks measuring performance in a lab-based environment are typically goal-directed. In the context of information retrieval, such tasks are commonly referred to as search tasks or navigation tasks.

Given that such tasks are somewhat artificial, it is desirable to increase ecological validity through means that can be controlled. For example, in this thesis, prototype evaluations conducted in the lab targeted participants that were generally interested in the topic covered by the navigation tasks used. Moreover, to help participants relate to a given task, they were embedded in small scenarios to increase intrinsic motivation - such as: *"A colleague told you, he just read an interesting article on the pros and cons of drinking coffee every day. Since you worry about your consumption, you want to have a look at this article yourself…"* (see Main Study 1).

3.3.2.3. Exploratory Tasks

Another approach to lab-based evaluations focuses on giving participants more freedom to explore a given prototype system. Such exploratory tasks are particularly useful for capturing more natural user behaviour (given the constraints of a lab), which provides more insights into individual differences of tested subjects than goal-directed tasks. Insights into more natural user behaviour are of particular interest, when testing novel interaction techniques or generally prototype systems that do not conform to 'typical' applications that subjects will be familiar with.

3.3.2.4. Challenges of Natural Exploration

However, utilising exploratory tasks to elicit natural behaviour may require the use of larger sample sizes, if quantitative aspects shall be analysed. Larger sample sizes enable comparisons of groups of subjects that demonstrate similar behaviour.

The following is an example of an exploratory task, embedded into a scenario (as explained above), and as used in this thesis: *"After your holidays, you have gained a few pounds. You are not happy and want to lose weight. You have seen from your friends that diets don't work. You believe that exercise is a much better way of getting back into shape. There is a lot of useful information on sport and fitness in these pages. Please* *have a look and find information on what sport is suitable for you and fits into your lifestyle"* (see Main Study 1).

3.3.2.5. Real-world Usage

The goal of field studies is typically to evaluate a prototype system with the highest level of ecological validity possible and to measure long-term effects of a prototype system (and the behavioural change it attempts to elicit).

As such, giving users defined tasks, whether clearly goal-directed or more exploratory, would take away from that validity and resemble a more experiment-centric remote testing that may be better suited to a lab-based study. Instead, field studies should employ real-world usage of a prototype system, where users make use of it as part of their normal daily activities.

More relevant qualitative data can be acquired by allowing subjects to use a prototype system within their natural (work) environment and over a longer period of time (e.g. weeks to months), compared to a typical laboratory session.

Enabling subjects to use a prototype system for a number of weeks rules out novelty effects (as typically encountered in lab-based studies) and thus helps generating more reliable data. In a longitudinal study, subjects are given time to familiarise with the prototype system and potentially integrate it into their daily work practice. Moreover, the longitudinal nature of such field studies allows for continuous feedback by subjects, in more or less structured form - ranging from simple feedback emails to more substantial subject interviews.

Another vital aspect of the potential necessity to conduct field studies is the availability of appropriate subjects. While many laboratory studies suffer from the use of subjects which are either under- or overqualified, unmotivated or 'professional participants', field studies allow targeting more appropriate subjects more easily.

Brown et al. (Brown & Reeves, 2011) provide a useful discussion of the methodological challenges of field studies. They argue that *"relationships between investigators and participants, relationships between participants themselves, and the nature of trial instructions"* may be sources of interference, which are typically not discussed. The authors argue that *"frequently, there is an attempt by reviewers to find the 'fatal flaw' in a methods section"* which as a result, causes methods sections to be written *"in a highly defensive manner"* (Brown & Reeves, 2011). The authors' assessment suggests that research with a sound methodology may benefit from a more open and minute discussion of methodological details and procedures, as to stand out among less well documented research.

3.3.2.6. Challenges in the Wild

Of course, an evaluation of real-world usage in the wild requires a level of prototype design and implementation that goes far beyond what is required for lab-based experiments. In the lab, usage conditions can be strictly controlled, the equipment used is specified, and certain behaviour of a prototype system may even be manually manipulated. In the field, a prototype system to be evaluated needs to be robust enough to withstand technical challenges that are sometimes hard to account for.

While a browser-based prototype may 'only' need to be compatible with a range of browsers, a prototype system for the desktop may have more complex requirements. For example, in this thesis the KnowDis prototypes were specifically developed to be used in field studies (see Main Study 2), and the extension/add-on for Microsoft Outlook that I developed for this study had to be compatible with different versions of MS Windows, different versions of MS Office, and work within the corporate network or through VPN for remote workers. As such the KnowDis prototype was evaluated by knowledge workers in Europe, America, Asia and Australia.

4. Preliminary Study 1: Dynamic versus Static Contextualisation of Information

The biggest issue for personalised websites is the applied method of visualisation. Scanning a grid or table-based layout spanning two, three or more columns is a task of high cognitive load, and users have developed efficient techniques to facilitate interaction with these websites. It is widely accepted that websites are mainly scanned rather than read and that effects like change blindness (see section 2.2.4) or banner blindness (Norman 1999) affect the perception of a website. This is a big thread to personalisation, since most websites make extensive use of navigational elements. The actual content gets more or less visually hidden behind the navigational framework of a website and processes of visual search (Duchowski 2003) become crucial. Often it takes users a substantial amount of time to understand the structure of a website, which spans visual layout, navigation and structure of content.

4.1. Introduction

This preliminary study tests the scale-ability of the Focus-Metaphor interface (FMI) for information spaces larger than the original static prototype, as reported in (Laqua & Brna, 2005). For clarity, the size of an information space will subsequently be defined by the number of 'articles' it contains.

In a scenario where more articles are available than can be displayed concurrently using the minimalistic visualization style of the FMI, the contextual elements will dynamically map onto a subset of available articles (henceforth referred to as 'dynamic FMI'). Within this study, an FMI consists of seven contextual navigation elements, a central content element, and no further traditional navigation menu or any other user interface elements (see Figure 37).

Figure 37: Focus-Metaphor Interface

For information spaces with more than seven articles, a subset of all available articles needs to be mapped onto the contextual navigation elements. To ensure that all articles are accessible via the contextual elements, with each interaction, a different set of articles is being mapped.

This approach results in an entirely dynamic user interface, without any static navigational elements as commonly found on the web.

By choosing a common topic for all articles within an information space, this study emulates tailoring the contextual display of information to the user and to a specific task context.

It is crucial to test the impact the dynamic allocation of content in the dynamic FMI condition has on user behaviour, as this change means that no static navigational elements exist within the entire user interface. This is a significant departure from how traditional web portals or web applications function. While even the original static FMI did not feature a traditional navigation menu, the fact that contextual navigation elements did never change their content meant that these elements do function more like traditional static web navigation. The potential problem investigated by this study is whether the omission of a familiar type of web navigation could cause user confusion, and prevent users from effectively using the dynamic FMI.

The study reports on the observed impact of dynamically changing contextual navigation elements (dynamic FMI condition) on user behaviour in comparison to the original static FMI condition used in (Laqua & Brna, 2005).

4.2. Research Question

***(RQ 1)** For (an) FMI interface, what are the effects of dynamically updating contextual elements during an information exploration task?*

4.3. Prototypes

The implemented prototypes model a simple information space based on textual and figurative elements. The prototypes have been implemented using Adobe Flash and ActionScript, using pre-compiled static articles.

Both prototypes utilize a Focus-Metaphor interface (FMI) to display content (see Figure 37). On the surface, the UI itself is identical for both prototypes.

The number of displayed contextual navigation elements (see Figure 38) is kept at a constant number of seven elements to maintain the minimalist nature of the UI. But the size of the underlying information space and thus number of available articles varies between both prototypes (see section 5.6 Independent Variables).

Figure 38. Example of a contextual navigation element

Contextual navigation elements of the FMI

Each contextual element provides snippet-like previews of the actual mapped article consisting of only text, only an image, or a combination of image and text. The contextual elements are arranged around one central content element (see Figure 39).

Content element of the FMI

The content element displays the currently selected article in detail and is displayed in the centre of the screen. As the user interacts with one of the contextual elements, the currently visible content element in the centre of the screen disappears, the contextual element selected changes its state to become a content element and then moves into the centre of the screen.

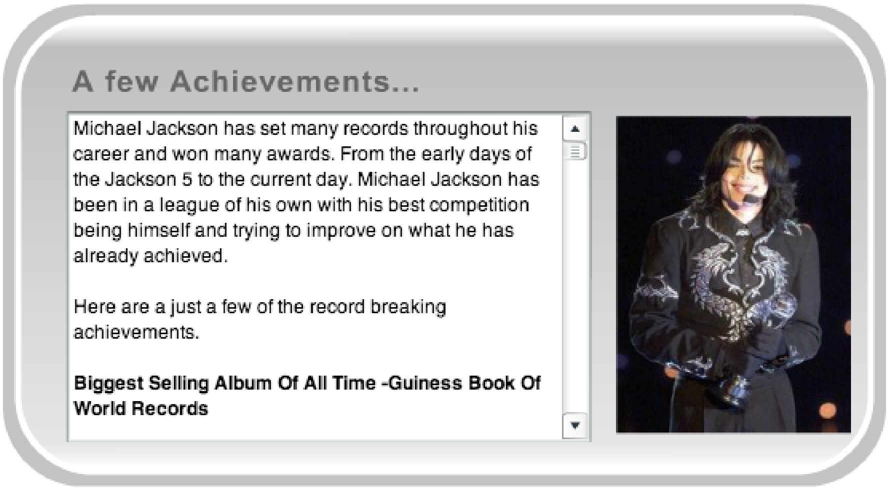


Figure 39. The central content element with sample content

No additional functionality was integrated, as the sole purpose of this prototype has been to observe participants' information retrieval behaviour - how they interact and navigate between the articles available in each prototype.

4.3.1. The Static FMI Prototype

The first version of the prototype ('static FMI') visualises an information space of exactly seven articles, mapping each article persistently to the same contextual navigation element. This mapping does not change during the session of the experiment. This **static FMI** is used as control group for the study. The seven articles in this version represent a subset of the articles used in the dynamic FMI prototype such that they reflect a good mixture of article types (specifically with regards to text-to-image ratio).

4.3.1.1. Behaviour of the Static FMI

In the static FMI, when a contextual navigation element is selected, two elements change. First, the article currently displayed within the content element in the centre of the screen is switched back to its representation as a contextual element and in its original position as part of the circular arrangement. Second, the selected contextual element changes its state to become the new content element (see Figure 39), displaying the article chosen by the user in detail. In addition, the new content element moves to the centre of the screen. In the initial state, upon first loading the static FMI, no content element is visible, as the user has made no selection yet.

4.3.2. The Dynamic FMI Prototype

The second version of the prototype (‘dynamic FMI’) visualizes an information space of 35 articles by dynamically mapping a random set of seven of these 35 articles onto the seven available contextual elements. The mapping of these randomly chosen seven articles is transient and only persists until the user selects one of those seven contextual elements. With each selection, a new set of seven randomly chosen articles is mapped onto the contextual elements. This version will be referred to as **dynamic FMI** and is used as treatment group (see section 5.6). A randomisation function was used to select the articles to be dynamically mapped to and displayed by the contextual elements. This approach allowed measuring the greatest possible impact in terms of an ‘unpredictable’ dynamic user interface, and how it compared to the static version of the Focus-Metaphor interface as used in (Laqua & Brna, 2005).

4.3.2.1. Behaviour of the Dynamic FMI

In the dynamic FMI, when a contextual navigation element is selected, all elements change. First, the article currently displayed within the content element in the centre of the screen is hidden. Second, the selected contextual element changes its state to become the new content element (see Figure 39), displaying the article chosen by the user in detail. In addition, the new content element moves to the centre of the screen. Third, the other contextual elements are mapped to a new random set of articles. In the initial state, upon first loading the static FMI, no content element is visible, as the user has made no selection yet.

4.3.3. Content used for Both Prototypes

The celebrity Michael Jackson was chosen as informational domain to ensure familiarity across participants, and because a lot of content was readily available. The prototype consists of 35 articles, covering various topics of Michael Jackson’s career.

4.4. Method

4.4.1. Experimental Design

This study was conducted in the lab measure gaze and interaction behaviour of participants when using two different versions of a Focus-Metaphor Interface (FMI). The same visual layout has been used during the experiment and the amount of accessible information has been altered through the randomly selected display of content in the dynamic version (experimental group). By comparing an interface that provides randomly personalised content (dynamic version – X[1]) with one that keeps all information static and thus makes the interface predictable (static version – X[2]) the main objective has been to *measure differences in visual attention* (behaviour Y) on navigation and content sections of the interface (through the measure of gaze time).

X[1]: Dynamic experiment version (Focus-Metaphor): animated (moving into position) content element in the centre of the screen; not animated (fixed position) contextual navigation elements arranged in a circular manner around the content element; 35 articles in total randomly loaded as contextual navigation elements.

X[2]: Static control version (Focus-Metaphor): animated (moving into position) content element in the centre of the screen; not animated (fixed position) contextual navigation elements arranged in a circular manner around the content element; 7 articles in total with explicit allocation to the contextual navigation elements.

The study used a between-subjects design. Participants were randomly assigned to either the experimental or the control group. The experimental group used the dynamic FMI henceforth referred to as X[1] and the control group used the static version, henceforth referred to as X[2].

4.4.2. Participants

The study was completed by 24 participants (16m, 8f). All participants were university students (avg. age 20.7 years) from a variety of ethnic backgrounds. The most common first language spoken was English (14 participants), followed by Gujarati (5 participants). To allow fair comparisons, none of the participants did have prior knowledge or expectations of the specific prototypes used or the Focus-Metaphor in general. However, in preparation for the experiment, participants were given a brief warm-up session to familiarise with the way the prototypes worked and to rule out effects that are due to the novelty of Focus-Metaphor interfaces.

4.4.3. Independent Variable

UI type

Two versions of a Focus-Metaphor interface have been tested. A static FMI (see section 4.3.1) has been used as baseline and compared against a dynamic FMI (see section 4.3.2).

4.4.4. Dependent Variables

Eye tracking was employed to analyse differences in users' attention through the measure of gaze time. Navigation and interaction behaviour with both versions was measured via mouse clicks. In order to compensate for variations in experiment session duration, all raw data has been normalised for all analyses.

4.4.5. Scenario

Participants were free to explore and interact with the prototype. The single scenario-like task given to participants was to find out new and interesting information on Michael Jackson, aiming to simulate a real-world information exploration scenario.

4.4.6. Procedure

Prior to each session, participants were given a scenario form, which briefly described the experimental procedure. Participants were also given a brief demographic questionnaire prior to the experiment. Eye-gaze data was collected using a fairly old eye tracking system by LC Technologies. Although this eye-tracker did not require any head mounted parts, this equipment was very sensitive to any head movements. Thus a chin-rest was used to minimise head movements and to increase validity of the data. Participants were asked not to move their head and to look at the screen at all times after calibration until the end of the experiment session.

Once participants were calibrated, the experimenter reminded them to look for some information about Michael Jackson that they would find interesting. Participants were told that they are free to stop the experiment after a reasonable amount of time, if they felt that they would like to stop because they have read enough (e.g. not after 10 seconds). If participants did not initiate the termination of the experiment themselves, the experimenter stopped the session after approximately five minutes.

After completion of the experiment session, which lasted approximately 30 minutes overall, participants were compensated £5 for their time.

4.5. Results

As expected, participants assigned to the session utilising the static version X[2] more frequently terminated the session early than those participants assigned to the session utilising the dynamic version X[1]. While the measured average session duration for the dynamic version was approx. 4 1/2 minutes (X[1]_time ~ 274 seconds), the average session duration for the static version was closer to 4 minutes (X[2]_time ~ 253 seconds). The resulting difference of average session duration between the two versions was 21 seconds, or less than 10%.

The analysis of the gaze data (through heatmap and attention analyses) shows that participants exhibited a different strategy in scanning the interface, between versions X[1] and X[2].

4.5.1. Heatmap Analysis

Due to limitations in the eye tracking software used, no traditional heatmap analysis could be performed. Instead, a custom heatmap analysis was devised, and utilised to highlight differences in scanning behaviour between the two versions. This heatmap analysis is based on a screen grid of 768 cells (32 columns, 24 rows), with attention distributions being collected on a cell-by-cell basis.

Instead of generating traditional heatmaps for versions X[1] and X[2], which would require visual inspection to identify differences, the screen grid data was utilised to generate 'difference heatmaps[3]'. This technique combines two traditional heatmaps by subtracting average attention for each cell in one version from average attention on the

[3] The term difference heatmap is introduced by the author. There is no awareness of existing usage of this term in related literature. The difference heatmap technique has been proposed for integration into Tobii's analysis software at the European Tobii User Meeting and might well appear in future versions.

same cells in the other version. The resulting heatmap contains both attention surpluses and attention deficits of one version over the other. This data is then separated into two distinct difference heatmaps that separately show attentional surpluses of each version over the over version (see Figure 40). In contrast to a traditional heatmap, which simply visualises the accumulated gaze interest on the screen, the difference heatmap integrates the attention data from both versions, thus clearly highlighting differences in attention distribution (see Figure 40 and Figure 41).

The left heatmap in Figure 40 shows the attention surplus of X[1] (static version) over X[2] (dynamic version), and the right heatmap shows the attention surplus of X[2] over X[1]. This analysis visually confirms that participants of the control group using the static version spent more time on the central content element (and thus on the content) than participants of the experiment group using the dynamic version. In contrast, the experiment group spent more time on the contextual navigation elements (and thus on the navigation) than the control group.

Figure 40: Heatmaps visualizing the attention losses on content (left) and attention gains on navigation (right) of the dynamic version X[1] versus the static version X[2].

Figure 41: The large heatmap (top) integrates gains (green) and losses (red) in visual attention (of version X[1] over version X[2]) in a single difference heatmap.

4.5.2. Attention Analysis

Areas of interest have been defined to compare attention distributions for content and contextual navigation areas between the dynamic version X[1] and the static version X[2]. This analysis provides insight into the time users actually spent reading information in the *content area* (X_read), and the time they spent navigating via the *context area* (X_navi), The context area describes the circular band containing all contextual navigation elements, and surrounding the central content element. The left part of Figure 42 shows the amount of time participants spent reading information in X[1] and X[2]. There is a significant difference between **X[1]_read = 72.1%** and **X[2]_read = 83.9%** with standard errors for the means of $\boldsymbol{\sigma_{Y1\_cont} = 3.3}$ and $\boldsymbol{\sigma_{Y2\_cont} = 1.1}$.

This result denotes a decrease of time spent on content by 11.8 percentage points when dynamically shuffled content is displayed instead of static content. The right part of Figure 42 shows the amount of time participants spent navigating in X[1] and X[2]. There is a significant difference between **X[1]_navi = 13.7%** and **X[2]_navi = 10.2%**. The standard errors for the means are $\boldsymbol{\sigma_{Y1\_navi} = 1.6}$ and $\boldsymbol{\sigma_{Y2\_navi} = 0.8}$.

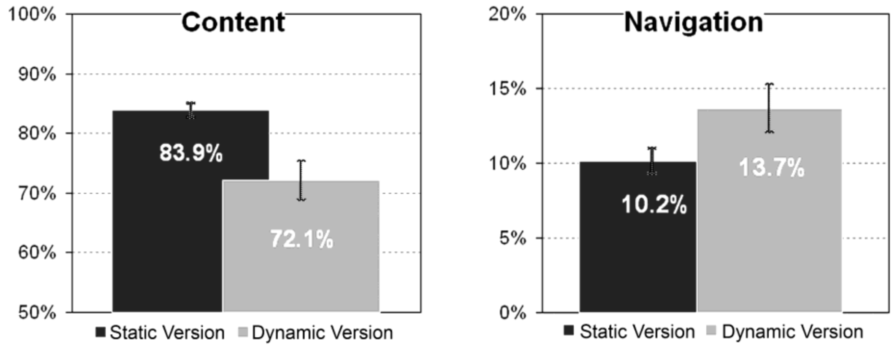


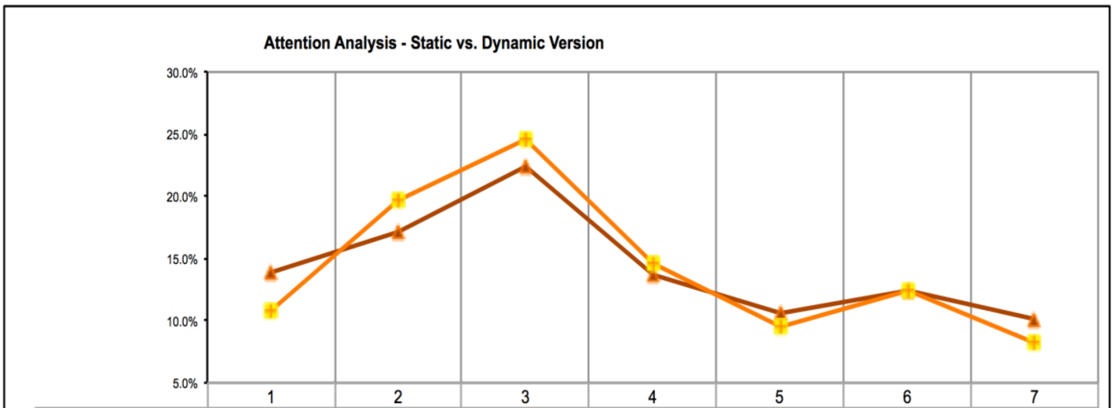
Figure 42: Comparison of user attention on content (left) and on navigation (right)

The results confirm that participants spent more time attending to the contextual navigation elements (Y[1]_navi) in the dynamic version X[1] and spent less attention on the central content element (Y[1]_read) as a result (Table 4). This can be explained by the randomly changing contextual navigation elements, which may have 'confused' participants to some degree. But more importantly, and possibly more likely, the changing elements simply meant more information to explore and thus could also just point to higher engagement with the information exploration activity.

Table 4: Comparison of user attention in X[1] and X[2] on content and on navigation

	X[2] (static version)			**X[1] (dynamic version)**		
	in %	σ_M	Confidence interval	in %	σ_M	Confidence interval
Content	83.9	1.1	$81.7 \leq \mu \leq 86.1$	72.1	3.3	$65.7 \leq \mu \leq 78.5$
Navigation	10.2	0.8	$8.5 \leq \mu \leq 11.8$	13.7	1.6	$10.5 \leq \mu \leq 16.8$

A more detailed breakdown of the attention data for X[1] and X[2] shows that both versions exhibit similar patterns of attention distribution (see Figure 43). In both the static and the dynamic version, visual attention is highest on element 3, which is in the top left corner, followed by element two, which is the left-most element. This analysis of the attention data per element position also reveals that overall attention distributions across the 7 contextual navigation elements are very similar for both the dynamic version X[1] and the static version X[2].

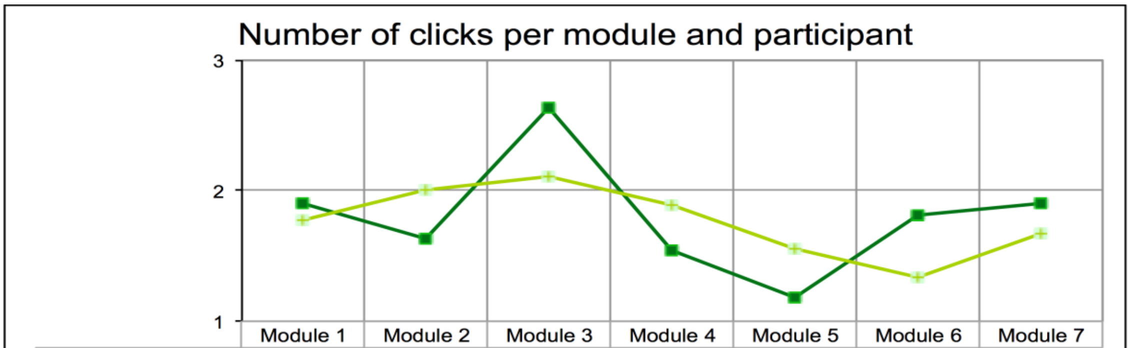


	1	2	3	4	5	6	7
dynamic attention	13.9%	17.1%	22.4%	13.6%	10.6%	12.3%	10.1%
static attention	10.9%	19.7%	24.7%	14.6%	9.6%	12.4%	8.2%

Figure 43: Attention distribution across contextual navigation elements in X[1] and X[2]

4.5.3. Interaction Analysis

The analysis of interaction data (mouse clicks) reveals a preference for certain elements within both interfaces (Figure 44). Possibly to be expected, element 3 in the top left corner was the most selected element. This trend is stronger for participants using the dynamic version X[1] than for participants using the static version X[2]. A somewhat unexpected observation is the higher number of interactions with elements 6 and 7 in the dynamic version X[1]. Those elements are positioned in the bottom right corner of the interface, and common knowledge on typical patterns on attention distributions on web sites or reading behaviour would suggest those areas to receive the least attention. The attention analysis somewhat supported that original expectation, as both elements 5 to 7 show the lowest visual attention across all elements.



	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6	Module 7
dynamic version	1.91	1.64	2.64	1.55	1.18	1.82	1.91
static version	1.78	2.00	2.11	1.89	1.56	1.33	1.67

Figure 44: Interaction (mouse click) distribution across contextual navigation elements in X[1] and X[2]

4.6. Discussion

Preliminary Study 1 was conducted to investigate *the effects of dynamically updating contextual elements during an information exploration task (RQ 1)* for a Focus-Metaphor Interface (FMI). The study demonstrated that it is possible to successfully use the FMI concept on larger information spaces by mapping articles dynamically to the available contextual navigation elements without detrimental effects to the user experience that might have made the dynamic version of the FMI unusable.

The reason, why subjects who interacted with the dynamic version spent 10% longer reading the provided material than those with the static version is very likely due to the simple fact, that the dynamic version offered a much larger choice of articles. As participants were given time to freely explore all information contained in the assigned version, the overall time on task does provide little insight into actual qualitative differences between both types of experiences – of the dynamic and the static FMI version. While the interaction distribution analysis points to some differences in how both versions were used, this can be explained by the selection patterns of participants. Whereas element 3 clearly was the most popular choice across participants and interface type, in the static version, its content did not change, so the value in selecting it again was low. However in the dynamic version, every selection of an article randomly populated the contextual navigation elements with new articles. As a result, participants learnt to simply keep clicking on the same element to get to more information.

Overall, the various analyses on the interaction patterns observed, and attention distributions measured across both static and dynamic context items suggest that the effects of dynamically updating contextual elements of an FMI do not significantly alter user behaviour for information exploration tasks. These findings suggest that it is feasible to support information exploration activities through the pro-active display of contextual navigation elements in a just-in-time information interface that maximizes the amount of information relevant to an information goal, and at the same time minimizes the amount of information extraneous to an information goal.

The next step in the further development of the FMI concept will need to be evaluations that more closely emulate real-world usage scenarios. This will require the utilization of larger information spaces with broader topical scope. Main Study 1 represents significant follow-up research on the study reported in this chapter, utilizing a much more advanced prototype, the testing of a range of scenario-like tasks and comparison to a more traditional web interface (see Chapter 6).

5. Preliminary Study 2: Gaze Controlled Content Spaces

5.1. Introduction

For several decades people have used keyboards and mice as primary input devices for personal computers. With continuous advances in technology, novel forms of interaction are emerging, most recently touch and gesture-controlled user interfaces. A few years ago, people were using pens or a single finger to control tablet PCs, PDAs, public terminals or early smart phones. Today, multi-touch interfaces like the Iphone, Ipad or a plethora of Android devices are used by hundreds of millions of people every day. With advances in eye tracking technology and particular research into eye gaze interaction, such as (Sibert & Jacob, 2000; Hansen, Johansen, Hansen, *et al.*, 2003; Miniotas, Špakov & MacKenzie, 2004a; 2004b; Junker & Hansen, 2006; Kumar, Paepcke & Winograd, 2007), interacting with devices using eye gaze may be the next big shift.

Already, manufacturers like Tobii and SMI are starting to push gaze interaction by reducing costs or by providing platforms and components to enable the integration of gaze-tracking technology into everyday devices such as laptops, TVs or even tablets. Recently, some early mass-market prototypes of consumer products that embed gaze interaction technology started to appear. Tobii is working with Lenovo on a laptop with integrated gaze-tracking technology (http://www.tobii.com/group/news-and-events/tobii-in-media/tobii-presents-eye-controlled-laptop/) and with Fujitsu on a similarly equipped tablet device (http://www.engadget.com/2012/09/19/tobii-fujitsu-and-ntt-docomo-partner-on-eye tracking-ibeam-tablet/).

But as eye-tracking hardware starts to become more ubiquitous, we will need to find ways to interact with information effectively using eye gaze. Much of the previous research on gaze interaction has focused on *assistive technologies for citizens with motor impairments (http://www.cogain.org/)*. For this audience, gaze interaction may be the only available input modality. Due to the invaluable benefit gaze interaction of any type offered to this community, not much attention has been given to the general user experience of such systems. The design of the GazeSpace prototype has been specifically aimed at able-bodied audiences, who will have much higher expectations on the overall quality of the interaction and general usability of such systems. Touch-based devices existed as a niche until the concept of multi-touch and improved hardware made the experience appealing to a mass audience. The same pattern is like to occur again with gaze interaction, unless the quality of the user experience is addressed. It should be noted that those shifts towards mainstream adoption as pushed by Tobii and other companies have started as recently as 2011/2012 - the research reported in this chapter was conducted in 2007.

5.1.1. Motivation

As technology advances, and input modalities change, a new paradigm for human-information interaction should be able to cope with those advances. It thus seems vital to evaluate the suitability of the concept of just-in-time information interfaces for other input modalities. The effective development of a paradigm for just-in-time information interfaces that is meant to tackle the problem of information overload implies that ideally, a range of input modalities are considered.

Back in the early stages of this research, it was already apparent that touch and eye-gaze may be input modalities that will grow in significance in the future. While at the time of conception of this thesis the use of a keyboard and mouse was the prevalent input modality, this has since started to shift towards devices using multi-touch surfaces. However until the introduction of the iPad in 2010, touch-based devices were somewhat clunky and ineffective to use. Given that eye tracking has been one of the main methods used in the various lab-based studies, it was obvious that the investigation of a just-in-time information interface variant, which supports eye-gaze interaction, would be promising.

5.2. Eye-gaze Interaction

Eye gaze interaction is commonly regarded as a potential complement, if not a replacement for traditional input techniques. However, when discussing input techniques for computing systems it is crucial to distinguish two main steps in the interaction process. As today’s user interfaces (UI) mostly apply a desktop metaphor, users are required to point (e.g.: move a mouse) and select (e.g.: press a mouse button). It is important to consider this distinction when analysing potential eye gaze interaction techniques. Many existing eye gaze systems demonstrate satisfying results for gaze-based pointing (or gaze-pointing). However, this only fulfils half of the requirement for the described *point and select* interactions. Successfully implementing the selection part of the interaction is more challenging than the pointing part. Current eye gaze systems commonly make use of static dwell-times to achieve this second part of the interaction process (Sibert & Jacob, 2000; Hansen, Johansen, Hansen, *et al.*, 2003; Miniotas, Špakov & MacKenzie, 2004a; Miniotas, Špakov, Tugoy, *et al.*, 2005).

Most existing eye gaze systems are designed for medical contexts (e.g.: Tobii P10), where they enhance quality of life for people with disabilities, who cannot use traditional input techniques. The most common applications for patients are eye-typing to communicate with their surroundings and means to control their environment (e.g.: light switches, or motor controls for wheel chair). With increasing accuracy (resolution of gaze pointing), flexibility (freedom of head movements) and decreasing costs (Babcock, Li & Winfield, 2012), applications *"will soon be practical for able-bodied users"* (Kumar, Paepcke & Winograd, 2007). But as able-bodied audiences have much higher expectations for quality of interaction and general usability, challenges to beat traditional input methods and user interfaces arise. This partly explains why much of existing research focuses on rather abstract tasks when evaluating prototypes using dwell-time selection (Sibert & Jacob, 2000; Majaranta, Aula & Räihä, 2004; Hansen, Johansen, Hansen, *et al.*, 2003).

Testing simple eye-gaze selection tasks ensures that limited cognitive effort is required. Although one can argue that simple selection tasks help modelling future scenarios, e.g. of selecting menu items, the abstract nature of these experiments often excludes the cognitive component: people need to look at a number of elements, make a choice, and then select the appropriate element. Experiments testing selection based on colour or single letters (Sibert & Jacob, 2000) minimise cognitive load and thus simplify task complexity.

Research on more realistic use-cases, involving more complex tasks, commonly combines gaze-pointing with alternative means for selection, such as hotkeys (Kumar, Paepcke & Winograd, 2007), speech (Miniotas, Špakov, Tugoy, *et al.*, 2005) or EMG clicking (Junker & Hansen, 2006).

Kumar's EyePoint system (Kumar, Paepcke & Winograd, 2007) enables users to browse the World Wide Web by replacing mouse interactions with a combination of gaze-pointing and hotkeys for selection (using a keyboard). EyePoint proposes a "look-press-look-release action" to cope with accuracy limitations of current eye trackers when used with standard user interfaces. This incorporates (a) looking at an area of interest, (b) zooming into this area by clicking a hotkey, (c) selecting the desired element within this scaled-up area by looking at it and (d) releasing the hotkey to confirm selection of the chosen element. The main benefit of this approach is its compatibility with standard desktop interfaces. However, this solution has a number of limitations: (1) It requires separate keys to control the interaction and (2) it complicates the interaction process by approx. factor 2: point-click (mouse) vs. look-press-look-release (EyePoint). Compared with traditional mouse-based interaction, EyePoint increased task completion time and resulted in much higher error rates (Kumar, Paepcke & Winograd, 2007).

5.3. Research Question

Existing research into eye-gaze interaction typically accepts the concept of static dwell times as a given but ineffective method for the use with realistic information interfaces. While these previous studies thus augment static dwell times with additional controls for more effective interaction, this study has taken a different approach:

(RQ 2) *For (an) FMI interface, what are the effects of selection by gaze based on dynamic dwell times on user-preference? (RQ 2.1) And specifically, are there any differences in user preference between the two implementations of dynamic dwell times – 'static interest accumulation' and 'dynamic interest decay'?*

This study investigates the effect of using dynamic activation times on user preference, for gaze-based interactions with a focus-metaphor interface (FMI). Two variations of dynamic activation time algorithms have been implemented (see section 5.4.3) and evaluated. These algorithms – 'static interest accumulation' and 'dynamic interest decay' - have been tailored to how the user is scanning an interface, and allow reading text contained within a navigational element without accidentally activating it (the Midas-touch problem). Such algorithms should enable the design of gaze interaction interfaces that enable the exploration of information spaces without resorting to additional input mechanisms.

5.4. The GazeSpace Prototype

GazeSpace has been developed to provide a simple eye gaze interface that offers an appealing alternative to using the mouse when browsing content spaces (e.g.: to read blogs and news, look at pictures or video clips).

5.4.1. The User Interface

GazeSpace combines real-time gaze tracking with a just-in-time information interface prototype. Building on a previous work on Focus-Metaphor interfaces (Laqua & Brna, 2005) and on the work reported in chapter X, the screen layout allows for the same seamless exploration of information spaces. The novel layout also suggests being beneficial to overcome the Midas Touch problem, commonly found in eye gaze interaction systems. The main information area (content element) in the middle of the screen is surrounded by contextual navigation elements (see Figure 45). When selecting a contextual element, its state changes: It enlarges into a content element and moves to the centre of the screen replacing the previous element. This previously central element switches back to context state and moves to the periphery of the screen.

Figure 45: The GazeSpace interface

5.4.2. Layered Architecture

A Tobii X50 eye tracking system and the Tobii Software Development Kit (SDK) have been used to develop the GazeSpace prototype. The SDK enables access to real-time gaze data through the Tobii Eye Tracker Components API (TetComp). GazeSpace is a 2-layer application consisting of a system layer and a visual layer. The system layer has been developed using Microsoft VisualBasic (VB) which enables real-time access to gaze data (50Hz results in 50 gaze points per second) through TetComp. The visual layer has been developed using Adobe Flash (formerly Macromedia). The Flash user interface (UI) is embedded within the VB application. Relevant information about a user's gaze and the state of the UI is communicated between VB and Flash (using Actionscript). The layered approach has been chosen for two reasons: (1) Flash alone is not capable of accessing system API's or DLL's, which are required to receive real-time gaze data, and (2) VisualBasic does not provide the flexibility for fast prototyping, which would enable a similar user experience as Flash does.

5.4.3. Algorithms

The GazeSpace prototype aims to enable people to explore information spaces entirely by using eye gaze. To cope with the increased task complexity, instead of using static dwell-time for activation, GazeSpace integrates a more flexible gaze-interest-threshold. This approach comes with another benefit: it is more robust (if the eye tracker looses the user's gaze, the last working state of the interface can be "frozen"). Visual feedback ensures that the user is aware of the problem (this will be explained in detail in the feedback section below). After repositioning, she can return to the last working state. Instead of using fixations, the implemented algorithms work on the actual raw data, delivered through TetComp. Two variations of this technique have been implemented:

The **static interest accumulation algorithm (SIA)** collects gaze over each target area (navigational elements) until a predefined threshold is reached for the first element. When the user moves away from one target area before threshold is reached, the interest counter remains at this level until (A) the user comes back and new interest is added, or (B) another target area reaches threshold (gets activated). In case B, the interest counter is set to 0. The logic behind this algorithm aims to follow natural decision making processes, where a user looks at certain navigational choices before deciding which element to activate. After getting a first impression about relevant elements (by gazing at them), the user's actual decision to activate one of the elements should be faster. Accordingly, by collecting interest (gaze) along this decision making process for each element, reaching the interest-threshold will be faster.

The **dynamic interest decay algorithm (DID)** works similar to the first algorithm (SIA). However, instead of simply stopping the interest counter for an element when a user looks elsewhere, an additional decay function is used to decrease the collected interest for each element based on a predefined timing value (e.g.: accumulated interest is reduced by 50% every 50 collected gaze points outside the target area). This approach follows the logic, that when a user is not looking at an element for a longer amount of time (after first focusing on it), she "looses interest". After a few seconds, information about the previous element will have left short-term memory and thus more time is required again to make a decision (longer activation time).

5.4.4. Integrated User Interface Feedback

Existing eye gaze systems provide a very basic appearance, with usability and aesthetics being subordinate to functionality. GazeSpace is designed to improve usability by providing continuous visual feedback to the user. Gaze-pointing and activation are facilitated through a coloured border for each navigational element, providing feedback through a dynamic gaze-over state (similar to mouse-over state). Being transparent in neutral state, colour intensity increases continuously whilst a user is looking at an element. When reaching interest-threshold (by collecting a user's gaze), colour intensity approaches 100%, signalling the user that the element is about to be selected (see Figure 46).

Figure 46: Contextual UI element in neutral state (left) and after collecting "attention" through gaze (right)

As all stationary (not head-mounted) eye tracking systems can only cope with moderate head movements, it cannot be guaranteed that the user's eye will be tracked the whole time. We believe that providing feedback, on whether the system is "active" or not will enhance overall usability of the eye gaze system. For this reason, GazeSpace integrates a coloured background providing visual feedback over the general tracking state of the system: light red – no gaze found, light green – gaze tracked (see Figure 45).

5.5. Method

5.5.1. Experimental Design

We conducted an experiment in the lab to measure system performance and user satisfaction with the GazeSpace prototype (see Figure 47). In order to test the algorithms using a between-subjects design, we created two prototype versions with separate content spaces:

C1: Seven UI elements providing information on ePassports (see Figure 45), and
C2: Seven UI elements providing information on the government's ID card scheme.

Both content spaces were designed to create a similar appearance and offer similar amounts of pictorial and textual information. Pairing of algorithm and content space has been counterbalanced (session combinations: A1/C1, A1/C2, A2/C1, A2/C1). Each participant conducted two sessions (e.g.: A1/C2 and A2/C1), ordering of sessions has also been counterbalanced.

Figure 47: Experimental setup for the GazeSpace experiment

5.5.2. Participants

20 participants took part in the experiment. The average age was 24.2 (age range 18 – 50) and the sample was balanced for gender (10 male / 10 female). 16 participants were undergraduate or postgraduate students, two in part-time employment and another two in full-time employment. 19 of the participants reported their self-rated experience with computers as expert (9) or above average (10). Nine participants reported being aware of eye-tracking technology and five reported to have taken part in eye-tracking experiments before.

5.5.3. Independent Variable

Two different activation algorithms have been tested via two versions of the GazeSpace prototype. While visually, both versions of the prototype looked them same, the two algorithms for the activation of contextual navigation elements were tested – a static interest accumulation algorithm (SIA) and a dynamic interest decay algorithm (DID). The algorithms are discussed in detail in section 5.4.3.

5.5.4. Dependent Variables

User feedback on both prototype versions was collected after each session through questionnaires requesting participants' overall impressions of the system and feedback on the general usability of each prototype.

5.5.5. Tested tasks

Each session (e.g. A1/C1) lasted approx. 10 minutes during which participants were given information tasks relevant to the topic for that session by the experimenter (see

Table 5).

Table 5. Information Tasks

ePassport questions for sessions A1/C1 & A2/C1	**ID Card questions** for sessions A1/C2 & A2/C2
1) When will they be issues in the UK?	1) How much does an ID card for retired people cost?
2) Do these passports have an electronic chip in them?	2) Do under 16s need an ID card?
3) How long is the interview?	3) What biometric data will be stored on the card?
4) How many passport applicants were there last year?	4) How could ID cards help to fight terrorism?
5) In what year will the interviews be compulsory?	5) When will the ID cards be issued?
6) Can you tell me a benefit of introducing ePassports?	6) Will foreigners staying in the UK have to get an ID card?
7) Are the ePassports reader machines used at air ports 100% secure?	7) Will the personal information be stored in just one database?
8) What kind of personal information does the chip hold?	8) Will it be compulsory to carry an ID card?
9) How much does an ePassport cost?	9) What percentage of the population in France carries an ID card?
10) How long is the warranty of the microchip?	10) Is reducing identity theft a benefit of ID cards?
11) Are there any other countries that use ePassports?	11) Do the liberal democrats support ID cards?

5.5.6. Procedure

All participants were given an ethical guidelines form to read and sign in order to participate in the study. The form described the study, and informed them about their right to withdraw from the study at any point. It also informed participants that at the end of the study, they would be compensated £5 for their time. After calibration of the eye-tracking equipment, participants could familiarise themselves with the GazeSpace system using content different from the one used during the experiment sessions. During this phase, participants received a brief introduction to how the system works and what the visual feedback means. The aim was to facilitate task-driven exploration of the content spaces. When participants confirmed that they were familiar with how the system works, the first session commenced. Due to the nature of the experiment – participants were only able to interact with the system using their eye-gaze and did not have access to a mouse or keyboard – it was unrealistic to display the session tasks on screen and allow participants to walk-through the whole session by themselves using only eye-gaze. Realistically, this would have required custom-building an eye-gaze enabled experimental framework. Thus, tasks were read to participants by the experimenter. Feedback was requested on whether participants understood the task. The experimenter also offered to repeat the task, if desired.

After each session, participants had to fill in a questionnaire evaluating the currently tested system. At the end of both sessions, participants were also asked whether they found a difference in the interface (apart from content) and if yes, which of the sessions they preferred.

5.6. Results

The following analysis focuses on subjective quantitative and qualitative findings. With the aim to develop an appealing and simple to use eye gaze system, collecting meaningful user feedback on GazeSpace has been very important in this experiment.

5.6.1. User Satisfaction

To capture participants' overall reaction to GazeSpace, 6 questions from the *"Questionnaire for User Interface Satisfaction"* (QUIS) have been used (e.g.: "The system was: Frustrating – Satisfying"). In addition, aspects of ease of use, accuracy, etc. were evaluated using 25 tailored usability questions (largely based on the Computer System Usability Questionnaire – CSUQ). A 6-point Likert scale has been adopted throughout. Figure 48 shows the clustered and normalised results.

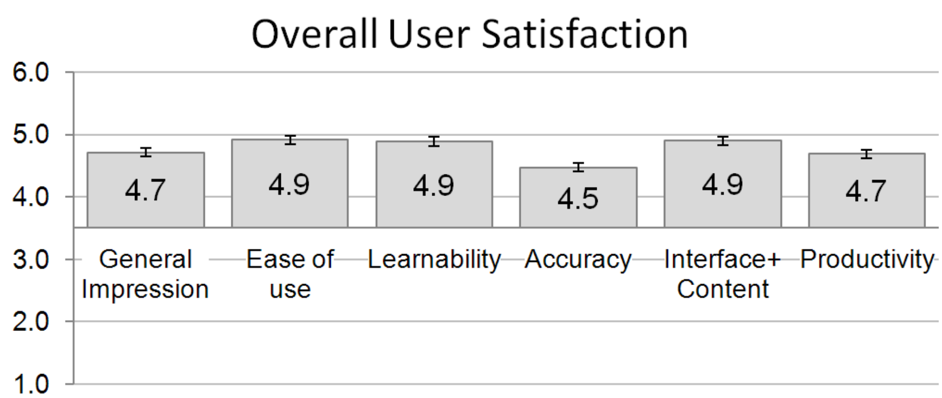


Figure 48: Questionnaire results

Participants consistently rated the GazeSpace prototype positive and particularly high for ease of use and learnability. This correlates with participants explicitly stating that GazeSpace is "very easy to use", "easy to understand", "self explanatory" and a "fun experience". However, there were of course also negative comments such as "speed of activation (not sufficient)", "eyes got tired after a while", "time to activate a bit long". The variety of comments shows that perception of the interaction experience using GazeSpace differed quite a lot between participants. The main aspect varying strongly was interaction speed. The user interface feedback (coloured border and background) created diverse responses: Some found it useful; others found it "annoying". The analysis of the related usability questions showed that participants rated the coloured border feedback significantly higher than the coloured background feedback ($t_{19} = 2.65$, $p < 0.01$). After conducting the experiment it also became clear that despite using a state-of-the art eye tracker, quality of calibrations and tracking ability varied strongly. While some participants were able to run a complete 10 minute session without hardly any interruptions of the gaze tracking, others had to cope with constant interruptions. This obviously biased participants' feedback. We therefore conducted a separate analysis of the questionnaire feedback using vision as independent variable (see next section).

5.6.2. Impact of Vision

Throughout all categories of user feedback, differences have been found when comparing feedback for participants with normal vision (7) and participants with corrected vision (13, using glasses or contact lenses). These differences have been significant for general impression ($t_{18} = 2.06$, $p < 0.027$), learnability ($t_{18} = 2.06$, $p < 0.027$), accuracy ($t_{18} = 2.50$, $p < 0.012$) and Interfaces & Content ($t_{18} = 2.16$, $p < 0.023$) (see Figure 49).

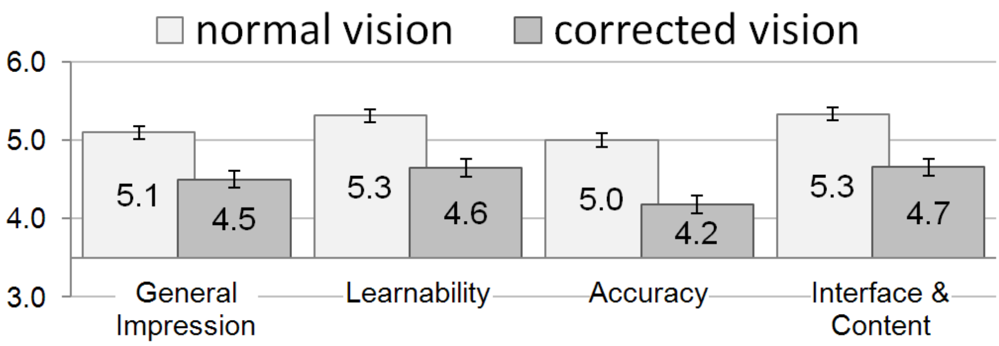


Figure 49: Impact of vision on user satisfaction

5.6.3. Impact of Algorithms

The comparison of the two tested algorithms (SIA and DID) revealed a higher preference for the static interest accumulation algorithm. While 9 participants preferred the SIA algorithm (45%), only 4 participants preferred the DID algorithm (20%). Interestingly, 7 participants (35%) did not notice any difference in the user interface of the two sessions (apart from content). Participants in favour of the SIA algorithm usually stated faster interactions as the reason.

When looking at the questionnaire feedback, using algorithms as independent variable, superior ratings for the static interest accumulation (SIA) algorithm for learnability and for interface & content have been found (see Figure 50). Although the SIA algorithm was preferred, participants surprisingly rated it less accurate. However, it needs to be noted that these differences are not significant.

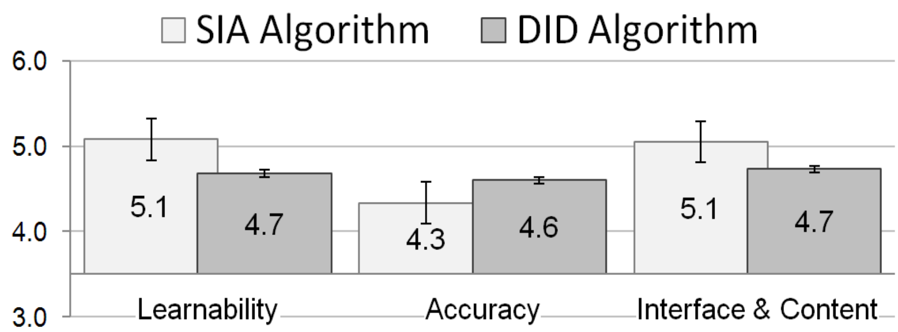


Figure 50: User satisfaction based on algorithm

5.7. Discussion

This study investigated *the effects of selection by gaze based on dynamic dwell times on user-preference (RQ 2).* The results demonstrate that it is possible to successfully use the FMI concept for gaze-based interactions with information when utilising dynamic dwell-time algorithms. Participants were able to use the prototype system well, interact with it and also found the experience overall quite satisfying.

5.7.1. General Prototype and Interaction Concept

Overall, the generally high user satisfaction with the GazeSpace system - independent of the version of the algorithm used – suggests that the use of dynamic activation times is a promising concept worth further exploration. Results suggest that the combination of a just-in-time information interface and gaze-based interaction has the potential to be a highly usable and enjoyable system for information discovery and exploration.

Results show that participants throughout liked the user interface and the interaction technique. Participants' comments such as *"aesthetically pleasing"*, *"user friendly"* and *"nice interface"* further suggest that the user experience will be crucial when targeting future eye gaze systems at able-bodied audiences. The design of the user interfaces should complement the novelty of the interaction technique and rich feedback on the system's state will further complement the user experience. Although feedback on the general tracking state (by change of background colour) seems to be quite useful, findings suggest a redesign into a more subtle and less disturbing feedback feature.

5.7.2. Differences between SIA and DID Algorithms

To cope with real-world information tasks, GazeSpace incorporated new algorithms using a more dynamic gaze-interest threshold instead of static dwell-times, which are typically used. Thus, this study also investigated *any differences in user preference between the two implementations of dynamic dwell times – 'static interest accumulation' and 'dynamic interest decay'? (RQ 2.1).* The comparison of the *static interest accumulation (SIA)* and the *dynamic interest decay (DID)* algorithms - both specifically designed to work with the just-in-time information paradigm – showed that users preferred the SIA algorithm overall (45% of votes vs. 20% of votes for DID). Participants also rated SIA less accurate than the DID algorithm. It needs to be noted, that the comparison of SIA and DID did not find significant differences in user preference between the two algorithms.

Nevertheless, the nature of the algorithms and some of the concrete feedback on user preferences suggests that the dynamic interest decay algorithm may be the more promising algorithm as it received a higher accuracy rating. This algorithm was intended to more effectively account for individual differences in user's exploration of the interface. The fact that participants more frequently preferred the static interest accumulation algorithm is also consistent with the fact that by the nature of the algorithm it would on average lead to faster activations (as attention decay is not taken into account). However, future adjustments to the DID algorithm could account for users' desire of overall faster dynamic activation times.

6. Main Study 1: Task Performance and User Experience for Traditional and Focus-metaphor Interfaces

6.1. Motivation

This study is a direct continuation of the research conducted in Preliminary Study 1 (Chapter 4), which demonstrated the general feasibility of dynamically displaying article previews in the contextual navigation elements of a Focus-Metaphor Interface (FMI). Consequently, this study has been designed around a further developed prototype, to evaluate the FMI comparatively against a traditional web-based interface, using a larger information space, and against common information tasks. Blog reading[4] has been chosen as a popular, yet specific activity that many people engage in regularly on the Web to allow for a more focussed, task-based evaluation.

The World Wide Web is increasingly about social interaction and collaboration. Blogging is a key activity in this Social Web enabling collective contributions of any type of information. Blogs have empowered millions of users to share their knowledge and experiences. But meaningful blogging experiences are as much about accessing information (reading) as they are about contributing information (writing).

The blogosphere (entirety of all blogs) faces the general problem of imbalance between ease of information contribution and meaningful information seeking. Millions of individual authors create millions of small and unique blog sites, and compete for attention in this messy space. Every contribution to this universal conversation - the actual content of a blog post - is wrapped into an individual visual design and a tailored structure of information through means of categories or tags. The dynamic nature of blogs quickly buries older content in archives or at best category lists reflecting the individual mind sets of their authors. In a sense, blogs are much like streams of individual thoughts. The main problem with information spaces as dynamic as the

[4] Due to the focus on text-centric information tasks in this research and by recommendation of the assessors of the thesis' interim reports, a typical blog interface was selected as proxy for traditional web sites that users encounter when conducting information search or information exploration tasks.

blogosphere is information discovery (Brooks & Montanez, 2006). Finding useful information can be hard and time-consuming often with a negative impact on the interaction experience.

Particularly in information spaces as dynamic as the Blogosphere, information seeking going beyond undirected browsing is problematic. Commonly, users' desire to explore a variety of information sources to feel confident in their judgement on complex problems (Bystrom & Jarvelin, 1995). Users also find it harder to formulate clearly what the problem is (Bystrom & Jarvelin, 1995) – something essential for effective search engine usage. With increasing task complexity, these two issues, (1) perceived quality of answers being bound to personal needs, and (2) people wanting to feel they make the choice, hold growing significance.

6.2. Research Question & Hypotheses

(RQ 3) For a Focus-Metaphor Interface (FMI), what are the effects of interaction-driven dynamic updating of contextual elements on task performance and user preference; and how does user interaction behaviour differ?

Two specific hypotheses have also been formulated to make specific predictions about the outcome of this study and to support the investigation of RQ3:

> (H1) For information search tasks, participants using the FMI will make fewer errors and complete tasks faster (than participants using the BlogUI).
>
> (H2) For both task types (information search and information exploration), participants using the FMI will rate the FMI as more usable in terms of usability criteria such as *ease of use*, *learnability*, and *productivity* (than participants using the BlogUI will rate the BlogUI).

In addition, an analysis will be conducted to explore differences in user interaction behaviour between FMI and BlogUI.

6.3. The FMI Prototype

In the blogging context, the FMI is mapped onto a blog space to provide convenient access to large amounts of blog articles (Figure 51). Contextual interface elements are arranged around the primary content element, which displays a selected article. The contextual elements function as navigation (activated through clicking) and provide previews onto the underlying content much like snippets on search engine result pages (SERP).

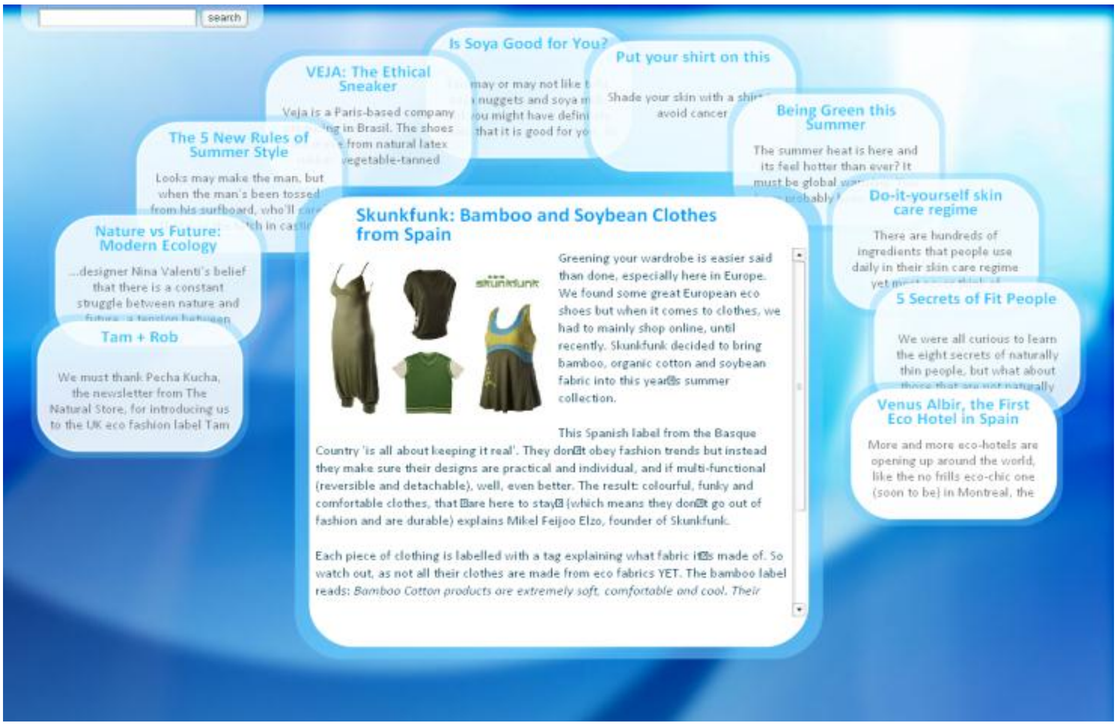


Figure 51: Just-in-time Contextual Blog Interface (FMI)

When selecting a contextual element, its state changes: It enlarges into a content element and moves to the centre of the screen, replacing the previous element. The display of contextual elements is dynamically adapted to the new primary content element using full-text similarity matching for the entire information space. The applied algorithm is inspired by a similar posts plugin for Wordpress (Marsh, 2006) and uses MySQL's full-text index and MATCH capabilities. This approach enables the dynamic adaptation of contextual elements to the currently displayed article. Selecting an article adjusts the context to the most similar / related articles. In a sense, the FMI "re-ranks" relevance of alternative articles and adjusts the contextual navigation accordingly; much like has been proposed for the display of search results by Teevan et al. (Teevan, Alvarado, Ackerman, *et al.*, 2004).

Providing only similar content to choose from in the navigation should facilitate orienteering and support focusing on a specific task. In order to enable efficient task switching, a search tool has been integrated that provides the same functionality as traditional blog search. Switching to a completely different topic using directed search mimics Teevan's concept of *teleporting* (Teevan, Alvarado, Ackerman, *et al.*, 2004).

6.3.1. Technical Details of the Prototype

The development of the high-level FMI prototype included the intricate combination of back-end and front-end development, using among other things:

1. a Flash/ActionScript front-end
2. a Flash Remoting implementation in PHP (similar to web services, and the only viable option for rich-internet applications back in 2005/2006 as it allowed client-server communication via large binary objects).
3. a MySQL database utilising full-text indexing and a range of complex dynamic SQL queries to enable real-time search and filtering.
4. the integration of an existing RSS archiving component into the server environment

6.4. Method

6.4.1. Experimental Design

We conducted a study to investigate information experiences in blog environments, and whether a Focus-Metaphor display can improve them. A corpus of approximately 160 blog articles on lifestyles was used to create a blog environment for this study. The content came from popular blogs and blog-like news-sites that cover health related topics such as exercise, fitness, workouts, healthy foods, diets, drinking, environmental issues and fashion. The original blogs and web sites have both male and female audiences. The standard blog interface used in this study (Figure 52) deploys a Wordpress installation with a 2-column layout. A "traditional" theme has been chosen to be representative for the majority of blogs in the blogosphere. A list of articles is displayed in the left column and a category list is provided in the right column. Each blog article can be accessed through 2+ categories. Search has also been integrated as alternative means for navigation.

Figure 52: Standard Blog Interface (BlogUI)

6.4.2. Participants

60 participants completed this study (31m/29f) from 18 to 67 years (median 28). Participants came from a range of educational backgrounds and had varying levels of computer experience (though all used computers at least occasionally). Since the tasks involved a substantial amount of reading, we recruited only native English or bilingual speakers and said study was for those with an interest in a healthy lifestyle. Payment was £8 for a 1hour session.

A *computer expertise* (CE) measure was calculated through a number of demographic questions on computer usage and literacy from a pre-questionnaire. Normalised value ranges from 0 (very novice) to 1 (very experienced). The average CE score for our participant sample is 0.4. For comparative analyses, participants with a score of <= 0.4 have been labeled “novice” and participants with a score of > 0.4 have been labeled “expert”.

6.4.2.1. Impact of Computer Expertise

Other studies in HCI that focus on eye tracking metrics (Cutrell & Guan, 2007; Halverson, 2003) commonly rely on sample sizes of 15 to 20 participants (or even less).

Their participants are often recruited from easily accessible and homogeneous participant groups, such as university students or employees of technology companies (e.g. Microsoft). In addition, these studies often focus entirely on eye tracking and potentially performance metrics, excluding any subjective evaluations.

With a sample size of 60 participants it seems both feasible and necessary to conduct some post-hoc analysis, considering the inconsistencies between task performance measures and user feedback as discussed in the previous sections. In a pre-experiment questionnaire, we collected a range of demographics, which allow us to calculate a computer expertise (CE) score. To validate the applicability of the CE score, we calculated some related statistics. There are 53% novice and 47% expert participants.

6.4.3. Independent Variables

6.4.3.1. UI Type

Two user interface types have been tested. A *traditional blog interface* (BlogUI, see Figure 52) has been used as baseline and compared against an experimental *contextual focus-metaphor interface* (FMI, see Figure 51).

6.4.3.2. Task Type

Information experience has been tested for two different task types: 1) A more flexible and user-centred condition, where *information exploration* has been tested through topical scenarios. 2) A more guided and goal-oriented condition, where *information search* has been tested through specific search tasks. Exploration tasks and search tasks represent contrasting scenarios. They have been tested in separate sessions.

6.4.4. Dependent Variables

Task performance measures have been taken for task errors and task completion times for information search tasks. A task error describes the case where for a given information search task, a participant is not able to find and identify the relevant information target. All information targets are identical for both UI types, but represented as a blog post in the BlogUI and as a content element in the FMI. Each occurance of a participant not completing a particular task is being counted as one task error. Task completion time is calculated in seconds from the start of information search task to its successful completion. Thus task completion times are only calculated for successfully completed tasks. Fixation counts, gaze time and average fixation durations have been measured as part of the eye tracking analysis. Navigation and interaction strategies have been measured through use of search usage and other types of interaction possible with the respective UI type. Subjective evaluation measures have been taken through detailed usability questionnaires.

6.4.5. Tested Scenarios (Task Types)

6.4.5.1. Information Exploration Tasks

The exploration tasks provided participants with a problem scenario and topic (see example below). Participants were then given time to *"explore information that ... provides useful insights concerning the given task"*. Participants were allowed to stop the task themselves (usually if they could not find more interesting information or felt they read enough) or were stopped after 5 minutes to keep the overall experiment time in a reasonable timeframe. After completing a task, participants were asked to briefly reflect on the information found (e.g. "what was useful, or not?", "which article was most interesting?", etc.). The aim was to foster a deeper (and more realistic) involvement in the given tasks (and stimulating intrinsic motivation).

Example of Information Exploration Task: *"After your holidays, you have gained a few pounds. You are not happy and want to lose weight. You have seen from your friends that diets don't work. You believe that exercise is a much better way of getting back into shape. There is a lot of useful information on sport and fitness in these pages. Please have a look and find information on what sport is suitable for you and fits into your lifestyle."*

6.4.5.2. Information Search Tasks

Search tasks provided participants with a specific scenario describing a particular article to be found (see example below). Participants were free to choose and switch between the integrated search and other means of interaction. If a participant was unable to find the target article, she was allowed to stop the current task and proceed with the next task. In addition to 6 standard search tasks, we included an additional difficult task:

Title and image of the target article did not provide an obvious link to the problem statement in task 3 (difficult task).

Example of Information Search Task: *"A colleague told you, he just read an interesting article on the pros and cons of drinking coffee every day. Since you worry about your consumption, you want to have a look at this article yourself…".*

A detailed list of all information exploration and information search tasks used can be found in Appendix B (see sections 10.5 and 10.6).

6.4.6. Procedure

The study was conducted in two parts - essentially as two between groups studies (see Figure 53 for details). The first study tested the information exploration tasks, and will henceforth be referred to as information exploration study. The second study tested the information search tasks, and will henceforth be referred to as information search study.

Results of a small pilot study showed that switching between types of tasks is prone to errors. Particularly the more complex procedure of the information exploration tasks required careful explanation to the participants. It was thus deemed that the best order to be running both studies was to let participants conduct the information exploration study first and the information search study second (as the search tasks are more self-explanatory).

Both studies were conducted using the same 60 participants. The participants were randomly assigned to one of two groups of 30 participants each. Group one conducted the FMI condition of the information exploration study and then subsequently the BlogUI condition of the information search study. Group two conducted the BlogUI condition of the information exploration study, and then subsequently the FMI condition of the information search study (see Figure 53 for illustration). The average CE score for participants in group one is 0.41 vs. 0.38 for participants in group two (see section 6.4.2).

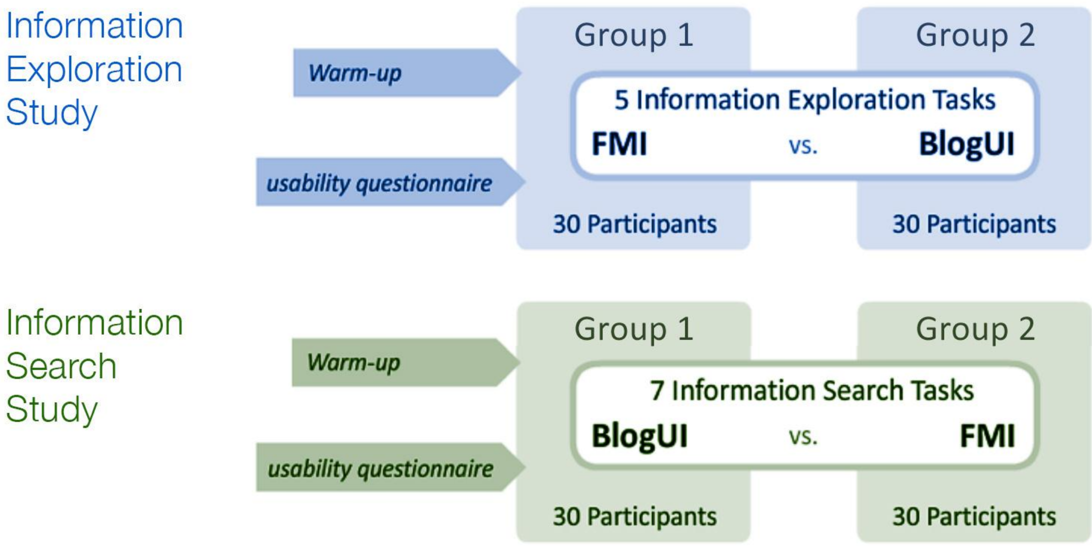


Figure 53: Schema of Experimental Design

The studies were conducted in a usability lab using a Tobii X50 eye tracker. Warm-up sessions in the beginning of each study ensured that all participants were introduced to the test environment and were given time to familiarize with the functionality of the user interface. Participants were calibrated separately for each study to improve the quality of the collected data.

6.5. Results

Task performance was measured using task rates and task completion times for the goal-oriented information search study. User feedback was collected after the completion of each study through detailed usability questionnaires.

User interaction behaviour such as articles read, searches made and categories chosen was analysed for both the information search study, as well as the more behaviourally oriented information exploration study. Traditional eye tracking measures such as fixation counts and average fixation durations were also analysed to support findings of performance and behavioural measures. Areas of Interest (AOIs) have been defined for navigational elements (Navigation AOI) and the content sections (Content AOI) within each respective UI.

6.5.1. Task Performance Analysis

The information search study required participants to conduct seven tasks. As mentioned in the description of the information search tasks, six of the tasks were "standard" tasks, while one task was more difficult, with the respective information hidden within the final few paragraphs of one particular article. The error distribution across individual tasks (see Table 6) highlights the problems participants had to successfully finish this difficult task (T3). Due to the structural differences of T3, we analysed task performance for the remaining tasks separately. A detailed reflection on T3 and its design implications can be found in the discussion section.

Table 6: Task Errors for individual search tasks (across all participants)

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
FMI	0	0	13	0	0	2	0
BlogUI	3	0	9	1	1	3	2

6.5.1.1. Task Error Rates

Overall, participants using the FMI (for traditional tasks 1, 2, 4-7) made 2 errors, resulting in an error rate of 1.15% (SD=10.69%). Participants using the BlogUI (for traditional tasks 1, 2, 4-7) made 10 errors, resulting in an error rate of 5.56% (SD=22.97%). In comparison, participants using the contextual navigation of the FMI made ~**80% fewer errors,** than participants using the traditional BlogUI. A two-sample t-test underlines the significance of this difference with $t_{57} = 2.00$, $p < 0.02$.

One additional aspect to consider is the distribution of task completion errors over the various tasks. Using the FMI, in 5 out of the 6 traditional tasks, all participants successfully completed their tasks. In contrast, using the BlogUI, only in 1 out of the 6 traditional tasks, none of the participants failed to complete the task successfully.

This finding hints at a general task-independent problem with the BlogUI for information search tasks. In contrast, the concentration of task completion errors for the FMI on specific tasks (task 3) hints at a task-dependent problem for information search tasks. Future iterations of the FMI prototype should investigate this phenomenon to further minimise task completion errors.

6.5.1.2. Task Completion Times

For successfully completed tasks, participants' task completion times were ~**19% faster** using the FMI (39 seconds, SD=24s) than for participants using the BlogUI (48 seconds, SD=38s). This difference is significant with $t_{338} = 1.97$, $p < 0.01$ (using a two-sample t-test). A breakdown of task completion times for each task can be found in Table 7 below.

Table 7: Mean task completion times for individual search tasks (across all participants) in seconds

	Information Search Tasks			
	FMI		**BlogUI**	
	Mean	*SD*	Mean	*SD*
Task 1	46s	*22s*	81s	*52s*
Task 2	48s	*32s*	56s	*37s*
Task 4	26s	*10s*	30s	*17s*
Task 5	39s	*23s*	42s	*40s*
Task 6	39s	*23s*	34s	*31s*
Task 7	36s	*23s*	43s	*22s*
Overall Mean	**39s**	***24s***	**48s**	***38s***

Another interesting finding is the much larger spread of successful task completion times for the BlogUI. Ranging from 6 seconds for the fastest to 212 seconds for the slowest task completion time, the spread of 206 seconds for the BlogUI is 44% larger than the spread for the FMI, with 115 seconds (9 seconds for the fastest, and 124 seconds for the slowest task completion time).

This finding is particularly interesting, when considering the very low error rate for the FMI. Although this user interface and its interaction technique were completely new to all participants, not only did participants make fewer errors, and completed tasks in less time overall, but participants also showed more consistent task completion times, than in the more familiar blogUI.

As both measures, error rates and task completion times, indicate a clear performance advantage for the FMI (over the BlogUI), I accept the first hypothesis (H1).

6.5.2. User Feedback Analysis

Differences measured for task performance indicated a clear advantage of FMI over BlogUI in form information search tasks. However, to successfully shift users' information experiences, users need to be comfortable with using a novel UI and adopt it in the long term. It is thus vital to not just measure performance but also users' perceptions of any novel UI concept. Other approaches of focus+context visualisations have proven in the past to be superior to traditional UIs but not popular with users. To build a rich picture of participants' subjective preferences, a usability questionnaire was administered after completion of information exploration tasks and after completion of information search tasks to evaluate the particular UI tested for those tasks (see section 6.5.2.1). The usability questionnaire consists of two parts:

1. First impression was captured using a set of 6 questions from the *"Questionnaire for User Interface Satisfaction"* (QUIS) (e.g.: "The system was: Frustrating – Satisfying").
2. Usability criteria *ease of use*, *learnability*, *UI and content*, *productivity,* and *engagement* were captured using a set of 24 standard and tailored usability questions (largely based on the Computer System Usability Questionnaire – CSUQ). A 6-point Likert scale was used for these questions.

After completion of both information exploration and information search tasks, participants were given a final questionnaire to compare both UI types directly (see section 6.5.2.2). Participants were asked to provide the most negative and the most positive aspects of each system – a summary of that feedback can be found in Table 9. Further, participants were asked to choose between both UIs with regards to the following three questions:

- Which interface did you like the most?
- Which interface did you find better to navigate?
- Which interface did you find easier to use?

For each of the three selections, participants were asked to provide a brief rationale.

6.5.2.1. Usability Questionnaire Analysis

Across information search and information exploration tasks, the overall usability rating is 4.79 (SD = 1.15) for the BlogUI, and 4.69 (SD = 1.15) for the FMI. This difference is not significant. The overall rating of both UIs is quite positive, considering the baseline at 3.5 for a 6-point Likert scale. Results broken down by various usability criteria in Table 8 show no quantifiable advantage for most usability criterion for either of the two

UIs. The only significant difference between BlogUI and FMI is for learnability in the search condition, favouring the BlogUI ($t_{57} = 2.00$, $p < 0.05$).

Table 8: Usability Questionnaire Results

	Information Search Tasks				Information Exploration Tasks			
	BlogUI		FMI		BlogUI		FMI	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Ease of Use	4.7	*1.1*	4.7	*1.2*	4.8	*1.2*	4.8	*1.1*
Learnability	5.0	*1.1*	4.7	*1.1*	5.1	*1.1*	5.0	*1.0*
UI & Content	4.6	*1.1*	4.6	*1.3*	4.8	*1.2*	4.8	*1.0*
Productivity	4.7	*1.2*	4.7	*1.2*	4.8	*1.1*	4.6	*1.1*
First Impression	4.5	*1.1*	4.6	*1.0*	4.6	*1.1*	4.6	*1.1*
Overall Average	**4.7**	***1.1***	**4.7**	***1.2***	**4.8**	***1.2***	**4.7**	***1.1***

In light of the very one-sided results of performance, and considering that most participants had a rather strong view on which UI they preferred, the measured conformity of overall subjective evaluations came at a surprise.

6.5.2.2. Direct Comparison Questionnaire Analysis

The overall results of the direct comparison show no clear preference for either of the 2 UI Types. However, there seems to be a slight preference for information search tasks independent of the UI used. This could be due to shorter task times, search tasks being more straightforward, or simply the fact that these tasks were tested 2nd to the information exploration tasks. Nevertheless, the experimental design counterbalanced such effects for a comparison between UIs (see Figure 53).

The analysis of feedback provided by participants regarding the most positive and most negative aspects of each UI is quite mixed (see Table 9). While there were strong positives for both FMI and BlogUI regarding the quality or amounts of useful information, as well as the ease of use, some differences can be made out that relate to the nature of the UI. Asked about the most positive aspects of each UI some participants felt that the BlogUI was *"familiar"* and *"looks professional"*, others felt that the FMI was *"fun, quicker, easier"*, *"fun to read"*, *"futuristic looking"*, and *"feels positive, happy to use"*.

Asked about the most negative aspects of each UI, some participants felt that the BlogUI was *"boring"*, required *"too much scrolling"*, and that its topics were *"unclear"* or *"overlapping"*, while others suggested that the FMI was *"messy"*, *"hard to go back"*, and *"easy to get lost"*.

Table 9: Participant feedback on negative and positive aspects of each UI

	BlogUI	**FMI**
Positive	Clearly structured Clear categories Index, easier to navigate Keywords at end of article made it easier to navigate Larger variety of information More information Familiar Looks professional Like using scrollbar	No time to lose interest High speed Options contain more information Clearly displayed information No contradictory articles Better overview, more info on screen Very precise, more information for a topic Fun, quicker, easier Fun to read, no long list of links Futuristic looking Less searching Easier to navigate More intuitive Feels positive, happy to use More natural to use Easier for novices Easier to use without keyboard No scrolling down during search
Negative	Category structuring overlaps or is unclear Articles contradict each other Too much scrolling Boring Not enough topics Can't get a short summary of articles	Hard to go back Easy to get lost No mouse wheel support Messy Synopsis is sometimes misleading No index Did not feel in control My eyes weren't comfortable Little boxes did not provide enough information Annoying animation Longer list for previous articles is needed No overview of all topics/articles Easy to waste time Can't browse Hard to notice search box

When asked *"Which interface did you find easier to use"*, a slight preference for the FMI (53% vs. 47% for BlogUI) was measured across task types. Although this difference seems marginal, it should be mentioned that a quarter of participants favouring the BlogUI mentioned *familiarity* as the key reason why they found it *"easier to use"*. Familiarity might also explain why 58% of participants preferred the BlogUI (vs. 42% preferring the FMI) when asked *"Which interface did you find better to navigate"* - despite clear advantages of the FMI for task completion times and error rates.

When asked *"Which interface did you like the most"*, 49% of participants chose the FMI compared to 51% of participants who chose the BlogUI. Considering the fact, that the BlogUI is a much more traditional type of web-based UI, and none of the participants used the FMI prototype in advance to this study nor were familiar with its concept, familiarity could also play a crucial role in explaining the difference for the learnability measure in the results of the main usability questionnaire.

As the analysis of the usability questionnaires shows very similar results overall and the direct comparison questionnaire also does not surface any clear preference for either the FMI or the BlogUI, I reject the second hypothesis (H2).

However, when considering participants' contrasting statements, such as *"easy to get lost"* or *"did not feel in control"* versus *"more intuitive"*, *"more natural to use"* or *"very precise"* (see Table 9) - all on the FMI by different participants, it becomes apparent that strong individual differences do exist but are somewhat obscured by the overall results. Additional analyses of the results have thus been conducted by comparing participants' results based on their computer expertise.

6.5.2.3. User Satisfaction by Computer Expertise

The re-analysis of the usability questionnaire broken down by expert and novice participant responses reveals contrasting evaluations of the BlogUI and FMI. Overall, experts rate the BlogUI higher than the FMI, and novices rate the FMI higher than the BlogUI (see Table 10), however those differences are not significant.

Figure 54 visualises accumulated ratings for the various usability criteria by UI type and computer expertise (see section 6.4.2.1) based on normalised differences from the overall mean for each UI type. Results show, that the homogeneous evaluation observed in the initial analysis of the questionnaire data (see Table 8) is rooted in contrasting ratings between experienced and novice participants. Whereas experts demonstrate a general preference for the BlogUI, novices generally prefer the FMI. Furthermore, in the direct comparison, ~57% of novices found the FMI easiest to use (up from 53% for all participants), whereas ~66% of experts found the BlogUI better to navigate (up from 58% for all participants).

Table 10: User Satisfaction by CE Score

	Expert participants (CE score > 0.4)				Novice participants (CE score <= 0.4)			
	BlogUI		**FMI**		**BlogUI**		**FMI**	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Ease of Use	4.9	*0.8*	4.6	*1.0*	4.7	*1.1*	4.9	*0.8*
Learnability	5.1	*0.9*	4.8	*0.9*	5.0	*0.9*	5.0	*0.7*
UI & Content	4.7	*1.0*	4.5	*1.0*	4.7	*1.0*	5.0	*0.8*
Productivity	4.9	*1.1*	4.5	*0.9*	4.7	*1.1*	4.8	*1.0*
Engagement	4.9	*1.0*	4.5	*0.9*	4.5	*0.8*	4.5	*1.0*
First Impression	4.6	*0.8*	4.4	*0.8*	4.5	*0.9*	4.8	*0.9*
Overall Average	**4.80**	***0.82***	**4.52**	***0.80***	**4.64**	***0.83***	**4.82**	***0.71***

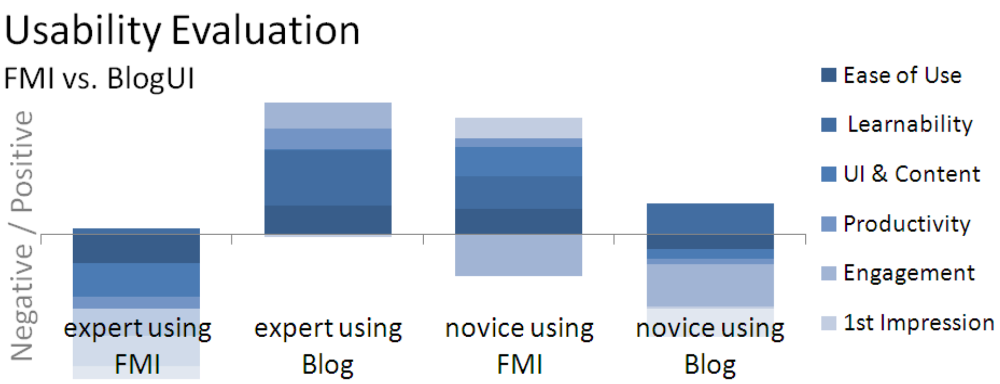


Figure 54: Usability Evaluation by UI type and IT score

6.5.3. Information Interaction Analysis

An exploratory analysis of users' interaction behaviour has been conducted to augment the insights gained through the analyses of task performance and questionnaire responses. In both UIs, participants interacted with information through the use of search and additional UI-specific interactions. Participants were free to choose at any point which means of interaction to use. In the BlogUI, a category list has been the central navigation element. Choosing a particular category would load a list of related articles, which participants could scroll through. In the FMI, the main interaction mechanism has been its contextual navigation (elements).

6.5.3.1. Search Usage

For the use of search, our findings show that participants conducted significantly more searches (approx. 47%) using the FMI, with 1.52 searches per task for the FMI versus 1.03 searches per task for the BlogUI ($t_{56} = 2.00$, $p < 0.005$) in the information search condition (see Figure 55).

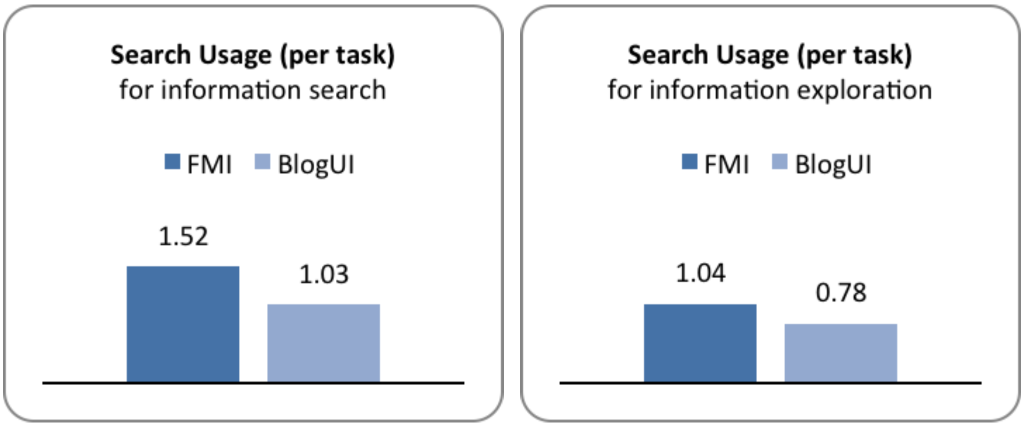


Figure 55: Search Interaction per task for both information task conditions and both UIs

Participants also conducted approx. 34% more searches using the FMI in the information exploration study than in the information search study. However this result is not significant, due to strong individual differences in user behaviour: Approx. a third of participants did not use search at all during information exploration, whereas other participants heavily relied on search (with a peak of 4.4 searches per task for both UIs). The lower scores for search usage during information exploration tasks are particularly interesting when considering the overall task duration of approx. 5min for these tasks – compared to a task duration of 48 seconds (BlogUI) or 39 seconds (FMI) for the information search tasks.

6.5.3.2. Other Interactions

In the information exploration tasks, participants preferred alternative means of interaction for both UIs (FMI: contextual elements, BlogUI: category list) (Figure 56). In contrast, when faced with a specific search task, usage of the search tool increased (see Figure 55).

Using the BlogUI, participants clicked on average on 2.3 categories during each exploration task (approx. 5 min). Using the FMI, participants interacted on average with 7.2 articles per task (through selection of a contextual navigation element).

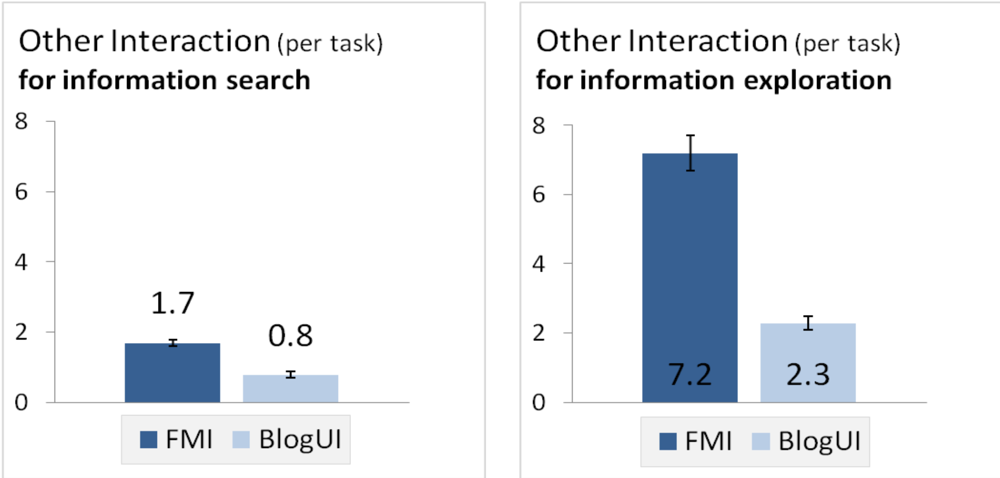


Figure 56: Other Interaction per task for both information task conditions and both UIs (category selection for BlogUI and context selection for FMI)

The distinctly different visual representations and types of interaction on the same blog space make a direct comparison of category interaction (BlogUI) and context interaction (FMI) difficult. In the FMI, selecting a contextual element will display the one selected article plus previews of related articles by updating the contextual navigation elements. In the BlogUI, selecting a category item will load the respective category page with a number of related articles (up to 10). Here, the participant has no direct control over which specific article(s) will be loaded, and particularly, which article will appear on top of the page and will therefore be visible "above the fold".[5]

The interaction technique underlying the BlogUI allows participants to scroll through a list of given articles, more or less one by one. The amount of articles skimmed over varies greatly across participants with some barely scrolling down and others scrolling down to the last element. Hardly ever did a user request a second result page (when more than 10 articles would belong to this category).

This distinct difference in interaction behaviour promotes a more passive consumption of information when using the BlogUI. By offering another article directly following the previous (through scrolling down), the "system makes the decision" for the participant. Considering the lack of overview, as participants cannot perceive which articles are contained within the list (without scrolling to the bottom), and the lack of screen estate, which rarely allows more than 1 or 2 articles to be visible concurrently, the BlogUI encourages participants to simply read the next article.

[5] The part of the web page "above the fold" is defined by screen resolution and relates to the area visible without scrolling down.

In contrast, the interaction technique underlying the FMI requires participants to actively select each article they want to read in full. Selecting a contextual element will load the full article in the centre of the UI and load a set of related elements (which replace the previously displayed elements) in the contextual area of the UI. This approach may encourage a more active decision making process.

6.5.4. General Eye Tracking Analysis

The following eye tracking analysis has been conducted to provide further inside into the user behavior during information search and information exploration tasks using the FMI and BlogUI interface types. The following gaze time analysis (see section 6.5.4.1) focuses on participants' attention distribution across content and navigation elements for both user interfaces. Gaze time data has been used in favor of fixation count data to account more accurately the actual time participants spent gazing at various parts of the interface (taking into account differences in fixation durations). A separate analyis of average fixation durations in reported in section 0.

6.5.4.1. Gaze Time Analysis

A summative analysis of the gaze time data (see Table 11 and Table 12) reveals significant differences between how participants use FMI and BlogUI across information exploration and information search tasks.

Relative gaze times across information search tasks are:

- significantly higher on the content section in the BlogUI than on the content section in the FMI (60.7% vs. 37.6%, $t_{52} = 3.49$, $p < 0.001$).
- significantly lower on the navigation section in the BlogUI than on the navigation section in the FMI (39.3% vs. 62.4%, $t_{52} = 3.49$, $p < 0.001$).
- significantly higher on the content section than on the navigation section in the BlogUI (60.7% vs. 39.3%, $t_{26} = 3.71$, $p < 0.001$).
- significantly lower on the content section than on the navigation section in the FMI (37.6% vs. 62.4%, $t_{26} = 3.71$, $p < 0.001$).

Relative gaze times across information exploration tasks are:

- significantly higher on the content section in the BlogUI than on the content section in the FMI (87.8% vs. 67.2%, $t_{56} = 3.47$, $p < 0.001$).
- significantly lower on the navigation section in the BlogUI than on the navigation section in the FMI (12.2% vs. 32.9%, $t_{56} = 3.47$, $p < 0.001$).

- significantly higher on the content section than on the navigation section in the BlogUI (87.8% vs. 12.2%, $t_{28} = 3.67$, $p < 0.001$) and in the FMI (67.2% vs. 32.9%, $t_{28} = 3.67$, $p < 0.001$).

Table 11: Mean Gaze Time Distribution Analysis (in seconds)

	Information Search Tasks				Information Exploration Tasks			
	BlogUI		FMI		BlogUI		FMI	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Content	139.2s	*71.5s*	62.7s	*36.3s*	799.4s	*360.4s*	613.6s	*334.5s*
Navigation	95.5s	*78.0s*	108.2s	*56.7s*	94.6s	*54.2s*	260.6s	*105.0s*

Table 12: Relative Gaze Time Distribution Analysis (in percent)

	Information Search Tasks				Information Exploration Tasks			
	BlogUI		FMI		BlogUI		FMI	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Content	60.7%	*12.4%*	37.6%	*11.6%*	87.8%	*7.1%*	67.2%	*11.6%*
Navigation	39.3%	*12.4%*	62.4%	*11.6%*	12.2%	*7.1%*	32.9%	*11.6%*

The measured contrast between BlogUI and FMI is likely rooted in the conceptual differences of how navigation is used. The preview snippets in the navigational elements of the FMI allow people to skim through navigational choices before selecting an article. In contrast, the category links in the BlogUI do not provide any details about the underlying articles. Users are required to skim through a list of the actual articles (after choosing a category) to identify relevant information. Scanning or skim-reading is an important part of *"navigational decision making"* – particularly when browsing the Web. As a result, a substantial part of gaze time within the content section of the BlogUI is conceptually used for navigating. However, while some of the time spent on the content section in the BlogUI may be attributed to navigational behaviour, the differences observed in task performance suggest that there is a link between the type of navigation afforded by the UI and the type of task that needs to be completed.

Gaze Time Distribution for Contextual Navigation in FMI

The analysis of participants' attention distribution across individual contextual navigation elements in the FMI shows a quite homogenous distribution for information exploration and information search tasks (see Figure 57). For the search tasks, the analysis shows a preference for the first contextual element (which is also the most related one). The trend line for search tasks (with a quite good fit of R^2=0.74) indicates a steady but slow decay of attention (negative slope) from the first to the last contextual element.

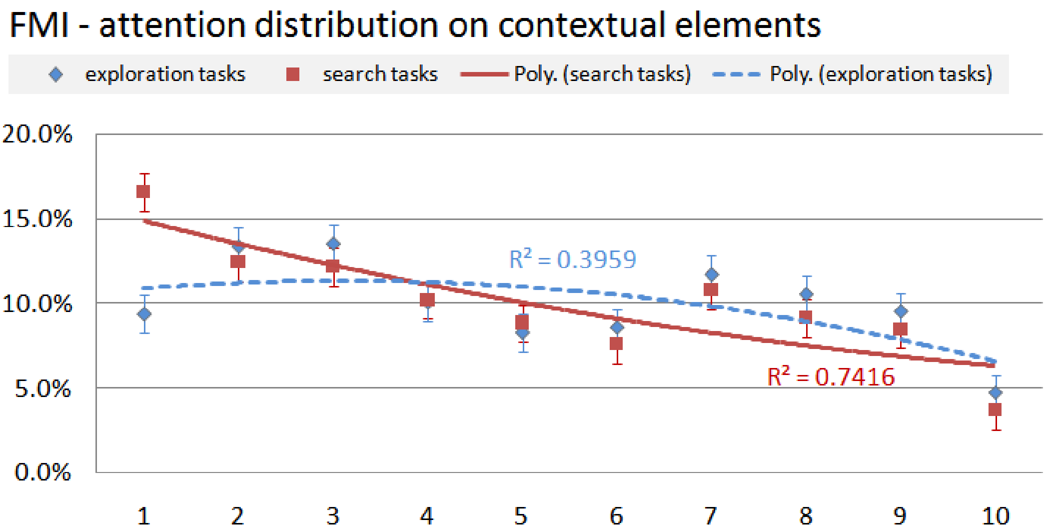


Figure 57: Overall Gaze Time Distributions for Contextual Elements of the FMI for Information Search and Information Exploration

Exemplary gaze plots for specific information search tasks using the FMI illustrate the homogenous attention distributions for two individual participants (see Figure 58). In particular, when looking at the individual fixations and how they are connected, which illustrates the path of the participant's gaze, it becomes clear that the eye 'navigates' from one contextual element to the next and does not 'jump' long distances within the UI.

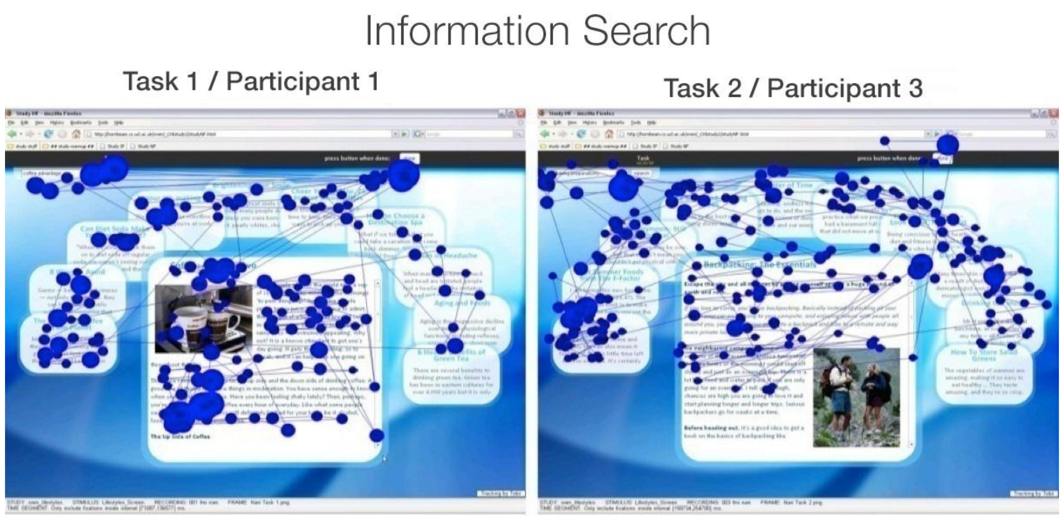


Figure 58: Individual Gaze Plots for Information Search (Task 1, Task2)

Gaze plots for the BlogUI (Figure 59) illustrate the much less homogeneous distribution of attention, primarily rooted in the visual nature of the UI. Participants' attention moves down the page to process articles in a more linear manner. This interaction requires a lot of scrolling, and can easily lead to "attention gaps", where individual articles are not fixated at all – intentionally or unintentionally. These findings provide some potential explanation for the observed differences in task performance between FMI and BlogUI as well the reported differences in user satisfaction between novices and experts (see section 6.5.2.3).

Figure 59: Gaze Plots for various pages in the BlogUI

6.5.4.2. Average Fixation Duration Analysis

The overall analysis of average fixation durations reveals significant differences between FMI and BlogUI for exploration tasks ($t_{56} = 2.00$, $p < 0.05$) and for search tasks ($t_{52} = 2.40$, $p < 0.02$) (see Figure 60 and Table 13).

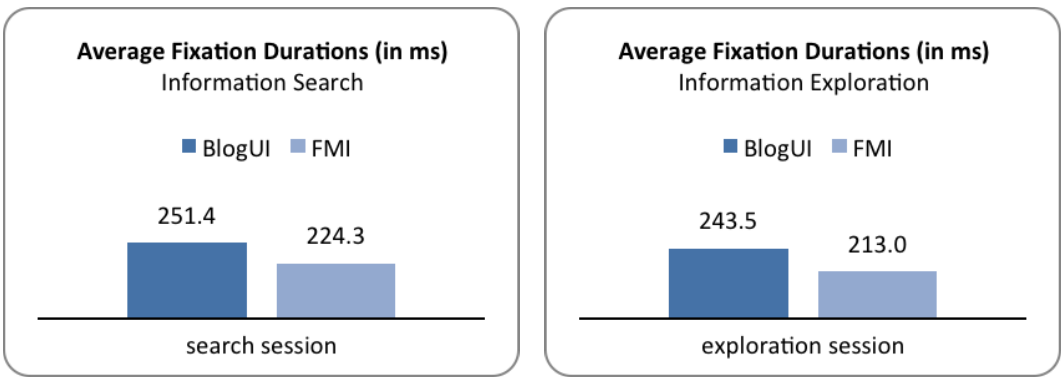


Figure 60: Average fixation durations - BlogUI and FMI for Information Exploration and Information Search condition

A further distinction of areas of interest (AOIs) for content and navigation in the respective UIs reveals that the navigation AOI is responsible for the increased average fixation durations in the BlogUI (see Table 13 and Figure 61). For information search tasks, participants' fixations on the category list in the BlogUI last ~48% longer than fixations on the contextual elements in the FMI (335.4ms vs. 227.4ms). This finding is significant ($t_{52} = 3.49$, $p < 0.001$). For information exploration tasks, participants' fixations on the category list in the BlogUI last ~53% longer than fixations on the contextual elements in the FMI (360.9ms vs. 235.3ms). This finding is significant ($t_{56} = 3.47$, $p < 0.001$).

Table 13: Average Fixation Durations (in ms)

	Information Search Tasks				**Information Exploration Tasks**			
	BlogUI		**FMI**		**BlogUI**		**FMI**	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Overall	251.4	*48.0*	224.3	*30.2*	243.5	*44.0*	213.0	*34.4*
Navigation	335.4	*100.0*	227.4	*30.3*	360.9	*110.5*	235.3	*39.2*
Content	210.6	*33.3*	209.7	*36.3*	234.4	*43.7*	203.4	*34.5*

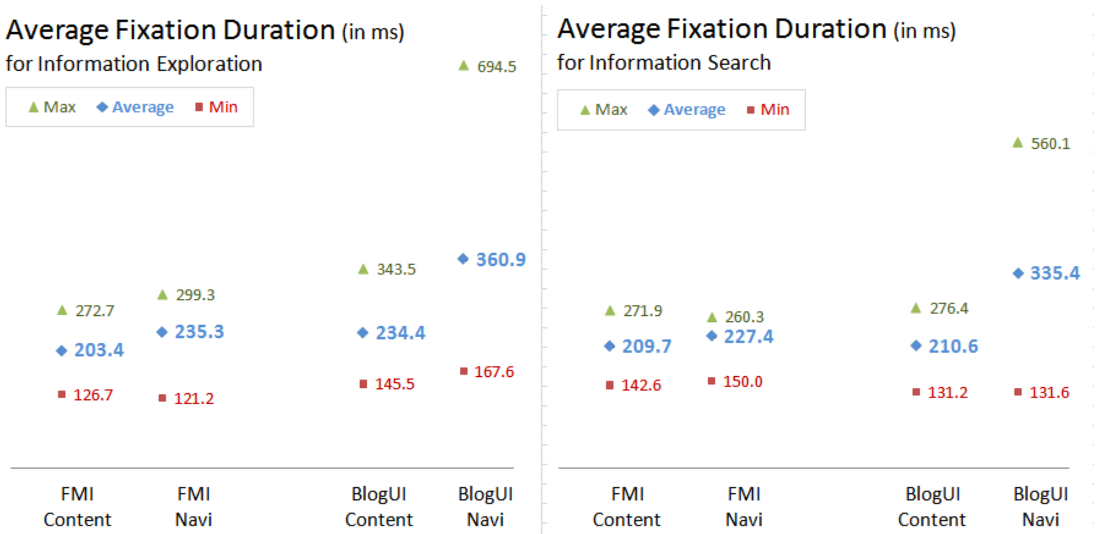


Figure 61: Average fixation duration breakdown for Content and Navigation AOI (for both UI types and both task types)

Increased average fixation durations for the search condition, independent of the UI used, are caused by participants' increased attention on navigational elements and the fact that average fixation durations are generally higher for the Navi AOI, than for the Content AOI. The more goal-driven nature of the search tasks required participants to find the right article (making more use of navigational elements), by skimming potentially relevant articles rather than reading them thoroughly. Thus, the BlogUI shows a ~250% increase in attention on navigational elements for the search tasks (when compared to the exploration tasks). The FMI also shows a ~100% increase in attention on navigational elements (see section 6.5.4.1).

6.5.5. Exploration of Differences by Computer Expertise

The evaluation of user satisfaction by computer expertise (CE) in section 6.5.2.3 revealed contrasting preferences for expert and novice participants. The following post-hoc analyses further investigate the impact of computer expertise on task performance (see section 6.5.5.1) and on interaction behaviour (see section 6.5.5.3).

6.5.5.1. Impact of CE on Task Performance

Average task completion success rates by computer expertise (CE) show only marginal differences between expert and novice participants for both UIs (Figure 62).

Figure 62: Breakdown of Task Completion Rates by Computer Expertise

Analysing task completion success rates of novice participants reveals that:

- Four novice participants using the BlogUI did not complete all six tasks successfully. One novice participant made two errors (67% task completion rate) and three participants made one error (83% task completion rate).
- Two novice participants using the FMI did not complete all six tasks successfully. Both participants made one error (83% task completion rate).

Analysing task completion success rates of expert participants reveals that:

- Five expert participants using the BlogUI did not complete all six tasks successfully. All of them made one errors (83% task completion rate).
- One expert participants using the FMI did not complete all six tasks successfully, making one error (83% task completion rate).
- The analysis of task completion times for novices and experts shows that experts completed search tasks significantly faster than novices (see Figure 63 and

Table 14):

- Using the BlogUI, experts were significantly faster than novices (41s vs. 55s, $t_{167} = 1.97$, $p < 0.05$)
- Using the FMI, experts were significantly faster than novices (34s vs. 44s, $t_{169} = 1.97$, $p < 0.05$)
- Novices using the FMI were significantly faster then when using the BlogUI (55s vs. 44s, $t_{165} = 1.97$, $p < 0.05$)
- Experts using the FMI were not significantly faster then when using the BlogUI (41s vs. 34s)

Table 14: Average Task Completion times (in seconds)

	Information Search Tasks							
	Novice Participants				**Expert Participants**			
	BlogUI		**FMI**		**BlogUI**		**FMI**	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Task 1	108s	*63s*	53s	*24s*	60s	*28s*	39s	*17s*
Task 2	61s	*32s*	63s	*38s*	53s	*41s*	33s	*13s*
Task 4	30s	*12s*	29s	*10s*	31s	*21s*	23s	*11s*
Task 5	52s	*52s*	43s	*23s*	32s	*22s*	35s	*24s*
Task 6	33s	*19s*	38s	*20s*	35s	*41s*	41s	*27s*
Task 7	52s	*24s*	37s	*18s*	35s	*18s*	35s	*28s*
Overall	55s	*44s*	44s	*26s*	41s	*31s*	34s	*21s*

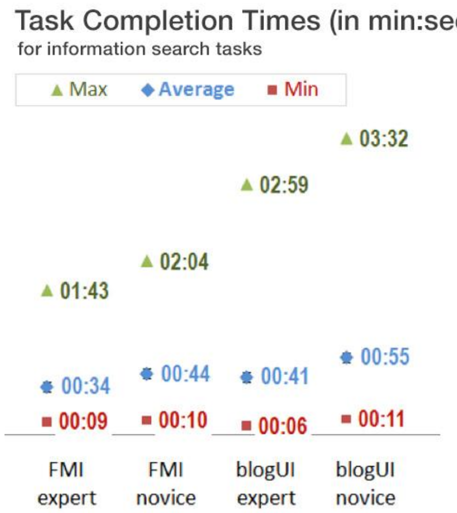


Figure 63: Breakdown of Task Completion Times by Computer Expertise (the standard error of SE_x=3sec for average task completion times is hardly visible in the graph)

To further investigate the strong impact of computer expertise, a mapping of all task completion times against participants' CE score has been created on a scatter plot (see Figure 64).

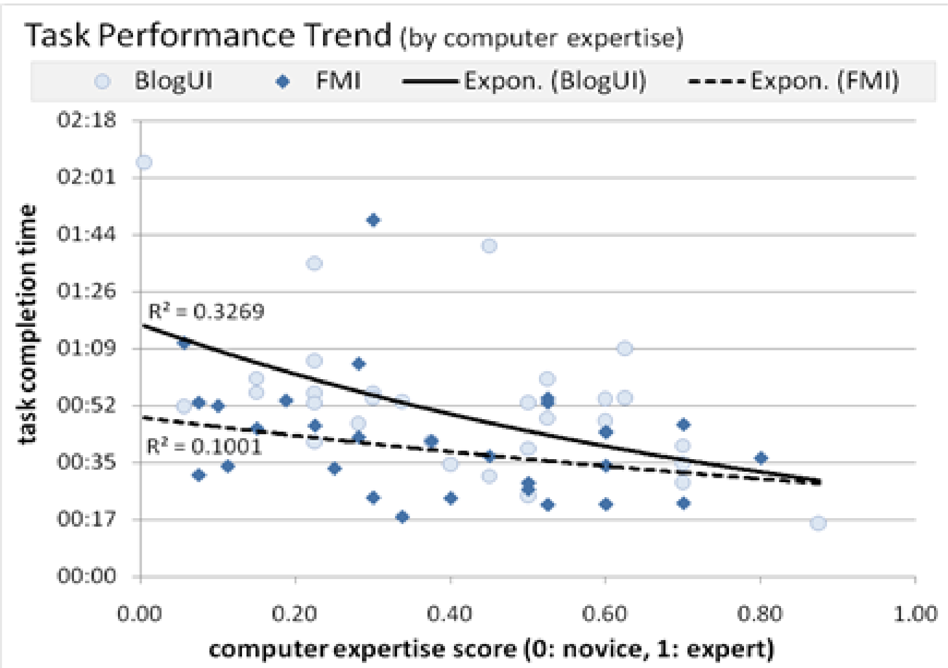


Figure 64: Task Performance Trend for Computer Expertise (CE) in the information search condition (using average task completion times for each participant, based on task 1, 2, 4, 5, 6 and 7)

- The most obvious difference is the bigger range of task times for the BlogUI, which is quantified by the larger standard deviations as shown in

Table 14. Trend lines suggest that task performance advantages of the FMI origin particularly from participants with lower CE scores. The contextual navigation of the

FMI seems to provide larger (task performance) benefit to participants with lower CE scores. This advantage of the FMI over the BlogUI seems to decrease with increasing CE score and basically vanishes for very experienced participants.

Low values for R-square indicate high variability in the data and suggest limited reliability in the trend lines. However, the accuracy of the trend lines is not that important for the above interpretation in the sense that we can deduct reliable task completion times for specific CE scores. What is more important is the general difference between the two trend lines, and the difference in their slopes.

It is also important to note that experienced participants are more familiar and trained with using the Web and in particular with using blogs: From 16 experienced participants using the BlogUI for information search tasks, 7 were using blogs (~44%), whereas none of the 14 novice participants was familiar with blogs (beyond the experience gained in the warm-up session). The obvious decrease in task completion times for the BlogUI with increasing CE score hints at learning effects of more experienced users.

None of the participants using the FMI for information search tasks was familiar with the FMI (besides warm-up session). As a consequence, it might be reasonable to assume that experienced users would gain back performance advantages over time, when using the FMI on a regular basis.

6.5.5.2. Impact of CE on Average Fixation Durations

A second trend analysis has been calculated for average fixation durations against the CE score, for information search tasks (see Figure 65). While these trends are very weak, their discussion can provide some addition insights into the differences between BlogUI and FMI. The trend line for the BlogUI shows that with increasing computer expertise, average fixation durations become shorter. Considering average fixation duration measures as an indicator for cognitive load, this finding correlates with the task performance trend. With increasing computer expertise, participants find it easier (thus shorter average fixation durations) to complete the given tasks (thus faster task completion times).

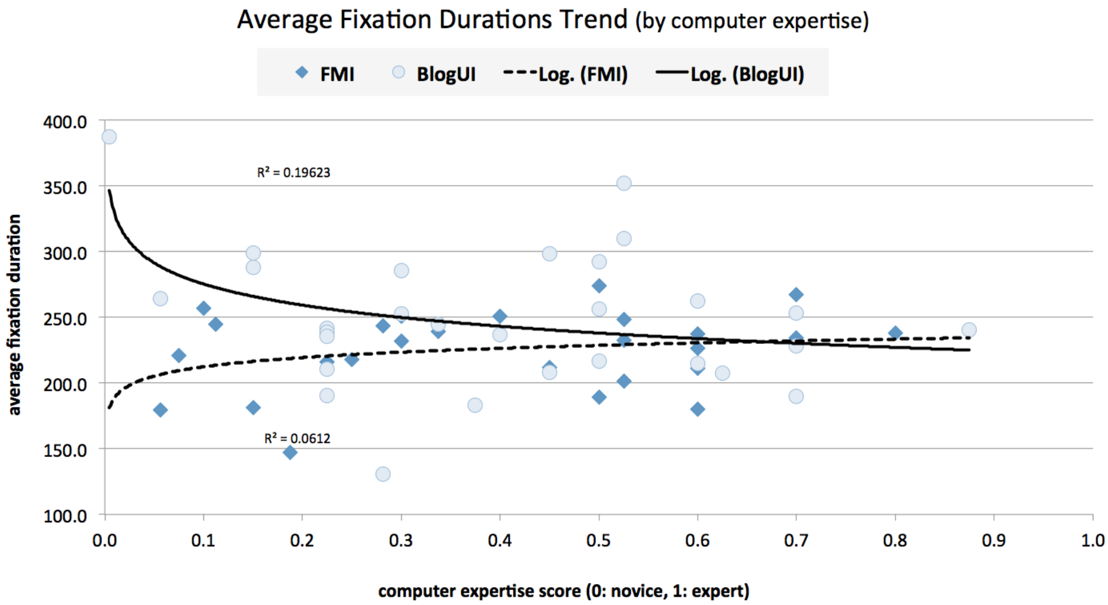


Figure 65: Average Fixation Durations Trend for Computer Expertise (CE)

The contrasting trend in the average fixation durations analysis for participants using the FMI is surprising at first, as with increasing CE score, average fixation durations become longer. Without the CE score-related effects, one could argue that general average fixation duration differences are merely an effect of visual layout differences between BlogUI and FMI. However, as the contrasting trends for the BlogUI and the FMI in Figure 65 demonstrate, differences might be grounded in (cognitive) schema "reuse" as well. Along with cognitive load theory, more experienced participants seem to be able to rely on existing schema to complete the given tasks with less effort - average fixation duration being an indicator for level of visual and cognitive processing - and in less time (task completion time) using the BlogUI. As the FMI is novel to all participants, even experienced participants could not rely on existing (cognitive) schema to aid them in completing the given tasks. However, as experienced participants are used to rely on their acquired schema to solve these tasks, this could potentially explain why average fixation durations are not just constant but even go up for experienced participants.

6.5.5.3. Impact of CE on Interaction Behaviour

Search usage during information exploration tasks has shown strong differences between individual participants as reported in section 6.5.3.1. The analysis of interaction behaviour by CE shows that differences can be found for expert versus novice participants using the BlogUI (see Figure 66 - left):

- On average, experts conducted ~130% more searches than novices using the BlogUI (5.5 vs. 2.4 across all tasks; not statistically significant due to high individual differences).
- The "most active" expert conducted 175% more searches than the "most active" novice (with 22 vs. 8 searches across all tasks).

These results indicate that more experts feel more comfortable navigating the BlogUI through search queries than novices do.

The analysis of search usage during information exploration tasks shows no similar difference between experts and novices when using the FMI (see Figure 66 - left):

- On average, experts conducted ~6% less searches than novices using the FMI (5.1 vs. 5.4 across all tasks; not statistically significant due to high individual differences).
- The "most active" expert conducted ~22% more searches than the "most active" novice (with 22 vs. 18 searches across all tasks).

One potential explanation for this contrast could be that experts (who are all familiar with blogs) have developed a schema to rely more on search when browsing blogs (e.g. because they have learnt that alternative means of navigation are not as efficient).

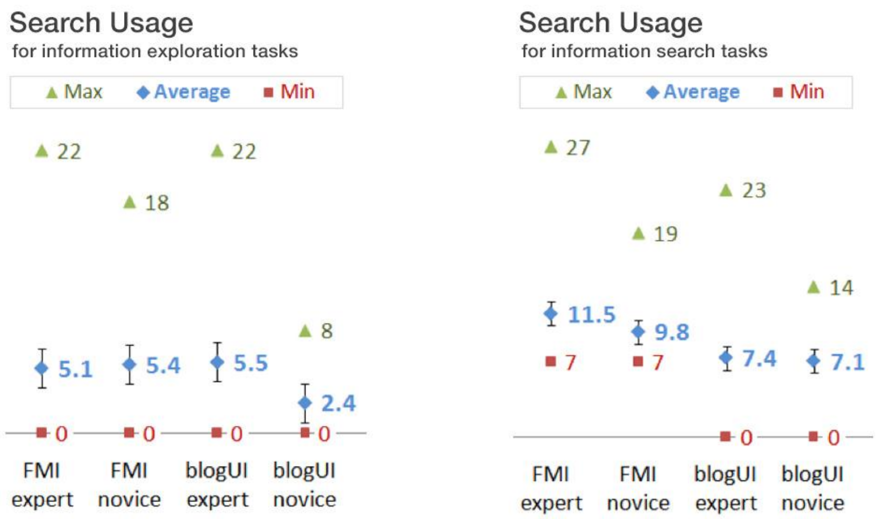


Figure 66: Breakdown of Search Usage for Information Exploration Tasks (left) and Information Search Tasks (right) by Computer Expertise

For information search tasks, differences in search usage between experts and novices appear to be more marginal when using the BlogUI (see Figure 66 – right):

- On average, experts conducted ~4% more searches than novices using the BlogUI (7.4 vs. 7.1 across all tasks; not statistically significant due to high individual differences).

- The "most active" expert conducted 64% more searches than the "most active" novice (with 23 vs. 14 searches across all tasks).

Interestingly, differences in search usage between experts and novices appear to be somewhat stronger when using the FMI for information search tasks (see Figure 66 – right):

- On average, experts conducted ~17% more searches than novices using the FMI (11.5 vs. 9.8 across all tasks; not statistically significant due to high individual differences).
- The "most active" expert conducted 42% more searches than the "most active" novice (with 27 vs. 19 searches across all tasks).

The observed strong increase for novice participants using the BlogUI when comparing search usage between information exploration tasks (2.4 searches) and information search tasks (7.1 searches) could be caused by the more goal-driven nature of the information search tasks. Asking participants to find a specific article seems to encourage most participants to use search.

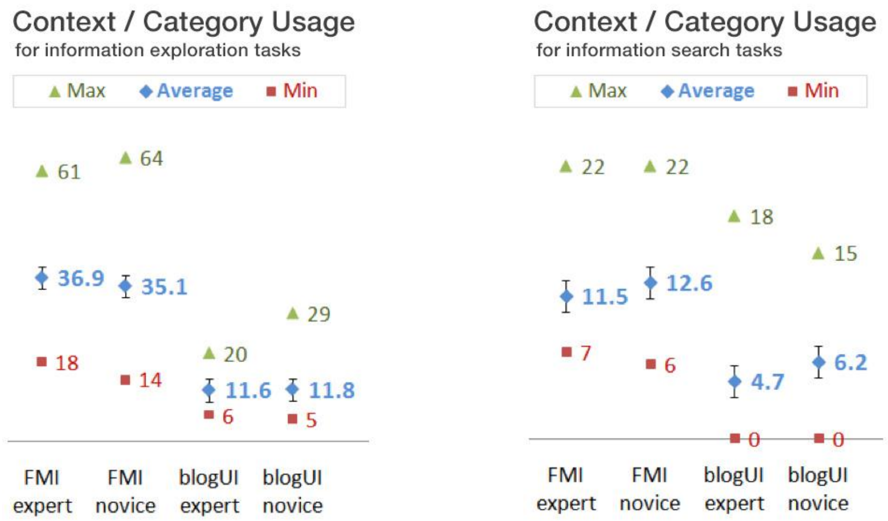


Figure 67: Breakdown of Context / Category Usage for Information Exploration Tasks (left) and Information Search Tasks (right) by Computer Expertise

The analysis of other interactions, such as context usage (FMI) and category usage (BlogUI) shows very similar results for expert and novice participants for both UIs (see Figure 67):

- Using the FMI, only marginal (not statistically significant) differences between experts and novices have been observed regarding the use of contextual elements for both information exploration tasks (36.9 vs 35.1) and information search tasks (11.5 vs. 12.6)

- Using the BlogUI, only marginal (not statistically significant) differences between experts and novices have been observed regarding the use of the category list for information exploration tasks (11.6 vs 11.8)
- The difference in use of the category list (BlogUI) between experts and novices for information search tasks is reasonably large (4.7 vs 6.2). However, this difference is still not significant due to the huge individual differences in user behaviour (see min and max values in Figure 67).

The generally lower values for context/category usage during information search tasks are caused by the much shorter individual task times (less than 1 minute vs. 5 minutes).

6.6. Discussion

Main Study 1 was conducted to investigate *the effects of interaction-driven dynamic updating of contextual elements on task performance and user preference? (RQ 3)* for a Focus-Metaphor Interface (FMI). This study also investigated how user interaction behaviour differs using an FMI and its interaction-driven dynamic updating of contextual elements versus using a traditional blog interface (BlogUI). As part of RQ 3, two concrete hypotheses were tested (see section 6.2).

The first hypothesis, that participants using the FMI for information search tasks will make fewer errors and complete tasks faster (than participants using the baseline BlogUI) was supported (see section 0). Using the FMI, five of the six standard tasks were completed successfully by all 30 participants. Using the FMI, only task 6 was not completed successfully by 2 participants, making 2 errors collectively. Using the BlogUI, five of the six tasks were not completed successfully by all participants. 9 of 30 participants using the BlogUI did not complete all tasks successfully, making 10 errors collectively. Participants using the FMI also needed significantly less time that participants using the BlogUI to successfully complete their tasks (see section 6.5.1.2).

The higher average fixation durations measured for participants using the BlogUI (see section 6.5.4.2) suggest that the FMI may be generally easier to use than the BlogUI as increased average fixation durations are commonly linked to increased levels of cognitive load. The found difference in average fixation durations might thus partly explain the strong advantage found for the FMI in the task performance analysis.

The second hypothesis, that participants using the FMI for both task types (information search and information exploration) will rate the FMI as more usable in terms of usability criteria such as *ease of use*, *learnability*, and *productivity* (than participants using the BlogUI will rate the BlogUI) was not generally supported (see section 0). Overall, both BlogUI and FMI were rated similarly high (4.8 vs. 4.7 on a 1-6 likert scale), and no significant difference in user preference could be found between both UIs. However, the subjective feedback collected in the form of most positive and most negative aspects for the tested UI clearly revealed contrasting perceptions about the BlogUI and the FMI. As a result, a more detailed analysis of user satisfaction based on computer expertise was conducted (see section 6.5.2.3). This analysis compared novice and expert participants' perceptions of both UIs and found some indication that novice participants rated the FMI as more usable than the BlogUI, whereas expert participants rated the BlogUI more usable than the FMI, although differences are not statistically significant.

6.6.1. Observations on User Interaction Behaviour

This study explored differences in user interaction behaviour between FMI and BlogUI. While some differences in the usage of search and other means of interaction specific to each UI were observed (see section 6.5.3), letting the numbers speak for themselves would not be a sufficiently meaningful judgment as to what differentiates the information experiences of using an FMI versus using a 'traditional' blog interface (BlogUI) for information search and information exploration tasks. Instead, a more high-level discussion of some of the conceptual differences between the FMI and BlogUI is much more important. In light of the results of the performance and user feedback analyses, such a discussion appears both relevant and appropriate. The minimal use of search in the information exploration scenarios supports the well-documented desire of users to forage in small steps by *orienteering* (Teevan, Alvarado, Ackerman, *et al.*, 2004). In the blog interface, this left users with the ability to select categories and scroll through lists of articles. In the FMI, users are required to select a contextual element to move from one article to the next.

While this seems like a subtle difference at first glance, the underlying interaction philosophy is very different. The fact that participants using the blog interface only selected 2.3 categories on average in a 5 minute period of scenario-based information exploration underscores that scrolling was people's primary type of interaction with a series of articles. The qualitative difference between those two types of interactions is that selecting a category is demanding a much more ‘active’ decision on the participant’s part, than scrolling through a list of articles. Moreover, a category label, no matter how well-assigned is still a stark abstraction of the underlying content. This suggests that there is more uncertainty as to the potential value ‘behind’ a certain label (meaning low information scent). Particularly with an increasing number of category labels, the perceived difficulty of making a good decision increases (see discussion in sections 2.2.5, 2.2.6, and 2.2.7). As a result, participants opted for a more ‘passive decision’ of simply scrolling article by article, which represents a decision of ‘marginal effort’. In contrast to the Blog interface, participants interacting with the FMI had only one type of decision to make - which of the context snippets to navigate to next. By design, this required participants to make active decisions. But the specific design of the FMI aimed to make such decision making as effective as possible:

1. The number of contextual navigation elements is limited to seven to minimise the likelihood choice-overload (see section 2.2.7).

2. Each contextual navigation element provided a title and brief snippet to allow for more informed decision making

3. The consistent and predictable layout of contextual navigation elements facilitated effective visual scanning, by displaying all available information ‘above the fold’.

Results of the eye tracking analysis clearly illustrate the contrasting interaction philosophies between blog interface and FMI. During information exploration tasks, participants spent only 7.7% of their time on the category-based navigation for the blog interface, versus 26.9% of time on contextual navigation elements for the FMI. While some navigational decision making within the blog interface may be attributed to skim-reading behaviour on the actual blog posts (and thus accounted for in the content part of the eye tracking analysis), the quality of decision making for such scroll-based skim-reading behaviour is impacted in part by the passivity described above. Skim-reading by scrolling through articles one-by-one does not allow for easy comparison of a number of articles, or even article previews. Moreover, the cost of going back to a previously skimmed-through article is increasing with the number of articles scrolled past since. This cost is not just related to the effort of scrolling back to a previously skimmed article, but also the potential cost of scrolling back down to the correct position that one left off after deciding to revisit said previously skimmed article.

Such assessment of decision costs and their impact on interaction behaviour is related to the information scent model by (Pirolli, Card & Van Der Wege, 2000; Chi, Pirolli, Chen, *et al.*, 2001; Chi, Hong, Heiser, *et al.*, 2007). In other words, the layout of information and interaction philosophy underpinning the FMI increases information scent compared to the traditional blog interface. Arguably, not all information exploration behaviour is explicitly goal-driven. Some users may prefer the passivity of browsing content through scrolling as it does not require active decision making. However, such behaviour typically results in users simply consuming 'the next best thing' rather than focusing more strongly on the quality or relevance of the content they attend to.

The eye tracking analysis further underlines the stark contrast in interaction effectiveness between both interfaces for information search tasks. As both interfaces are laid out and function substantially differently interaction behaviour in the less goal driven exploration tasks could be interpreted as baseline behaviour. While relative gaze time spent on the navigation goes up by approximately 130% (from 26.9% to 61.5%) when using the FMI for search tasks, gaze time spent on the navigation goes up by close to 390% (from 7.7% to 30%) when using the blog interface for search tasks. On their own, those increases could be seen as meaningless. Simply comparing them directly could even be misleading, as clearly, relatively less time is spend on navigation in the blog interface. Only by integrating the eye tracking analysis with the task performance analysis, does a clearer picture on the contrasting interaction philosophies emerge. Interacting with the blog interface (through some of the most common interaction paradigms on the web today), participants clearly struggle to find the correct articles (as expressed by the significant error rates and increased task times). However, interacting with the FMI, participants cope much more effectively with the information search tasks (as expressed by marginal error rates and decreased task times).

6.6.2. Effects of Dynamic Adaption of Context in the FMI

In the search condition, participants particularly used search interaction. Traditional search result pages do not allow the user to conveniently look at all returned results (Cutrell & Guan, 2007). In contrast, the layout of the FMI enabled participants to effectively scan and compare all returned search results and to efficiently identify the target article or refine their search strategy. Selecting one of the search results automatically adjusts the contextual navigation to provide choices that are more relevant to the selected article. To some extend, this approach of dynamic adaptation realises the promoted personalised re-ranking of search results, which Teevan et al. (Teevan, Alvarado, Ackerman, *et al.*, 2004) identify as beneficial to cope with users' individual information goals. However, mixed results for subjective evaluations suggest that the level of dynamic adaption (with each interaction) in the FMI might have gone one step too far. Users unfamiliar with this novel interaction strategy were partly overwhelmed by the constant adaptation of the contextual navigation. Particularly experienced participants who are very familiar with traditional web-usage might have experienced a lack of control over the navigation. The approach of immediate adaptation of contextual elements to the currently active article made it harder to explore a number of navigation options, e.g. after conducting a search and skimming over the returned search results. Consecutive iterations of FMI prototypes will have to investigate alternative adaptation techniques supporting its contextual navigation.

6.6.3. Effects of Participant Heterogeneity

Huge individual variances across measures such as task performance, interaction strategy, etc. have been found in this study. Some participants relied entirely on search, others entirely on using categories (or contextual navigation in FMI). Some participants interacted very frequently; others were much more passive and particularly in the BlogUI, would thoroughly go through article lists. The use of realistic scenarios and conduction of the study with participants genuinely interested in the domain used for the evaluation of the may have contributed to these individual factors to surface. Diversity of the participant pool (age range between 18 to 67 years; large differences in computer literacy) may have been another factor contributing the large individual differences in participants' performance and user feedback.

6.6.3.1. Other Observations regarding Participants

Participants were generally quite engaged in the given tasks, particularly in the information exploration condition, where they were free to read articles of their choosing. General comments on both conditions either favoured the exploration condition for its more in-depth information, or favoured the search condition for being easier and faster to complete. Participants contrasting views could result from different levels of engagement and interest in the topic. Overall, feedback suggests a slight preference for the UI used during information search tasks.

Interestingly, a number of participants were very positively surprised by the quality and relevancy of content to the subject matter. While this content was created editorially by selecting relevant articles from a number of life-style blogs and news sites, it is all readily available on the web. But as it is scattered over a dozen or so sites, participants' surprise over the quality of content simply underscores the argument that people do not naturally browse across a larger number of sites for such content, as the effort involved would be too high for the perceived benefit of increased quality content.

6.6.4. Limitations of Main Study 1

6.6.4.1. The Impact of the Nature of the Task

The particularly strong variances in the task performance for search task 3 indicate the impact that task design has on the actual results of such studies. Task 3 highlights participants' focus on headlines and image information. Many participants would not be able to find the right information, if they cannot deduct the relevance of an article quite easily from the article title or image information alone. Information that is not related to title, image or at least the first paragraph will only be discovered by experienced users or by chance. As reading blogs involves a lot of scrolling, the eye might coincidently fixate on the correct paragraph in a longer article. However, this strategy should not be regarded as reliable and satisfying as numerous participants were scrolling over the target paragraphs with fixations in close proximity to the "answer" without noticing it. Scrolling has often been equated with bad usability (Nielsen, 2005). Error rates and task performance times for Task 3 show the positive side of scrolling and how particularly experienced users mastered this interaction. Scrolling through long lists of text seems to allow experienced users to process large amounts of information in a very short time. Experienced users were thus able to find less prominently placed bits of relevant information more often and commonly faster than novice users.

6.6.4.2. Study Results are Hard to Generalise

Through the detailed analysis of both UI types, for two separate tasks, using performance, user feedback, eye-tracking and interaction measures, it has become obvious how particular design decisions can impact particular measures. More research is needed to test different traditional search interfaces (e.g. search interfaces, different blog, web UIs) and further iterations of FMI prototypes, to provide a more generalizable assessment of the FMI in comparison with existing user interfaces.

6.6.4.3. Lack of Real-world Context

As the analysis by computer expertise revealed, there were clear differences between experts' and novices' perceptions of each user interface. While novices seemed to prefer the FMI over the BlogUI, experts seemed to prefer the BlogUI over the FMI. In addition, some of the findings reported in the analysis of task performance by computer expertise (see section 6.5.5.1) suggests that experts did not benefit as much from the use of the FMI than novices did (compared to their performance using the BlogUI).

These observations lead to the assumption that the FMI is more beneficial to novice users than it is to experts. This could be related to the fact that experts have more strongly formed mental models of how to engage in tasks of information search or information exploration on the web, and that they require an interface to more strongly meet their expectations. However, this will always be a challenge when testing a new UI concept in the lab, in a somewhat abstract context.

It is thus desirable to evaluate the use of just-in-time concepts as utilized in the FMI prototypes in a more real-world context and in a more longitudinal evaluation. In particular, Main Study 2 (Chapter 7) has been conducted to address this issue, by evaluating the just-in-time concept as part of field studies with (expert) knowledge workers of an IT organisation over the course of several weeks.

7. Main Study 2: Study of a Knowledge Discovery Tool for Organisations

"The mere existence of knowledge somewhere in the organization is of little benefit; it becomes a valuable corporate asset only if it is accessible, and its value increases with the level of accessibility." (Davenport & Prusak, 2000)

7.1. Summary

This study examines the need for additional contextual information when reading or responding to email, and whether providing this information can reduce the sense of information overload. This study is based on a two-year process of iteratively designing, developing and testing a proactive knowledge discovery system (KnowDis) for large organisations. The study summarizes the outcomes of a research collaboration with CA Labs, of CA Technologies, which funded this 2 year process.

The contributions reported are primarily based two field studies of KnowDis, conducted in a large IT organisation. The studies demonstrate that the KnowDis prototype design did improve the user experience for participants overall by making work-related information search more efficient. However, while the KnowDis prototype design was useful for some knowledge workers and even integrated seamlessly into their day-to-day work, it appeared to be less useful and even distracting to others. This chapter also provides a characterization of knowledge workers' email usage behaviour during the field study, and how this behaviour differs for employees with varying degrees of workload.

7.2. Introduction

Most organisations heavily rely on email and instant messaging for asynchronous communications, and many employees organize their work around email and calendar tools (Inkpen, Whittaker, Czerwinski, Fernandez, et al., 2009; Whittaker, Bellotti, Gwizdka, 2006). Organisations also make use of collaboration and social networking tools for knowledge management and transfer (Leshed, Haber, Matthews & Lau, 2008; Millen & Fontaine, 2003).

Nonetheless, opportunities for collaboration are frequently missed and efforts duplicated, due to relevant information not being known by the appropriate individuals. Knowledge workers often need to interrupt their tasks to search for additional information, and then switch back and resume the task. These interruptions decrease productivity, and search often produces a large number of hits the user has to sift through, increasing the sense of information overload (Boardman & Sasse, 2004; Czerwinski, Horvitz & Wilhite, 2004).

Consequently, employees often do not look for supplementary information. They do not expect to find relevant information within corporate knowledge repositories, do not know how to look for information using corporate search tools and data collections, or do not even know enough to ask a colleague.

The overall goal of this research is to support knowledge workers by providing them proactively with information relevant to their current task. Since reading and responding to email is a substantial part of the daily work experience, this research focused on providing email users with information that is relevant to the email message they are reading, but that they might not be aware of. Contextually relevant knowledge is discovered automatically, in a proactive manner, from the web, intranet, or desktop, without requiring explicit attention or instruction from the user.

The risk of providing contextual information in this way is that it might distract users' attention from their focused activity, so information must be displayed in an unobtrusive way. At the same time, it should be easily noticed when the user requires it. In order to develop a system that meets these needs, and to assess the overall value of the system, this research combined an investigation of email and search work habits with the development of a prototype tool, called *KnowDis*.

7.2.1. PIM and Email Review

Knowledge is the key resource in today's knowledge society (Dubie, 2006). As Davenport & Prusak (Davenport & Prusak, 2000) have put it: "*a knowledge advantage is a sustainable advantage*". Knowledge management is not simply the amount or quality of information controlled by an organisation, but rather how employees can create, share, access, maintain and act upon that information (Levy, Hadar & Greenspan, 2010). Adding knowledge within an organisation through generation or acquisition is not enough: the knowledge must be discoverable, comprehensible, and transferable at the appropriate times within the immediate workflow context (Gupta, Sharda, Ducheneaut, *et al.*, 2006).

Unfortunately, the ideal is rarely met (Nielsen, 2007b). In a survey on Intranet usability, Nielsen found that *"poor search was the greatest single cause of reduced usability"* and *"accounted for an estimated 43% of the difference in employee productivity"* (Nielsen, 2002). Recent interviews on knowledge management indicated that 39% of the statements about knowledge creating, sharing, accessing, using, maintaining and infrastructure were concerned with access (Levy, Hadar & Greenspan, 2010; Outsell, 2005, Millen & Fontaine, 2003). The *"grasp of collective organizational knowledge"* decreases for companies larger than two to three hundred employees (Davenport & Prusak, 2000); global enterprises are fighting an uphill battle for effective knowledge management (Dubie, 2006). The result is limited attention, information overload, vague or ambiguous communication, and misplaced attention on less relevant issues. This in turn forces knowledge workers to constantly compromise and somehow simply *"muddle through"* (Hollnagel, 1992).

To alleviate the onslaught of information, a broad range of studies have explored ways to aid knowledge workers in large organisations. Those studies particularly focus on analysing new tools that facilitate searching for expertise or experts (Ehrlich, Lin & Griffiths-Fisher, 2007; Ehrlich & Shami, 2008), automate and share how-to knowledge (Leshed, Haber, Matthews, *et al.*, 2008), enable social bookmarking (Millen, Feinberg & Kerr, 2006), content aggregation (Brzozowski, 2009) or social networking (DiMicco, Millen, Geyer, *et al.*, 2008).

For knowledge workers in large organisations, the one salient aspect of their work is dealing with email – reading, replying, creating, organising, re-finding – typically for several hours per day, and interacting with dozens or even hundreds of people. Employees are flooded with email and other interruptions that distract their attention from important tasks (Whittaker, Bellotti & Gwizdka, 2006). Problems associated with email overload include stress, interruptions, lost productivity, and email obsession (Gupta, Sharda, Ducheneaut, *et al.*, 2006).

Setting aside time to develop personal information management (PIM) strategies through reflection (Boardman & Sasse, 2004) or well-defined training interventions

(Soucek & Moser, 2010) can reduce feelings of information overload. However, Gupta et al (Gupta, Sharda, Ducheneaut, *et al.*, 2006) argue that email overload is more dependent on the complexity of tasks linked to the processing of particular emails, rather than the number of emails being processed, and that a more task-centric approach to email management is needed. A number of researchers also (Gupta, Sharda, Ducheneaut, *et al.*, 2006; Whittaker, Bellotti & Gwizdka, 2006; Sproull & Kiesler, 1991) highlight the need for better (automatic) categorisation and prioritisation of email to cope with email overload.

7.3. Research Questions

(RQ 4) How do users respond to embedded proactive search in an email application?

A set of more specific questions have also been formulated to focus the investigation of the conducted field studies:

> (RQ 4.1) Do users find having a proactive search tool embedded in their email application useful or not? If not, why not? If yes, how is it useful and how does it integrate with their day-to-day work?
>
> (RQ 4.2) Do users find proactive search features distracting? If yes, how is it distracting? What can be done differently to make it less distracting?
>
> (RQ 4.3) Do users think their work-related tasks that depend upon information search become more efficient and more effective when proactive search tools are available?

7.4. Method

Over the course of two years, a proactive knowledge discovery system (KnowDis) for the enterprise was iteratively designed, developed and field-tested. The work involved was conducted in two phases (see Figure 68).

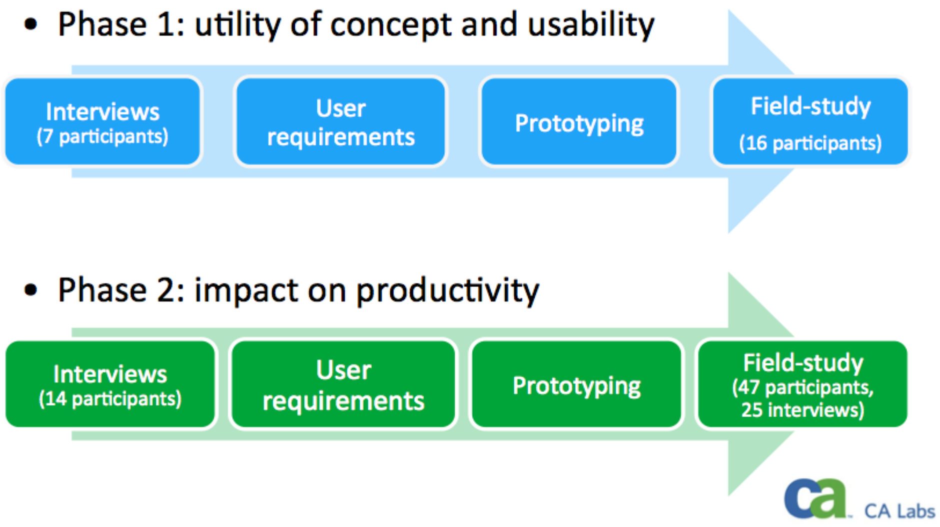


Figure 68. KnowDis work phases

Phase 1 started by collecting initial user requirements for such a tool through semi-structured interviews with seven employees of a large IT organisation (see section 7.5). Based on those findings, a first functioning prototype was built (see section 7.7) and pilot-tested with sixteen employees in the organisation (see section 7.8). The prototype was an add-in to Microsoft Outlook that extracted keywords from the active email and displayed links to relevant documents within corporate knowledge repositories. In field study one, participants used the prototype for some days (some participants used it considerably longer) in their day-to-day work. Consequently, de-briefing interviews were conducted with fourteen of the sixteen field study participants.

In **phase 2**, the feedback from the interviews of participants of field study one was analysed to refine user requirements (see section 7.8.4) and to inform the re-design of the prototype from phase 1. A second functioning prototype was then built (see section 7.9) and made available to knowledge workers within a large IT organisation as part of a seven week long field study two (see section 7.10). Employees wishing to integrate KnowDis into their day-to-day work had to fill in a detailed pre-study questionnaire before being provided with a user guide and video demos explaining the key features of KnowDis. And the end of field study two, participants were asked to fill in a post-study questionnaire and invited to take part in debriefing interviews. Twenty-five such interviews were then conducted with participants of field study two.

7.5. Initial User Requirements

Seven semi-structured interviews with employees of a large IT organisation were conducted to collect initial user requirements. Four of our interviewees worked in sales, two in development, and one in back-office support.

7.5.1. Interview Method

Although originally, initial user requirements for the project where meant to be collected as part of a focus group to be run at the organisation's EU headquarter, on the day of the focus group it was not possible to get everybody into a room at the same time. As a result, individual interviews were conducted instead. For these interviews, the detailed blue-print that was created to prepare and guide through the focus group was used instead to conduct the semi-structured interviews to understand how they manage corporate information in the context of their daily tasks.

7.5.2. About the Participants

P1 has been with the organisation for a long time, works in a back-office role in the services division of the organisation, managing implementers and selling consultancy. P1 receives "lots of data, little information" and uses a range of internal tools and Excel to analyse data, often doing work by experience. P2 works in pre-sales, dealing with customers via email and phone, as well as preparing and running customer presentations and demos. P3 works on an internal software product as senior software architect in a large team of 140-150 people, half of which are other developers. P3 liaises between product management and development to inform product strategy and direction. P4 joined the organisation just a year ago as pre-sales consultant, providing technical coverage to sales other people. P4's work requires in-depth knowledge about software specifications, software architecture, manuals and other documentation as well as detailed understanding of customer requirements. P5 works in sales operations, dealing with direct and indirect sales. Part of P5's responsibilities includes preparing weekly newsletters for global sales operations, web casts and to generally improve communication within the company. P6 also works in sales operations, dealing with direct and indirect sales. P6 prepares quarterly customer satisfaction surveys and regularly travels to sales meetings to meet with customers and collect their feedback. P7 rejoined the organisation a few months ago as client solution architect. P7 requires a lot of product information to familiarise with the relevant internal products, such as newsletters and online training.

7.5.3. Interview Structure

In the first part of the interview, we specifically asked about what types of information they needed for their daily tasks, and how they used tools such as email, corporate knowledge repositories, and collaboration tools such as wikis and SharePoint. In the second part of the interview, we introduced the concept of KnowDis as an augmentation to Outlook and asked about its potential relevance to the participant's email activities utilising a very early prototype. This was to gauge initial impressions on the initial design concept and to assess whether participants generally showed any empathy towards the proposed design. The following questions were used to guide the conversation of the first part of the interview rather than being checked off one by one.

1. How do you currently **use email**, and make use of CA **knowledge repositories**, either through search or other means? Do you have problems organizing your work?
2. How would you categorize the **amount of email** you receive? Do you have problems managing your email?
3. What **types of information** do you require for your everyday work? Do you have problems getting to that information?
4. Do you attend (internal) **seminars, workshops** or other work related **events**? Do you have problems finding out about those events?
5. How do you use tools, such as **SharePoint**, **Wikis** or other **collaborative work spaces**? Are there any problems with integrating those tools into your daily work flow?
6. How often are tasks such as reading email and finding information on a knowledge repository entwined in your daily work? Do you often interrupt email reading/responding to search for additional information that would help make sense of the email or is necessary for responding to the email? What sort of files or info do you search for? Do you search for that information on your PC, the CA Intranet, or the general Internet? Are you trying to clarify something, or is it related to a task generated by, or associated with, the email?

In the second part of the interviews, some more specific questions were used to test the concept of KnowDis and to better understand if and how participants used Outlook. These questions were contextualized by sharing a mock-up and example use case for KnowDis (see Figure 69 and Figure 70).

- Do you use Outlook?
- How do you read your email – in preview pane, or open each email in separate window?
- What screen setup do you usually work with?
- Do you work in the office, or remotely?
- If you work remotely, are you usually connected to CA via VPN?

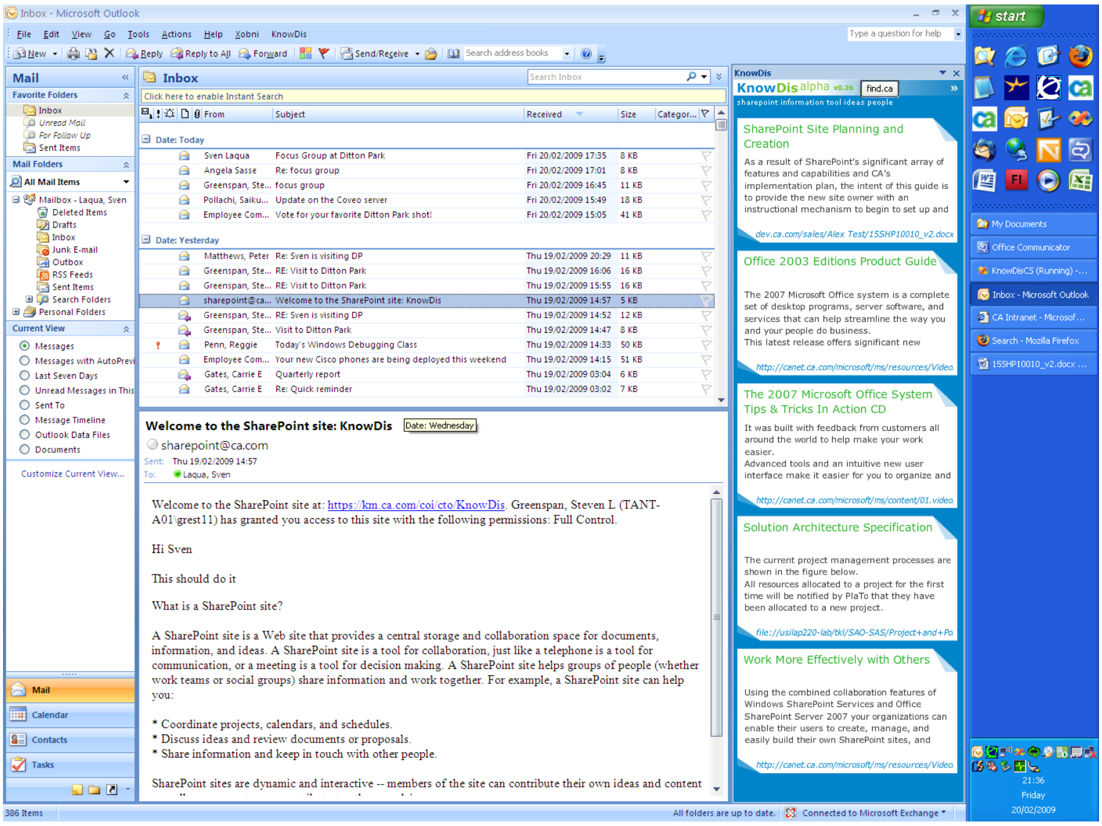


Figure 69. KnowDis concept mock-up

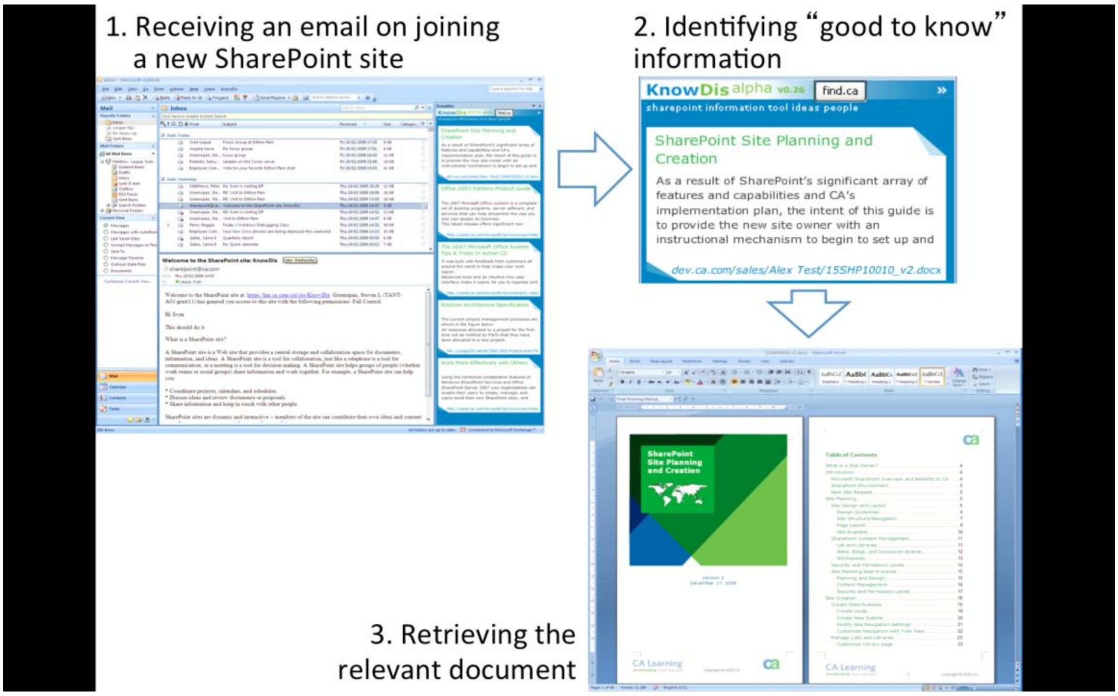


Figure 70. KnowDis use case example

7.5.4. Analysis Method

The interviews were then analysed thematically around the relevant themes such as use of email, use of internal knowledge repositories, retrieval of information related to work, general workflow and feedback about the concept of KnowDis. These themes are partly based on the guiding topics of the interview blue-print and partly emerged from participants' feedback.

7.5.5. Results

The following sections report on the feedback gathered during interviews, grouped by the themes deemed most relevant during analysis.

7.5.5.1. Email Usage

For all participants, email is central to their work. It is at the heart of their communication with the rest of the organisation, the exchange of information with colleagues and the portal to organise and coordinate work, as aptly described by one participant: "*Email is my everything. My Outlook. It's my to-do list, it organizes me 24/7*" (P5).

Participants broadly reported four general types of email they deal with - 1) direct communication with co-workers, 2) targeted, but larger audience ‘mass emails’, 3) general broad cast email such as corporate newsletters or distribution lists, and 4) external email with customers.

Email Overload

Most participants reported struggling with the level of email they receive. For some, it is common to receive between 100 and 300 emails per day (P5 and P6). For P2 around 20 to 30 emails accumulate overnight due to him working in a different time zone from his colleagues. He estimates receiving 99% of information relevant for his work via email and notes that he rarely has sufficient time to effectively prepare for customer meetings: “*A lot of luck is needed for things to go right, if there is not much preparation time. I get lucky a lot*” (P2). This participant seemed particular daunted by the amount of email he receives and the consequent information retrieval workload these email generate by stating that “*you will never find what you want, when you want it. It’s a fact of life, accept it*” (P2).

P4 in particular struggles with email containing too much information, such as links to 5 or even 10 articles. And P3 directly acknowledges that he cannot keep up or even answer all the email he receives. One of the specific problems with email aiding email overload is the inefficiency of the medium to effectively communicate with colleagues, often resulting in email chains (P1). Only one participant, who very recently re-joined the organisation reported email usage as “*not so bad yet*” (P7).

Information-sharing via Email

Central to the use of email is the sharing of information. For some, the frequent sharing of attachments is a burden. P2 estimates that 50% of his emails contain spreadsheets, screenshots with problem descriptions, or other attachments. While he is using a blackberry for email, attachments require him to defer dealing with those attachments until he is in the office. One particular problem with attachments as noted by P4 is that documents shared may have file names that do not always match the actual content, making it much harder to later identify the correct document in ‘here is that document’-type emails.

For P3, receiving lots of email with questions is particularly burdensome. Questions range from software-specific code-level type questions to generally strategy type questions, leaving him with the impression that "*people think you know everything*" (P3) and when he does not know the answer, it "*makes you feel stupid*" (P3). This participant estimates that roughly a quarter of such emails are related to general information - e.g. who is in charge of that aspect of the product - that the email sender should be able find out on their own. Newsletters are another source of information received via email that participants need to deal with and that participants may sign up for by change when trying to download a white paper or other work-relevant documents (P4).

P4 further notes that colleagues might share links to information in their email that then do not work - potentially due to access restrictions within SharePoint or other knowledge repositories. He would thus prefer to receive the actual documents, rather than links.

<u>Strategies to cope with Email</u>

Participants utilise a range of strategies to cope with the amount of email they receive. For example, P2 uses his blackberry to deal with email on the way to work as much as possible.

P3 lets email threads 'bubble along' hoping someone else might spend the time to answer a question that has been emailed to a group of people. However, he acknowledges that this strategy is dependent on the importance of the email and the person who sent it. Due to limitations of mailbox sizes in the organisation, participants regularly have to archive email in separate files. While some archive email opportunistically in case it might be useful later on (P4), others frequently run out of space, and then struggle finding their archived email (P5 and P6). P7 categories email archives as much as possible to be better able to re-find relevant emails.

7.5.5.2. Use on Organisational Repositories

In 2002, Nielsen (Nielsen, 2002) reported that company intranets that are not well designed significantly reduce productivity. Our interviews indicate that this has not changed much since - "*you know it's there, but cannot find it*" (P5). Interviewees reported struggling with the structure of knowledge repositories in the organisation - identifying which information is useful is the biggest problem. One participant illustrates the problem, saying "*people want a path [to information] to get their thoughts right*" (P1) and that employees cannot find the information using search "*because they don't get their plan right*" (P1). The intranet within the organisation is used for general internal news, for online training, such as on ethics, strategy, or for all types of information on internal products. It used extensively to (try to) learn about other internal projects, teams or initiatives. SharePoint is a central part of the intranet, alongside a range of other knowledge repositories as well as a range of internal search engines.

P1 further points to problems with access restrictions, which prevent people from finding useful information and suggests that there is are regional differences in the effectiveness of using SharePoint across the organisation - people in the US being more confident and people in the UK being more confused by SharePoint. P1 even argues that SharePoints pretends to aid collaboration. Similarly P4 acknowledges regional differences related to internal data sharing, suggesting that "*in local [UK] community, people share information*", but it is harder to understand what information is available in the US or Asia. This problem may be related to localisation of intranet resources. P4 points to another problem with SharePoint, in that folder structures may be hard to understand, if they are maintained by another team and no common hierarchies are defined. Some, like P5 and P6 have to use SharePoint a lot, but simply do not like to use it.

P3 argues that one of the underlying problems with internal information discovery is caused by poor data quality, a lack of meta data or otherwise more intelligent search. As an example he uses changes in product or project names. As a result, old documentation may not be found or even if it is found, it may not be perceived as relevant. Another similar problem is related to the ambiguity of versions of documents. P3 recalls searching for an official text snippet to be used to list 'company name and copyright information' to be embedded in products built by the organisation. However, he only managed to find an older version from 2006 but could not find the up to date version, nor could he identify, which department or individual would be responsible for this information.

P2 heavily depends on distribution lists to find "*a lot of useful information*" as "*[Internal] search never brings up the right documents*". One potential reason for this is the problem of changing code names of projects, which make it harder to track down relevant project information on the intranet. P1 suggests that as a result, people miss actually relevant information as they might not be familiar with old code names.

Another problem with intranet pages as reported by P4 is related to the 'perceived relevance' of information as per the font sizes used. The most relevant information may be in 'small print' whereas the general layout of information may try to guide the user elsewhere.

Enterprise Search

The ineffectiveness of the organisation's intranet search is a central aspect of the perceived limited usefulness of the intranet as a whole. Participants generally lack trust in the accuracy of search results. Search results are either not specific enough (P7), not relevant (P6) or wrong (P5) - the latter caused by wrong localisation, such as 'dental plan's for the US, whereas the employee might be in the UK.

While the organisation has recently upgraded its enterprise search technology, the feedback from participants is still mixed. Some find the new system "*really good*" (P3), others "*do not like it very much*" (P4) or simply prefer not to use it (P5). Some of the problems participants experience may be due to a lack of experience with the new system, which provides a range of advanced filtering options that participants either did not notice (P6) or did not use as it is hard to figure out how they would help (P4). One issue closely related the lack of trust into internal search is reported uncertainty about the scope of resources actually being indexed by the enterprise search engine. As a result, when information cannot be found that should be there, it is not clear whether a refinement of search parameters might eventually return the designed information or whether this part of the intranet simply is not indexed, thus continuously eroding trust into the enterprise search.

A related problem is the perceived complexity of tools such as enterprise search. It takes time to learn how to use the various filters effectively, which users may either not have or are not willing to spend.

Lack of sufficient sharing of relevant information is another problem impacting the usefulness of intranet search. P1 suggests that other people within the organisation do not necessarily share information or useful tools, as they may not receive a direct benefit from it or fail to see how it could benefit their colleagues. P5 further notes that some colleagues are very territorial. For example, people in sales may not want to share their contacts, but this affects her ability to use salesforce effectively, e.g. To allocate accounts.

One final problem intranet search generally faces is the benchmark it is typically compared to: "*Google is very good*" (P2). Google is used in a personal and professional context by all participants and its simplicity and effectiveness shape their expectations of what intranet search should be like.

7.5.5.3. General Workflow and Information Retrieval

Information retrieval strategies - Orienteering, Teleporting & Serendipity

P2 does not find internal search useful. He tries to avoid using it and prefers to use browse repositories directly, such as particular SharePoint sites, or wikis. P2 describes his strategy as "*making a mental tick*" when coming across information that might be useful in the future. However, he notes that when he forgets to make a mental tick, re-finding information works very badly.

P3 has two specific methods to "*get an answer*". He either asks someone who should know the answer, or uses internal web sites (wikis or SharePoint sites) to find the answer. However, P3 also acknowledges that generally "*it's not easy at all to find the information*". P5 and P6 even suggest that searching for information upsets them as they may spend 20 to 30 minutes to look for some specific information, thus putting them in a bad mood before starting the next task.

P1 reports that people within the organisation may have a "*power user*" which they habitually turn to for help and thus become reliant on those other people. For P4, this would mean to send an email if the information need is not urgent, and if it is urgent call or go by the office of his power user(s). P4 also acknowledged to preferably ask a colleague rather than searching for information himself just in case they might know, as even if they likely would not know, it is still quick and easy to just ask first.

On the flip side, P5 and P6 acknowledged that people regularly contact them for information. While both generally did not mind as they felt it is part of their job responsibilities, they did note that some colleagues repeatedly contact them about the same problem.

P2 indirectly confirms the comments by P5 and P6, suggesting that "*technical people are the worst to help you*" as they might reply to an information request by saying "*we sent you that document 3 weeks ago*" (P2).

P1 provides anecdotal evidence of serendipitous encounters with useful information during a phone conversation with a colleague in which they discussed data both of them needed to analyse separately for their managers. His colleague told him of a tool that would vastly speed up the specific analysis he needed to conduct and which had been created by a colleague. The tool extracts project status data and imports into a spreadsheet to enable quick analysis of update problems. P1 did not even consider the possibility that such a tool might exist and thus did not even look for it. P1 further noted, that even if he did look for such a tool internally, it would have been very hard to find as it would either not be shared at all on the intranet or hidden on a particular SharePoint site.

Local Information

Another problem noted by P4 is that of repeatedly searching for information that has already been stored locally. One possible explanation is that the person has forgotten about acquiring the particular information previously. Other participants reported struggling to re-find relevant locally stored information (P1 and P3). A number of participants (P1, P3, P5 and P6) have thus started to use Google desktop search, which has reportedly solved the problem of re-finding local information for them - "*works brilliant for me*" (P6).

Problems with On-boarding New Hires

P1 points to a potential cause contributing to some of the problems outlined above - such as ability to use internal tools effectively or store and share information consistently. For new hires, there is no clear coordinated introduction. While some people may get two weeks of detailed training, others may get half a day training after being with the organisation for three months already. Such inconsistencies across different parts of the organisation may be related to the overall size of the organisation, as well as the fact that acquisitions of new companies create a less homogenous environment. As a result, once new hires get to the point where they are able to make more sense of internal tools and processes, things may have changed already, with new policies in place, or new tools being adopted.

P4 further acknowledged that a year after joining "*it got much easier*" to deal with information problems, because he might know the answer already, know where to find it, or because he now knows people he can ask.

7.5.5.4. Concept-testing Feedback

The last part of the interview gathered concrete feedback on the concept of KnowDis as an add-in to Outlook.

Empathy Feedback

Participants generally expressed interest in the concept underlying KnowDis - "*I think that's a very good idea*" (P2), "*it's an interesting functionality to have, absolutely*" (P5). One participant explicitly suggested that he "*would try it*" (P3) when it is ready, while another did not even want to wait "*I would rather want to have such a tool sooner than later. Even if it did not work perfectly, give me version 1 now, and make an improved version 2 available later*" (P1).

Some of the potential benefits a tool like KnowDis might provide were related to saving time searching for information (P5 and P6), getting access to a relevant document without explicitly having to get a colleague to share it (P2), as well as possibly receiving more accurate recommendations of relevant files due to the fact that the entire email content is used as context for search rather than just a few search terms (P3). However, P7 noted that such functionality may only be useful for some email but not other 'basic' email, while P3 flagged that if recommendations are not relevant too often, users might start to ignore it.

Outlook integration

Participants acknowledged that the planned integration of KnowDis into Outlook would be the right approach as "*it's the main portal for users*" (P1) and as "*everything goes through email*" (P3).

Comparisons

P2 noted that KnowDis "*reminds me of Google Mail*" as adverts shown alongside email also parse email content to provide related recommendations, whereas another participant likened the recommendations provided by KnowDis to be "*TIVO-like*" (P1).

User Interface

Participants generally seemed to warm to the sidebar concept, but some noticed that KnowDis "*uses quite a bit of screen real-estate*" (P3), which may be acceptable "*if it works well*" (P3), but others felt that KnowDis "*should occupy less space of Outlook*

UI" (P1). One participant expressed a preference for being able to manually close and open the sidebar as for some email, he might not want to use the tool (P4).

Feature Suggestions

Participants proposed a range of potential features, which further suggests a certain level of empathy towards the concept of KnowDis. P3 and P4 explicitly wanted to see the integration of local search functionality as many colleagues in the organisation use Google Desktop which reportedly works well for them. P3 specifically brought up issues with re-finding particular emails that KnowDis might be able to help with. P3 and P4 further wanted to see the integration of filtering options, such as by file type. P6 and P7 raised the idea of using KnowDis as a reminder, which would re-surface certain emails within the sidebar after a few days when the information becomes relevant. P3 and P7 wanted to see more control over what parts of an email message are being used to provide recommendations. P7 specifically asked about the ability to recognise links and provide more detailed information on those links, thus helping him make a decision about the relevance of links and prioritise, for example when emails contain a lot of links. P3 wanted to be able to highlight any part within an email message and to only make KnowDis search for related information based on that selection.

Other ideas included a thumbs-up / thumbs-down voting mechanism to provide feedback on the usefulness of recommended items (P1), as well as the ability to provide recommendations whilst writing an email (P1and P3). The latter may be of particular interest, if the email being written is to request information that may consequently be surfaced even prior to actually sending the email.

7.5.6. Resulting User Requirements

The interviews revealed that employees develop their understanding of how the company operates through exchanges – by email, phone or face-to-face - with their colleagues. This social network is a crucial backbone for employees with information needs. Employees try to find their way around the company's knowledge repositories - spread across internal sites, databases, Wikis, SharePoint. But they often end up asking members of their social network where to find the information. Most admitted to habitually turning to the social network as a first resort (see section 7.5.5.3). Participant P1 argued to be highly dependent on help by a colleague he described as his "*power* *user*".

Participants acknowledged that integrating a tool which provides work-related information into Outlook was a sensible choice, as a significant part of their work day revolves around email. The amount of email most participants reported receiving and the amount of additional information retrieval tasks these email generate clearly suggest that a mechanism that makes these tasks more efficient and effective would be perceived as useful (see section 7.5.5.1).

Participants understood the concept behind KnowDis, and its place alongside their work email (see section 7.5.5.4). They expressed clear empathy towards to KnowDis, its integration into Outlook, and recognised the potential benefits: "*save time, rather than having to search for 20-30 minutes*" (P6). Feedback from participants such as P1 and P4 further indicates that the on-boarding process for new employees can be daunting and time consuming and a particular area in which KnowDis' potential benefits may be particularly relevant (see section 7.5.5.3). The need for a tool such as KnowDis was further underlined by the fact that participants reportedly struggle to find relevant information in internal knowledge repositories due to bad search, poor data quality, access restrictions or other issues (see section 7.5.5.2). Participants further expressed the need for some flexibility in terms of how KnowDis integrates into Outlook, as to not occupy too much space, or even get out of the way completely when not needed (see section 7.5.5.4 - user interface). A range of additional feature suggestions was raised by participants (see section 7.5.5.4), most of which represent rather nice-to-have type features, which would not be likely to make it into a first proof-of-concept type prototype (see section 7.5.5.4 - feature suggestions).

7.5.7. Method for Persona Generation

Two persona profiles emerged from the detailed topical analysis of interviews with knowledge workers within the organisation - *the information seeker* and *the information guide*. These personas are not meant to identify distinct types of knowledge workers as such, but rather particular work modes or implicit roles that are taken on by these knowledge workers to a lesser or greater degree. While an information seeker persona would spend most of her time in the mode described by that persona, an information guide persona would spend a significant amount of time in that mode. The information seeker is a well-known persona that is at the core of most information retrieval research - commonly faceted further to describe information seeking behavior in more detail. In contrast, the information guide persona seems to a be a much less recognised yet fundamental part of information retrieval processes. This second persona emerged from the analysis of primarily two facets of what participants were reporting during interviews. First, participants talked about ‘answering’ questions of colleagues. This alone does not appear particularly noteworthy as in itself answering a question someone else might have is a fairly typical activity. The second observation was embedded in the strategies participants described when trying to find some relevant information (being in information seeker mode). These information seekers utilise their social network strategically to offload information retrieval work, and increasingly so when they feel too busy to cope with all relevant information retrieval tasks.

7.5.8. Description of Personas

During the initial interviews with knowledge workers in the UK, two competing persona profiles emerged in the context of how participants experienced tasks related to knowledge management. The design of KnowDis aims to benefit both persona profiles as outlined in the following two sections.

7.5.8.1. Supporting Information Seekers

The first persona profile - the **information seeker** - can be interpreted as the ‘default end user’ for which nearly all information retrieval systems are designed. These systems are then evaluated using common information search tasks (such as navigation tasks), modeled around the needs of the information seeker persona.

But the interviews exposed some specific characteristics of information seekers in a corporate environment, which provide further context on how information needs arise, and are dealt with. Knowledge workers engaging in information seeking behavior commonly follow an ‘escalation strategy’ to complete their informational tasks. A typical escalation strategy looks like this:

1. Receive an information request by email.
2. Recall relevant information from own memory.
3. If unsuccessful, search for relevant information on work PC using desktop search.
4. If unsuccessful, search for relevant information on the Web (e.g. using Google).

5. If unsuccessful, ask colleague(s).

When receiving an information request (typically via email), or facing an information task in some other shape or form, the level of urgency (time available), priority (seniority of stakeholders), and other factors, influence the path taken and the amount of time spent on it. But due to a number of factors, information seekers regularly struggle to find the information that they need. Those factors included simple things like time pressures, which often do not allow for enough time to engage in a 'thorough' information retrieval process. In addition, the actual and perceived quality of the search engine plays a vital part in knowledge workers' trust in their own abilities to find the relevant information. This has been of particular relevance when the organisation's enterprise search engine needed to be used to gather information from corporate knowledge repositories.

One of the consequences of the uncertainty employees feel when performing information retrieval tasks is that they routinely reach out to other humans for help. In the first instance, they draw on their social network in the workplace to shortcut the escalation strategies listed above. Another problem is often the sheer number of tasks and the amount of time people actually have to engage in information retrieval processes, which leads information seekers to seek out help and delegate the information task to their social network. The mere fact that people like to tend to their social network for answers to more complex information tasks is not new, and a variety of enterprise social networks or collaboration platforms (e.g. Salesforce Chatter, SAP Streamwork, IBM Connections, Microsoft SharePoint, etc.) underline the value of tapping into colleagues' knowledge bases.

However, looking at the steps of the typical escalation strategy, such tools commonly target one of two improvements. At best, enterprise social networks will eliminate step 5, as relevant information may be found on a colleague's profile or data repository page of the enterprise social network. More likely, such tools will help identifying a promising individual among co-workers thus optimising step 5, and potentially speeding up step 4 and/or the shift from step 4 to step 5.

KnowDis' key difference to those systems stems from the pro-active nature in which it delivers potentially relevant information. By augmenting the actual email interface with related information, KnowDis attempts to intersect the knowledge worker's escalation strategy before step 3 and provide a short-cut to relevant information that would typically be found in step 3, step 4, or step 5. It could even be argued that the display of information related to an email can aid step 2 - recall from memory - as such additional information may just be what is needed , e.g. to remember the name of related project or product.

7.5.8.2. Supporting Information Guides

The second persona profile - the **information guide** - suggests that proactive augmentation of the email display is of particular relevance. Information guides are in some ways the counter-persona to the information seeker. They represent the kind of person in an organisation that is being swamped by information requests from information seekers. While asking a colleague for help first may seem to be common and acceptable behaviour, from an individual's point of view, if you are looking at the same behaviour from an organisational perspective, the information need has at best been migrated from one employee to another, but more likely than not, It has been multiplied, for example by emailing to a group of employees. At the very least, each of those employees – identified as information guide – now has to actively make a decision on whether or not to deal with this request or not, thus interrupting their own work flow.

On its own, each individual information request may seem like a very trivial act of interacting with a peer in one's social network. But as our interviews highlight, the 'abuse' of delegating information needs as a shortcut to the information seeker's escalation strategy, creates a distinct persona, which is burdened by the added work load: "*email is like a river, and some twigs pass you by*" and perceived work pressure: "*people think you know everything*", and not knowing the answer "*makes you feel stupid*". Asking colleagues for information (or more precisely for help to close a knowledge gap) is, of course, not a bad thing. In a productive work environment, colleagues should learn from each other, and communication is a catalyst for serendipity. However, when analysing the well-known concept of 'water cooler conversations', we find that its concept is based around two parties willingly engaging in a conversation. Both have taken a break from work and made themselves available to a conversation. To that end, information requests via email are more often than not an intrusion into existing work flows and ongoing tasks. As discussed in section 2.1.3, email overload is one of the major productivity killers and significantly contributes to information overload within organisations costing the US alone an estimated $997 billion in lost productivity (Spira, 2011).

7.6. Design Decisions

The initial interviews with knowledge workers in the UK explored the nature of their work and how information related to their daily tasks is managed. In those interviews, it became clear that Microsoft Outlook (as the corporate email program used in the organisation) acts as the central hub for communication and coordination throughout a typical workday and that a significant number of information-related tasks are ‘driven’ by email requests.

Those insights informed two key design decisions for the prototype system to be developed. First, potentially useful information should be provided directly in context to email being read. Second, in order to maximise ‘usage’ of the prototype system, it would be embedded into Outlook. This second technical decision was driven by the fact that Microsoft Outlook was the sole corporate email client being used across the organisation. While mobile email usage through smart phones was existent, it was marginal and thus excluded from design considerations.

An alternative design that was considered early on was a stand-alone prototype application (e.g. in the shape of a messenger application or similar to Google Desktop’s sidebar) that would offer the same functionality, but be application independent. Such a stand-alone version would have taken the project in a more or less different direction. Such an approach would have allowed for more freedom in designing the interface. Potentially, it would further have facilitated a different usage dynamic, as users could use a stand-alone application side-by-side to their email application, or in a more independent manner (e.g. when browsing the web, or editing a word document). However, a key argument for developing a prototype system to be embedded into Outlook was increased awareness and accessibility. The assumption made was that as knowledge workers would tend to Outlook throughout their work day, the likelihood of them also engaging with the embedded prototype system would be greater than for a standalone application, which may be opened up in the morning, but then forgotten about. This aspect of assumed greater exposure to and interaction with the prototype system was a driving factor in the chosen design for two reasons. First, it made sense in the context of the type of application to be developed, as it would simply make the delivery of potentially relevant more reliable. Second, it made sense in the context of the field studies to be conducted as any potential impact of the prototype system would have to be measured in real-world situations, while people are busy with their daily routines and within a reasonable time frame in which such studies would be conducted.

A separate application would have required more active decision making on the user's behalf only to attend to it for potentially useful information (if said application was not carefully lined up visually side by side other productivity applications (e.g. Outlook) being used. In addition, to go beyond the provision of contextual information for a single application, the implementation effort would simply have been out of scope for this project.

7.7. Description of First Prototype

A detailed user study guide was created for participants of the first field study. This study guide provides some general context about the project, a usage scenario for how KnowDis might help discovering "good to know" information, and a detailed description of the prototype (see Figure 71).

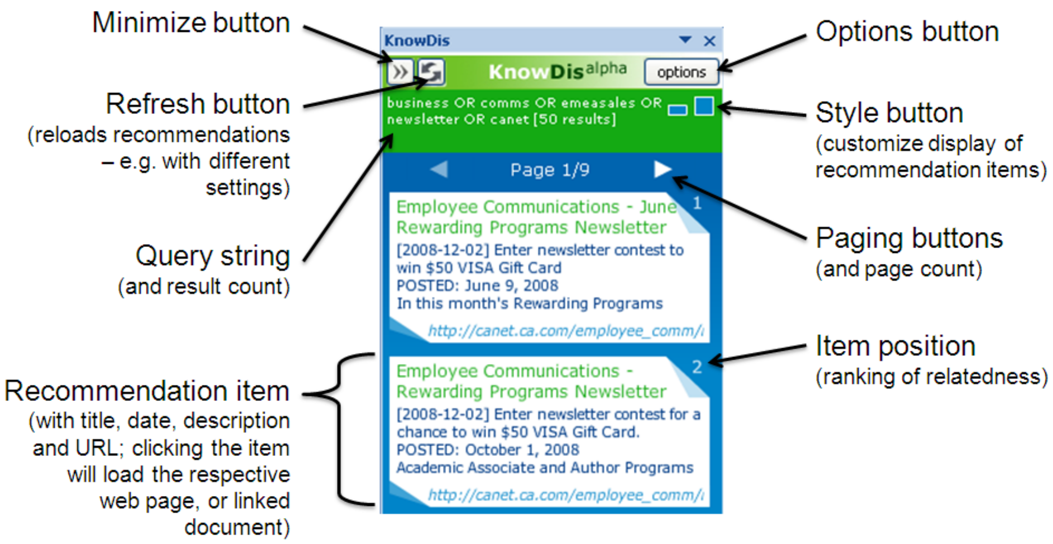


Figure 71. User Interface of the First KnowDis Prototype

The prototype allows defining the collections to be included when looking for related resources (see Figure 72). The collection settings can be adjusted in the "collections" tab of the options menu. The selection of collections corresponds with the collections available on find.ca.com:

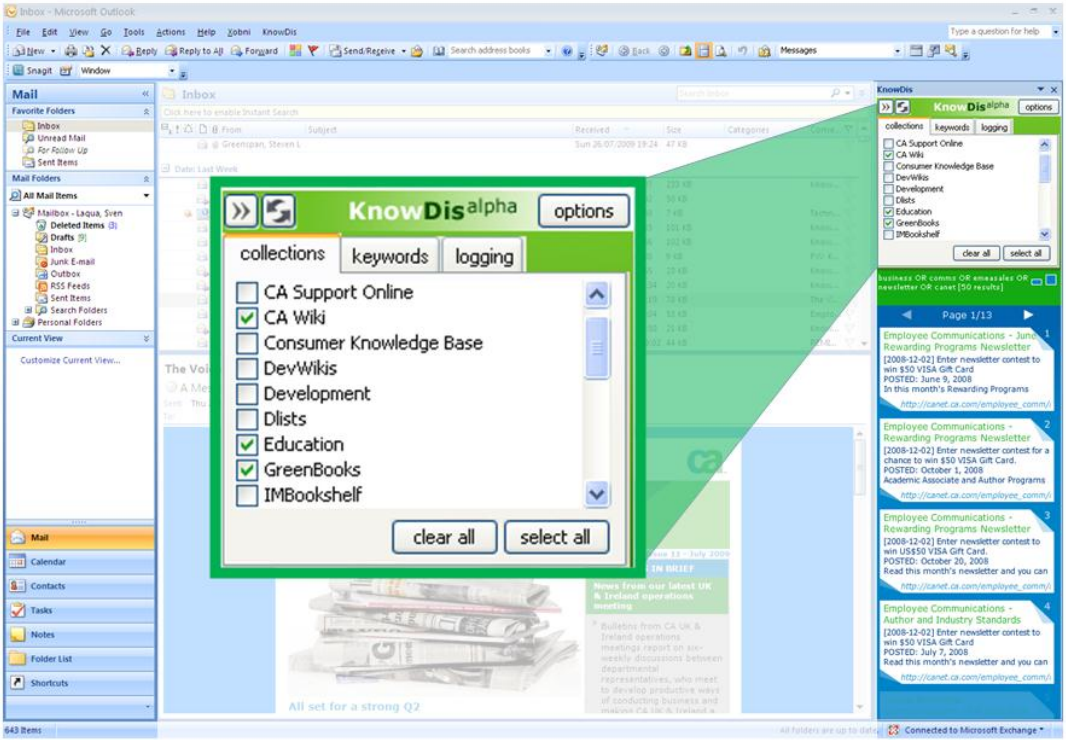


Figure 72. Collections Filtering Functionality

The prototype uses keywords generated from the selected email message to find related resources. Keyword generation in the prototype is handled by a custom-written function called ExtractKeywordsV3, which receives the raw email message text and the number of keywords to be generated as parameters. This function then runs through eight steps:

1. Removing line breaks
2. Removing URLs
3. Removing emails
4. Removing special characters
5. Removing multiple white spaces
6. Removing stop words
7. Filtering out names (utilising the user's contact list in Outlook)
8. Sorting and grouping remaining words

Finally, the algorithm extracts the desired number of keywords based on their frequency from the generated list and sends them to the UI. For the complete algorithm, see Appendix C, section 10.12.

Keywords generated from an email message can be adjusted in the "keywords" tab of the options menu (see Figure 73).

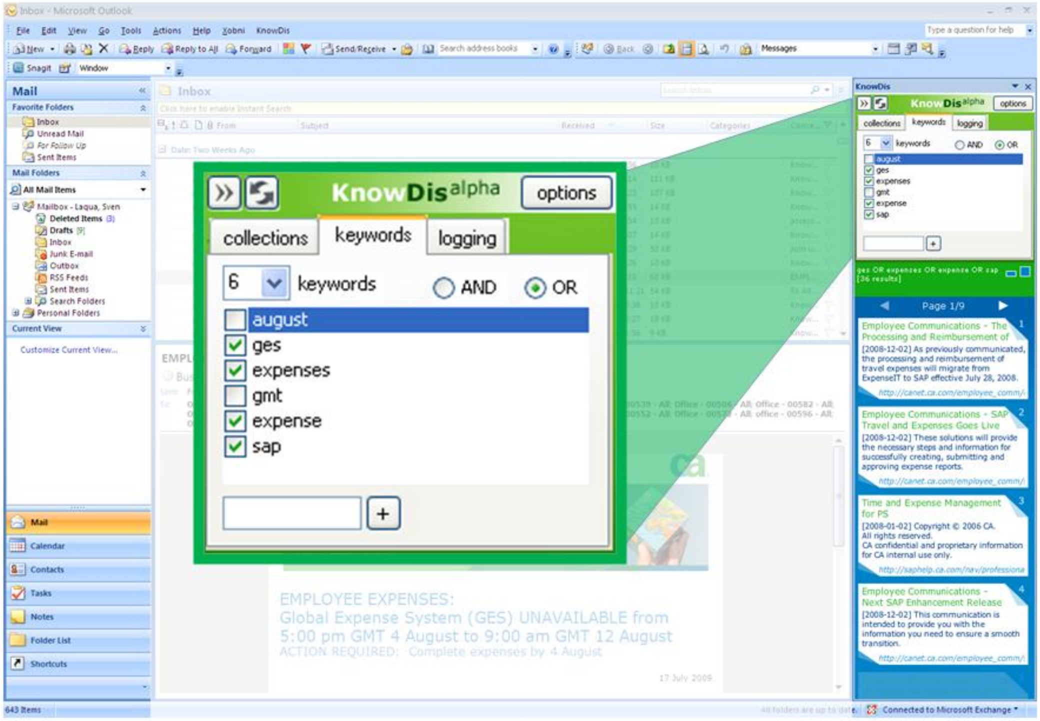


Figure 73. Keyword Settings

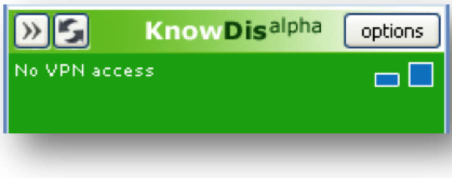
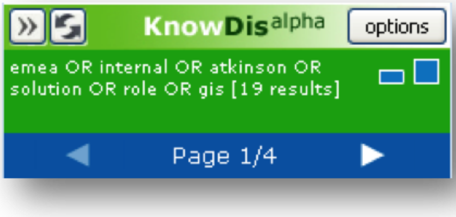
7.7.1. Installation Instructions

The user guide provided to participants contained the following installation instructions:

> *You will require Outlook 2007 to install and use the KnowDis tool. Any additional components that the tool requires will be automatically downloaded and installed during the install procedure (step 3).*
>
> 1. *Unzip the attachment "KnowDisinstall.zip" - This will create a "KnowDisinstall" folder, which contains all necessary documents.*
> 2. *Copy the file "knowDisinterfacev1.swf" to "C:\" - This file is part of the user interface of the KnowDis prototype. (Without this file located in C:\ the prototype will not work).*
> 3. *Install the Outlook add-on by executing "setup.exe" - The install process might require the download of .NET 3.5 SP1 and potentially other components, if they are not available on your system so far.*
> 4. *After successful completion of the installation start Outlook - The KnowDis sidebar should appear on the right side within Outlook.*
> 5. *Please read the user guide - It contains all important information about how to use the KnowDis tool effectively.*

7.7.1.1. Getting Started with KnowDis

The user guide also contained the following information about the prototype to help participants become more familiar with its functionality.

<table>
<tr><td></td><td>The KnowDis tool requires VPN access to retrieve recommendations from the CA Intranet. If VPN access is not available, the tool will stay inactive and inform you of the problem.</td></tr>
<tr><td></td><td>After the selection of a new mail item in Outlook, the KnowDis tool will be retrieving recommendations and indicate its busy-state by displaying a “loading…” message alongside the query that is being used to find suitable recommendations.</td></tr>
<tr><td></td><td>After the recommendations have been retrieved, the tool will indicate the number of recommendations found after the query string used and update the paging mechanism accordingly.</td></tr>
</table>

7.7.1.2. Understanding KnowDis Recommendations

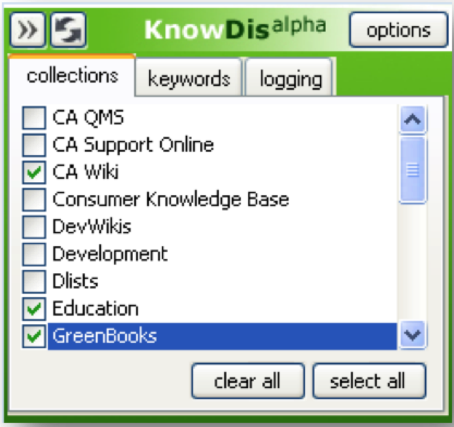
<table>
<tr><td></td><td>The display of recommendations offers 2 different display modes, which can be switched by using the style buttons.</td></tr>
<tr><td></td><td>The “standard mode” provides the user with a title, a description with a date, indicating the last time the information has been modified, and a URL.

Clicking either the title or URL will both load the underlying information (e.g. a web page, Office document, etc.).

The description is scrollable when the mouse is over the description text. This allows to access a bit more information about this item.</td></tr>
</table>

The “**mini mode**” of recommendations, provides the same information as the standard mode minus the description. This alternative allows displaying more choices concurrently at the expense of less detail (no description).

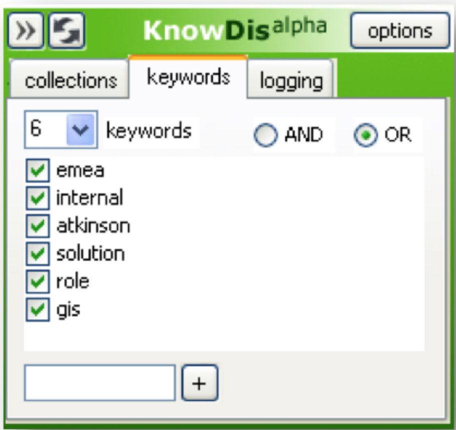
7.7.1.3. Adjusting KnowDis



The most important adjustment to the KnowDis tool is the (de-)selection of collections the user would like receive recommendations from.

Example:
If you are not a programmer/developer, you might not want any recommendations from collections that contain only such information.

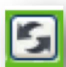
Those settings are persistent and will be stored as part of your user-settings.



Another way to adjust the recommendations you receive is by **controlling the keyword settings**.

The user can:
- adjust the number of keywords used,
- de-select keywords that should be excluded from a query,
- add keywords manually, that should be included, and
- switch between AND/OR Boolean operator to connect the keywords

Tip: the OR operator is recommended for larger amounts of keywords (e.g. 4 or more). The AND operator is recommended for smaller amounts of keywords (e.g. 1-3).



The **re-query button** allows initiating a query with the made adjustments for collections and/or keyword settings for the same mail item.

7.8. Evaluation I

The first functioning prototype, built based on the findings of the initial user requirements gathering, was pilot-tested with 16 employees in the organisation.

The prototype was an add-in to Microsoft Outlook that extracted keywords from the active email and displayed links to relevant documents within corporate knowledge repositories. After participants used the prototype in their day-to-day work for some days (some participants used it considerably longer), we conducted a further fourteen semi-structured interviews.

7.8.1. Field Study Method

The recruitment of participants for the first KnowDis field study was liaised by internal staff with the organisation. Our choice of participants focused on individuals that qualify as knowledge workers within the organisation.

In total, 16 participants (15 male, 1 female) took part in the evaluation of the first KnowDis prototype, covering business analysts (P4), software architects (P3, P5), other development roles (P8, P16), various management functions (P1, P2, P9, P13, P15) and a range of R&D roles (P6, P7, P10, P11, P12, P14).

Nine participants reported mostly working from an office, five reported mostly working from home, and two reported a mix of home and office working. One of the participants was 30 years or younger, four between 30 and 40, four between 40 and 50, and seven above 50 years old.

The study lasted three weeks, with participants making active use of KnowDis for little over one week on average (M = 5.4 days, SD = 4.5). On active days, participants received recommendations through KnowDis on average 40.2 times (SD = 31.5), primarily through selecting a new email in Outlook.

7.8.1.1. The Participant's role

Participants were asked to install and then use the KnowDis prototype for a period of two or more weeks on their main work computer on which they used Outlook on a day-to-day basis. By default, the prototype was visible and active within Outlook, but could be minimised or closed at any time. While the KnowDis prototype is minimised or closed, it would stay inactive and not analyse any email or run any search calls.

In order to gain feedback from the field study, participants were asked to fill in pre- and post-study questionnaires. A subset of participants were also asked to participate in semi-structured interviews to get more in-depth qualitative feedback from participants on some of the aspects surfaced by the questionnaires.

7.8.1.2. Privacy Matters

In addition to direct feedback via questionnaires and interviews, the KnowDis prototype also logs user activities such as clicking on recommended resources. This log data should provide additional objective behavioural information into how the prototype is being used. As such data may contain private or confidential information, the log data is stored on the participant's computer. A function has been built into the prototype which creates and displays an email message containing all of the participant's log data (see Figure 74). This approach allows the participant to inspect the data before sending it through for analysis (see Figure 75).

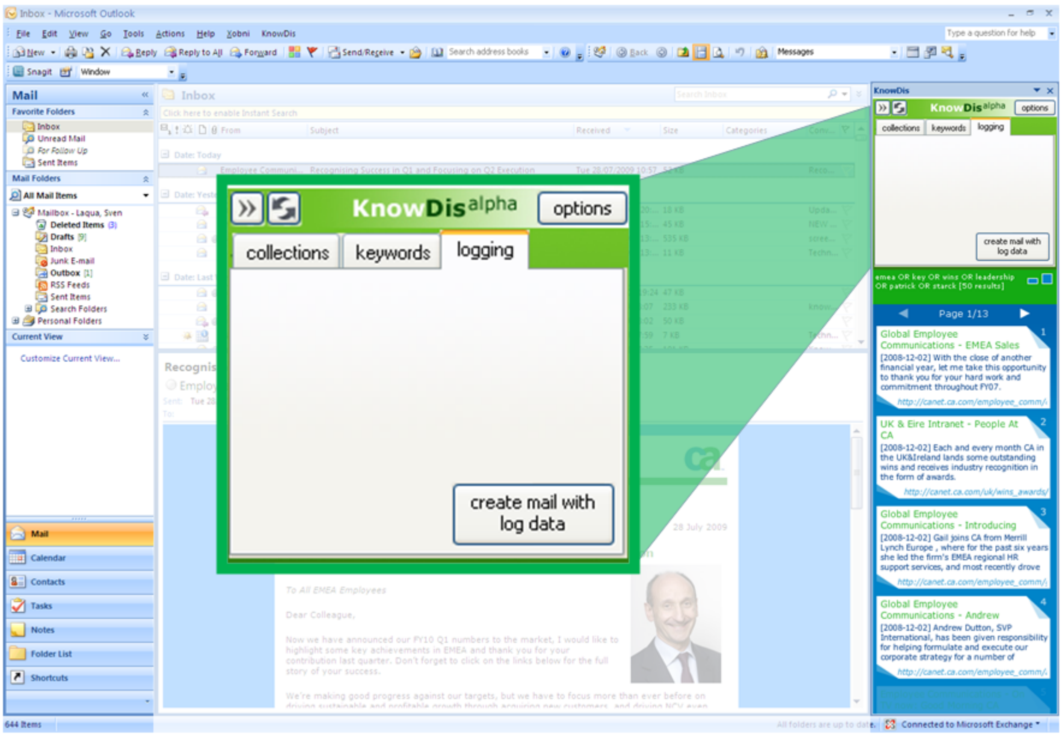


Figure 74. Sending log data - Within the options menu, the third tab, called "logging", provided functionality to generate an email message with the user's log data, by clicking the available button

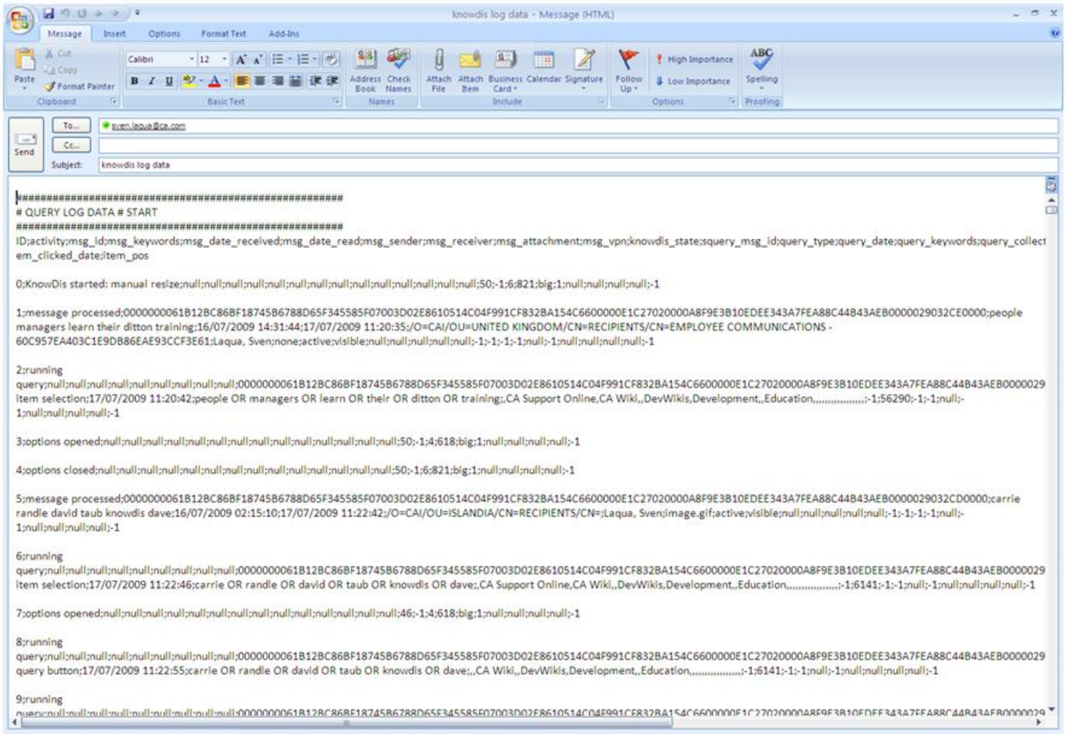


Figure 75. Email message containing the log data.

At the end of the field study, participants received an email, in which they were are asked to send their log data, if they are happy to share it. Participants were informed in detail about what data has been collected, and what data was not collected in addition to being able to inspect the log files.

They were also informed that although the log data would inevitably contain some personal information, the analysis of that data will be done in an anonymised manner. Specifically, participants were informed that the data collected consists of the keywords generated from their email messages when the KnowDis prototype was active as well as some metadata from those messages, such as date, anonymised sender information, and file names of any attachments.

Participants were further informed that log data would not contain the subject of email messages, actual message bodies, the actual attachments themselves, nor any identifiable information about the sender. Finally, participants were reminded that they should feel free to get in touch, if they had any questions regarding the privacy aspects of this study.

7.8.2. Post-study Interview Method

After the three week field trial, all participants were asked to participate in post-study interviews. 14 out of the 16 participants agreed to and took part in those interviews, which lasted approximately one hour each. The interviews were semi-structured and split into two parts. In the first part, the conversation focused around understanding the participant's work habits in more detail. To help with this, a list of seven guiding questions was used to inform the conversation:

1. Please talk a bit more about your work? What do have to do on a typical day?
2. Where do you look for information as part of your work? (Desktop, Intranet, Outlook, Google, Colleagues, etc…)
3. How do you manage/archive your work-related documents? (Desktop, Email, Bookmarks, SharePoint, etc…)
4. How do you typically retrieve information from Intranet resources, and which resources do you use? (Bookmarks, Search, etc…)
5. Do you recall an instance of coming across work-related information when it is too late (after this information would have been useful to you)?
6. Do you usually work with the same people over a longer period (e.g. long-term projects)? Or how much do people or groups you work with change?
7. If information necessary for your work does not reside inside documents, email or another person, please elaborate on how you extract that information, and from where.

In the second part, the conversation focused specifically on the participant's experience with and feedback on KnowDis. A list of ten questions was used to inform this second part of the conversation:

1. What were your expectations of KnowDis? And were those expectations met?
2. Did you come across any useful information, recommended by KnowDis?
   What was that information? Why was it relevant?
3. Do you remember an instance where you were reading an email and thought KnowDis should have some relevant information but it didn't?
4. Your log shows that you clicked on the following item <use from participant's log>… Was there a particular reason, why this was relevant?
5. Let's say, you would receive an email asking for the information contained in that recommended item, and let's say you did not have KnowDis recommending it to you – how would you try finding this information?
6. How did you make use of the collection settings? (e.g. leave as is, adjust once, change continuously)
7. Do you know which collections within KnowDis are relevant (or not) to you? Could you try naming the collections most relevant to you?
8. Did you use the keyword setting features? If yes, how? If not, why not?
9. Which version of the UI did you prefer, and why?
10. In which cases did you close KnowDis? (Why?)

At the end of each interview, the participant was asked for any other comments or questions about the study.

7.8.3. Results

7.8.3.1. Lessons Learnt

Participants agreed on email being their central hub of communication throughout the day. Although mobile usage of email (via blackberries typically) seems to be increasing and is quite significant for some individuals, Outlook, within this corporation acts as the main means to manage email and the large number of meetings and teleconferences.

A particular problem for all participants was space limitations of their Exchange account. As a result, participants typically archive their email in one or more local archives. At the extremes, we came across employees with 20,000+ emails in sent folders, 70,000+ emails archived and hundreds of email folders, managed across various logical email archive files. Folder management is typically based on projects, organisational bodies, or individuals.

Participants reported receiving 100 to 500 emails a day. Employees said that for 50% of replies, they needed to retrieve additional information, so improving this process seems particularly important. For most participants, email was the main information source: "*That's one of my largest pools of information, really, is my e-mail system*" (P3).

7.8.3.2. Knowing the Right People

However, needed information may not just reside locally in other email, but also in their personal files, on the corporate intranet, or somewhere on the Web. Accordingly, some of our participants had an 'escalation strategy': own memory > desktop search > web search > ask a colleague. Contacting colleagues is the most common fall-back method used to obtain required information. This has significant implications for productivity because domain experts are often overwhelmed with requests.

Some participants made use of specialised applications for their tasks, relying on information retrieval mechanisms embedded in those tools. Those participants articulated a strong desire for integration: "*can you integrate this with Google Desktop?*" (P3). The convenience of a single way of getting information is why many resort to asking colleagues; it may seem the fastest way to an individual employee, but has significant implications for organisational productivity.

7.8.3.3. Email & Email Overload

A number of participants reported struggling to cope with the amount of email they receive: "*I get a couple of hundred e-mails a day sometimes, and just trying to go through that actually can be a real problem*" (P1). One participant reported receiving more than 500 emails per day, suggesting that he has to "*look at all of them [and] end up personally replying to several hundred*" (P9). For P11, this amount of email leads to "*a great deal of context switching*", particularly where emails are perceived as so crucial that "*we often have to drop whatever we were doing and take care of it*", such as when helping a colleague with a customer request, or some "*high priority corporate initiative*" (P11). P12 feels particularly frustrated by frequent "*How do I do this?*" type questions related to the organisation's internal SharePoint installation, as he frequently ends up having to look up answers to them himself. Another participant reports checking email "*basically constantly, very rarely I don't check email, I don't display that screen from Outlook window for more than 5 minutes or so*" (P6) as he likes to "*give preference to emails unless I am writing a document or programming*" when he might not look at email for up to two hours.

<u>Coping Strategies</u>

In order to cope with the amount of email they receive, some participants archive "*stuff every day or two*" (P1) or use "*macros that move, filter and move those emails*" (P9) automatically into sophisticated folder systems that relate to projects, topics, or tasks. For P9, this method is crucial to enable him dealing with his email archives containing "*certainly in excess of 75 000 emails*" (P9). P12 reports that he makes "*folders and subfolders for pretty much everything*" (P12) and P8 reckons he has "*probably about 60 or 70 folders*" in his Outlook. Another participants creates "*folders for everyone who I get a lot of email from*" (P14). However, P7 notes that setting up effective filtering for emails (e.g. by project) can be hard, as some colleagues work on different projects thus making the set up of filters fine enough for emails to get to the correct project nearly impossible, resulting in things getting "*all kind of mushed up*" (P7). P11 tries to cope by being "*more organised with regard to when I'm going to pay attention to email*". For him this means setting "*aside large chunks of time for making things*" (P11) and not letting "*email drive your entire schedule, because that's just a recipe for failure*" (P11).

A number of participants report using Google Desktop to search their old email as they find Outlook "*really, really difficult to search*" (P14).

Archiving Strategies

For many participants, archiving large amounts of old email seems to be a crucial activity. P3 reports keeping archived email for a few years "*so that I can refer back as far as I need to*" suggesting that "*one of my largest pools of information, really, is my e-mail system*" (P3). This strategy seems to be grounded at least partially in him having had experiences with customers where discussions required him to "*prove certain things*" (P12). Another participant devised a particular strategy to save information received within email: "*I will change the subject line to have key words and then I'll send it to myself, and then remove the original email. So then if I need to find that again then I'll just use the search with the find capability of Outlook…*" (P8). P8 reports having around 24,000 sent emails in Outlook. For P7, who has to frequently switch between a number of projects, Outlook reportedly acts as "*a repository of my memory, so every time when I switch back [to a project] I look for emails that were exchanged recently and read them up and this is how get my context back*" (P7).

7.8.3.4. Workflow Feedback

Receiving Requests for Information

A significant amount of email participants receive represents requests for information. These requests may range from simply easy to deal with requests - "*80% or maybe even 90% of the emails that I get I can respond to without research and they take 20% of my time*" (P11) - to complex tasks that require additional research and carefully crafted responses: "*I can think of emails that are only one or two paragraphs in length that can take an hour to write because they need to be revised and quite often need specific data*" (P11). Often, the combination of too many emails and too many requests for information means participants start to feel overwhelmed: "*There are days when some of us get hundreds of emails and being able to figure out which ones you're going to get to in the thirty minutes that you may have at any given time, it becomes, you know if you can get good at that your life becomes better…*" (P11). As frequently such email requests come from "*higher up in the organisation*" or from "*a different part of the business*", there is additionally felt pressure to choose words carefully and be very "*specific about how things are phrased*" as to not cause others to misinterpret a response, as this would lead to a lot more email to clarify things and to "*putting out little fires that you unintentionally create*" (P11). P14 recalls an instance where the marketing team would ask for information about a research project that they should be able to find themselves: "*I don't know whether they search the internet first before asking me, they probably didn't but… I send them the documents and the links*" (P14). While some information requests may be due to senders not being able to find the information they are looking for, there seems to be a strong sense that "*people ask questions through email rather than research it themselves*" (P08). As these requests become habitual - "*I might get 20 to 25 emails a day from various people either giving me a nomination or asking me a question…*" (P08) - the burden perceived by recipients steadily increases: "*so what they do is they ask a question and then they almost assign the ownership to the person they ask … so then it becomes the recipient's obligation to find the answer*" (P08).

Finding Information

Participants report on a range on information finding strategies that broadly confirm and further strengthen the insights gained in the initial user requirements interviews. Some participants struggle with the relevance of documents on the intranet - "*it's dreadful… you get something with a low relevance and you get 200 pages of stuff*" (P7) and thus prefer reaching out to people who might be able to help first. For P10, who just joined recently, "*learning whom to contact for certain information*" (P10) is part of building up a network within the organisation. However, some participants reflect more on the impact frequent emailing of simple questions to colleagues might have: "*I don't want to bother them. I guess I kind of put myself in their shoes. If I get emails all day from people I don't know asking me simple questions then that would be pretty bothersome*" (P08). These participants "*don't waste other people's time*" (P12) and use contacting colleagues as a last resort: "*I only want to go to them for help when it's an important thing that I can't figure out on my own*" (P12). A few participants rely heavily on information stored locally, such as research articles, email archives, admin-type documents, event-related information, or documents on groups they belong to or projects they are involved in. While for some, organising large amounts of information locally seems to work well - "*my filing system usually helps me to get to where I need to go*" (P3), others seem to struggle more: "*I've got folders for each project and I've got folders for different kinds of stuff I'm looking into. I have about sixty folders, that's Level 1 folders and then there's also folders within folders… There's so many that I* *have trouble finding stuff…*" (P14). Consequently, desktop search is a more or less vital technology for these participants: "*I use the Windows Desktop Search… I can't work without it*" (P11).

Depending on their role, some participants spend a lot of their time looking for information on the web, researching relevant technologies "*doing Google searches, going into websites like Apache and researching the documentation and researching forums where people have had issues on things to find out what the resolution is*" (P5). To an extend, the preference for web rather than intranet search is based on the conviction that the relevant information cannot be found internally: "*they're very specifically related questions like database nuances and how to do something… I don't think I could get the information I needed inside [the organisation]*" (P8). But the ubiquitousness and effectiveness of Google also plays a vital role: "*I am more kind of a Google search person*" (P12).

Intranet Usage

The intranet within the organisation is made up of a large number of separate repositories, each more or less relevant to individual participants depending on their role. For some technical repositories are useful to find "*interesting stuff there on my product*" and to "*stumble across a whole bunch of other stuff that is interesting as well*" (P3). But when trying to make sense of the intranet at a higher level, search seems to fail most participants: "*I've occasionally done intranet searches, but to be honest with you, I mean, I haven't done it in a while, but they were absolutely rubbish. In the past, the results were dire, so it was usually a futile act*" (P3). P3's experience with searching the intranet is mirrored by many other participants, who argue that "*the intranet is a disaster*" (P7), because "*you get mostly garbage*" (P6) and "*anything… is next to impossible to find*" (P5). A few participants seem to be more patient, acknowledging that "*it often takes a long time to find*" relevant information because there ordering of results "*is really not that good so you might need to be on page 8 to find the document you want*" (P14). Another concern undermining participants' trust into intranet search is about what is indexed and thus findable using the search, and what is not.

"*The only thing to me, the biggest limitation, is that you're using [the organisation's intranet search engines] and they're pretty poor technologies so it's hard to build a really good system on top of those crappy technologies.*" (P14)

While searching the intranet is widely regarded as ineffective, a lot of this seems to have to do with bad past experiences and participants not being accustomed to how to use intranet search effectively, as illustrated by this response to the latest intranet search engine: "*No, never heard of it… Well I can see right now, I have just done a search on IPV6 there and I am seeing sorts of development related bits, which you would never find in the other search … Oh yes I am seeing all sort of stuff! Man I wish I knew about this! … Yes … Oh man! … This is the kind of stuff that I have needed all along and didn't even know it was there*" (P5).

Sharepoint Usage

Sharepoint is a central part of the organisation's intranet. But as one participant admitted, "*many, many people share documents as attachments*" (P9). Others acknowledged that while SharePoint may be usable when dealing with a few files, it quickly becomes unmanageable when dealing with a lot files across a lot of folders (P14). P9 partly attributed SharePoint's limitations to performance problems - "*it's slow*" - and to "*weak search coverage*". Another participant was not even sure, whether SharePoint sites were searchable on the Intranet at all: "*they're not exposed, are they?*" (P3). Another problem with SharePoint are its limited support for collaboration as documents could only be edited by one person at the time - "*So what typically happens is that every time you edit a document on the Sharepoint you give it a new name just so that you don't risk overwriting another part of someone else's work…*" (P14).

7.8.3.5. KnowDis Feedback

The evaluation of the first KnowDis prototype was primarily focused around a proof-of-concept.

General Impressions

Overall, user feedback on the first KnowDis prototype ranged from quite positive - "*my first impression was more of wow*" (P8), "*It worked pretty well… I want to continue to use it*" (P4), and "*getting it working and using it, it was a piece of cake. It was very simple.*" (P11) - to somewhat negative - "*I didn't get much benefit from KnowDis*" (P1).

A few participants likened KnowDis to Google Mail's contextual ads, acknowledging that applying the concept "*in the company context … would be very very good*" (P7), as it would surface links to new projects or other interesting information whilst reading email (P14). Most participants appreciated the concept of KnowDis as an add-on to

Outlook: "*One of the things I really like about KnowDis was that it brought together all these different search locations, made it really convenient and combining that with email made it a little bit more powerful.*" (P11) Only one participant did not feel that the general concept of KnowDis recommending related content would be applicable to her work: "*my job function didn't require me to have to be looking up old documents, or old webinars or whatever, so I think that just because it wasn't applicable to me, maybe that's why I didn't find it very useful*" (P12).

Relevance

Some participants found "*it pretty impressive that [KnowDis] was able to locate so many different [things], from different sources*" (P8), and appreciated "*how it brought together all of the different sources of material*" (P11). But a perceived lack of relevance of recommendations was probably the biggest area of concern for participants. The general potential participants could see in KnowDis was undermined by the lack of relevance of recommendations as attributed to problems with keywords as well as the underlying search: "*I think that KnowDis could be extremely useful. For me it was not, primarily because the keywords… And I believe that, with a better search, it could be quite useful*" (P9). P2 suggested that part of the problem of irrelevant keywords may be related to large email threads where "*keywords may be largely based on the content of a long initial email rather than on the content of short follow-up replies*" (P2). Nevertheless, some participants "*tried to keep it open as much as possible to see what type of content is going up*" (P4) and occasionally "*found something that was very interesting but not really related to … the email I was writing*" (P6).

Names

One particular issue with keyword generation was the inclusion of names: "*if I look at the keywords, some of them would be Daniel and Steven, Stewart, because they're the most in use there…*" (P6), as acknowledged by a number of participants (P6, P9, P11, P12). As a result, participants perceived recommendations as "*just name-based*" (P12) instead of content-based and felt that because of that, keywords were not very good (P11).

Customization

One of the key mechanisms to improve the relevance of recommendations KnowDis provided to participants was the adjustment of collection settings. Not all participants were aware of this functionality during the study: "*I might have missed something* *because I don't recall ever configuring the collections, so maybe I was hasty in installing it and didn't read all the documentation*" (P5). One participant did notice the settings but "*didn't change anything… because… all of those collections could potentially contain information that's interesting*" (P14). However, the potential impact this customisation could have on participants is best expressed by the commend of P7 after walking him through the process of re-enabling KnowDis during the post-study interview: "*Okay. Options yeah, okay. Oh I see, so you're showing CA QMS, support… Oh this is cool, so this is like…yeah it's like meta search, this is cool.*" … Oh this is good. I like it. Ah.*" (P7)

Email

Most participants agreed that "*email is a logical place to have [KnowDis]*" (P3), because it is the central place for organisational communication, exchanging information, and coordinating collaboration. However, one participant expressed surprise about the fact that "*KnowDis was coming in as an Outlook plug in*" (P7), as he was expecting something more like a dedicated browser-based application on the intranet.

Adoption

One participant embraced KnowDis as a tool that "might help me to understand patterns of communication better … it's in the corner of your eye and you can see that when you're in the context of a particular e-mail, that there's a whole bunch of … information that might be of interest" (P3).

P3 actively limited the search scope of KnowDis to only the repositories most relevant to his work. This participant was possibly more realistic about the potential benefits of KnowDis than some other participants suggesting that he was not looking for KnowDis to "*give me, within one click, all of the information that I would likely have to go off and search for myself.*" He further elaborates on his impressing of KnowDis, saying: "*I'm not saying it didn't do that, it's just that I wasn't looking at it in that respect. I was probably seeing it more as something that was – ah, okay, so there's been an e-mail discussion on this before, and it's come up on one of the distribution lists or there's a document out there that might be of interest*" (P3). His use of KnowDis also highlighted some of the issues with the general quality of the organisation's intranet that tainted some other participants' perceptions of KnowDis: "*So there was a click-through that I did – it was to a TechScan page that looked interesting, related to the subject of one of* *the e-mails, but actually when I got to the TechScan page, it hadn't been filled out – it was a blank. So I remember thinking – they should probably update that…*" (P3). Similarly, P14 suggested that he did not use KnowDis "*as a search tool per se, it was more something that alerted me to related stuff. I did play around with the keywords on a couple of occasions but I wouldn't do that for every email I'm reading… [KnowDis] already is useful in that when I was using it it did make me aware of other stuff that's happening in [the organisation]*" (P14).

<u>Crashes</u>

Some participants experienced problems with KnowDis disappearing after a restart of Outlook - "*it would just be gone*" (P1) or "*it didn't come up again*" (P7). While it was possible for participants to reenable the KnowDis add-on, it is not very straight-forward or self-explanatory how to reenable add- ons in Outlook, nor does Outlook itself report on any automatically deactivated add-ons.

"*I think I was travelling that week so I didn't take the time to figure out how to re-enable it until today when I sent you the log. So I went through the menus and I find the KnowDis add-on and I found out how to re-enable it… It certainly wasn't obvious how to re-enable it. I found your email which talked about sending the log and that gave me clues about how to do it.*" (P14)

Moreover, as this problems persisted for some participants (P1, P5), it meant they would eventually give up on using it: "*KnowDis died for me, like continually… you have to go to the Trust Centre to try to bring it back… after a while, you just stopped using it.*" (P1)

<u>UI</u>

Participants generally agreed that the integration of KnowDis into the Outlook user interface is appropriate: *"Setting the options and paging through and expanding and contracting the windows, that was really intuitive. It was very simple."* (P11) The primary feedback on visual changes to KnowDis were around it occupying less space when minimised (P1, P11) and the need for increased font sizes, as expressed by one participant, whose *"eyesight is not that great"* (P7). The KnowDis prototype allowed participants to adjust the size of the recommendations, switching between 'standard' and 'mini' (see section 7.7). Not all participants noticed the ability to change this part of the UI. While P1 preferred the 'mini mode' *"because you got more information"*, P14 argued that *"the small one I think it's too many things… it's too much to absorb… if there's a really long list, I'm less likely to see anything"*. Between the participants who experimented with the setting to change the size of the recommendations, a preference for the standard mode emerged (P4, P10, P14).

On Improving Scope

A number of participants expressed a strong desire for KnowDis to provide recommendations for more than the organisation's intranet repositories (P4, P6, P7, P8, P9). In particular, the inclusion of web search was perceived as desirable as for some tasks "*it would help if I could find something outside [the organisation]*" (P6). Other participants expressed a need to receive recommendations of documents (P6) or email (P9) on their computer. In addition to the extension of sources for automatic recommendations, P8 and P11 suggested that the inclusion of traditional search box that lets users manually retrieve recommendations "*given whatever you want to type in… would be the most useful feature*" (P8), "*because I did like very much how it tied all those different search sources together*" (P11). Finally, P5 suggested that it would be useful for KnowDis to identify and recommend people within the organisation that might be able to help with a problem, "*because a lot of the times you know you are dealing with a problem that others in [the organisation] have dealt with, but you don't know who they are to talk to them to get…*" (P5).

On Improving Personalisation

Some participants argued for greater control over how recommendations are generated, such as via the use of a list of priority words (P7, P9, P12). Such a list could be user-populated - "*I can put those facts in there if it will help KnowDis to find the relevant information*" (P12) - and may be related to information sources (e.g. blogs), project names or project topics (e.g. root cause analysis) or specific products and product groups - "*if KnowDis could actually track that for me it would be amazing*" (P7). Another angle to more effective recommendations was the suggestion to allow users to manually specify "*folders that KnowDis should be active in*" for the indexing of local information - "*if there is a way that you could set my CR Requests folder, that would be a good one because those are my tickets*" (P12). One participant extended that idea by suggesting the ability to automatically utilise users' commonly used folders or intranet resources to better tailor recommendations - "*What would be really cool is if you could trace ... what folders in my computer I usually look at or which part of the Wiki or* *which part of the SharePoint I usually look at [after reading an email]*" (P14).

On Improving Usability

While an increase in search scope (additional search sources) and search accuracy (improved personalisation) were the most crucial improvements participants requested, a small range of usability improvements were made as well. P4 suggested the ability to drag and drop something into KnowDis to retrieve recommendations, while P11 wanted to option to instantly preview or expand recommendation items in case the list of recommendations "*had a number of things that all looked like they had potential… then I would want to be able to expand it without clicking on the link*" (P11).

7.8.4. New User Requirements

The following three user requirements are the most critical improvements to KnowDis derived from the first field study:

1. First, a more robust and reliable KnowDis prototype (see section 'crashes' in 7.8.3.5). As the initial prototype regularly disappeared from Outlook for some participants, or stopped working after a while, this made it hard to reliably conduct a field study over a number of weeks.
2. Second, improving the relevance of recommendations seems to be a second critical aspect to be addressed (see section 'relevance' in 7.8.3.5). One of the issues raised in relation to the relevance of recommendations was the inclusion of names derived from email messages, which some participants attributed to the poor quality of recommendations.
3. Third, participants expressed a need for KnowDis to include recommendations for content from beyond the organisation's intranet (see section 'on improving scope' in 7.8.3.5). While some wanted to see the integration of web search, others asked for recommendations of local files or email that are stored on their computers. As part of these increases in search scope, a few participants also wanted to be able to conduct searches manually from within KnowDis.

Other user requirements, while also relevant and useful, were out of scope for the implementation of the second KnowDis prototype.

7.9. Description of Second Prototype

Based on the findings in Phase 1, we re-designed the initial prototype. In its second iteration, KnowDis is a fully functional prototype add-in for Microsoft Outlook that uses keyword extraction to make sense of the active email message, and proactively displays without any user intervention, links to information on the local machine, a company's intranet resources, and the Web (see Figure 76).

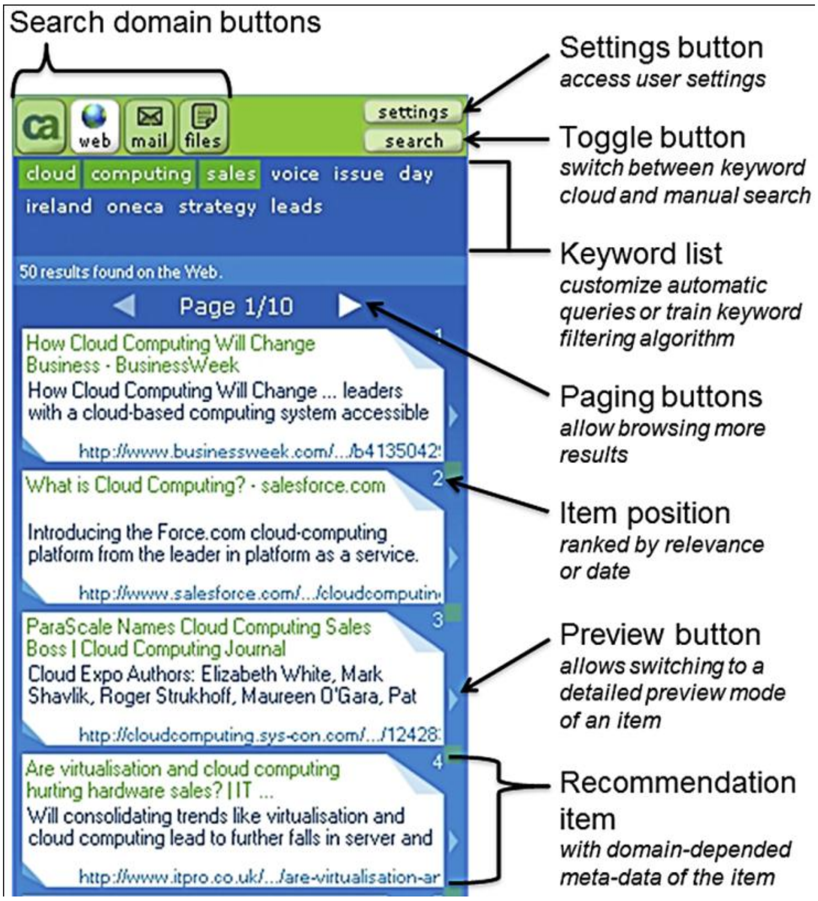


Figure 76. User Interface of the Second KnowDis Prototype

As an Outlook add-in, KnowDis becomes part of the Outlook User Interface (see Figure 77), its location, size and other aspects of the KnowDis UI being customizable (see Figure 78). The second iteration of KnowDis utilises background threading for improved performance, Add-in Express[6] for improved deployability across Windows environments, and a more flexible UI for greater customizability.

[6] Add-in Express is an extension for Visual Studio to facilitate the development of Microsoft Office add-ins.

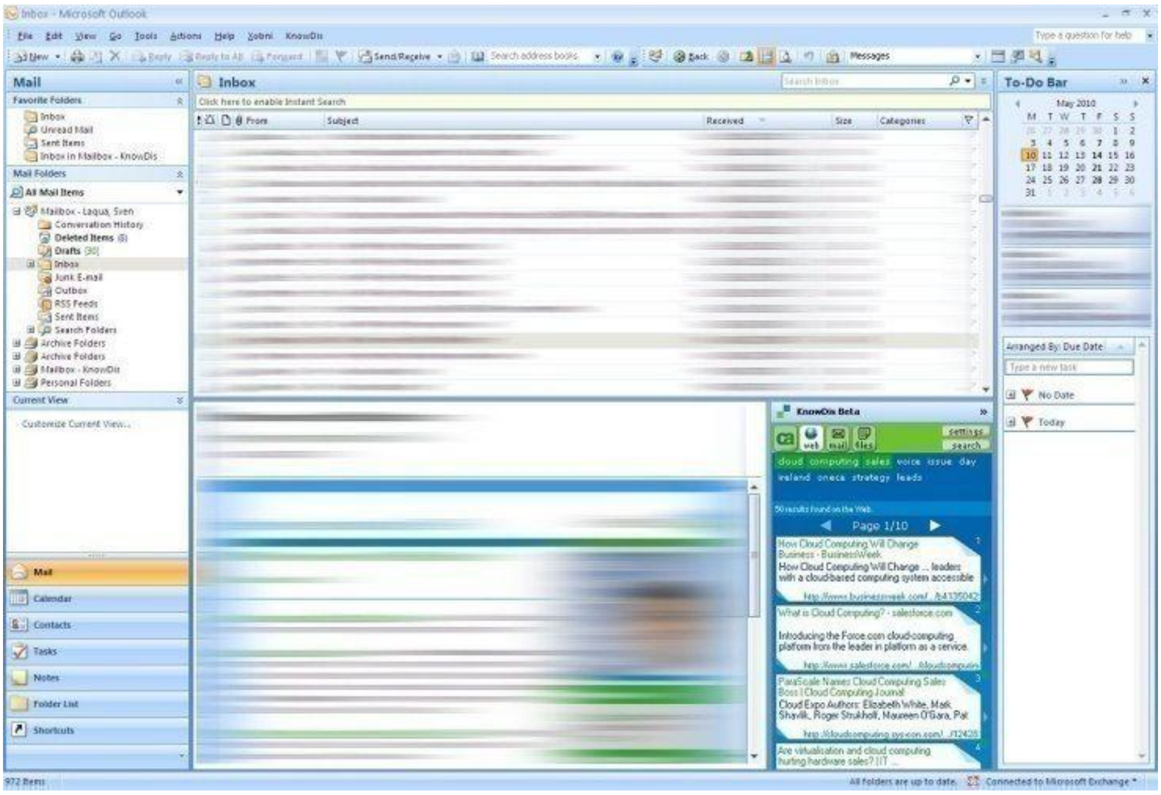


Figure 77. Integration of KnowDis in Outlook

Without having to switch the focus of attention away from reading email, relevant documents are made available, such as prior presentations, journal articles, competitor news, etc.

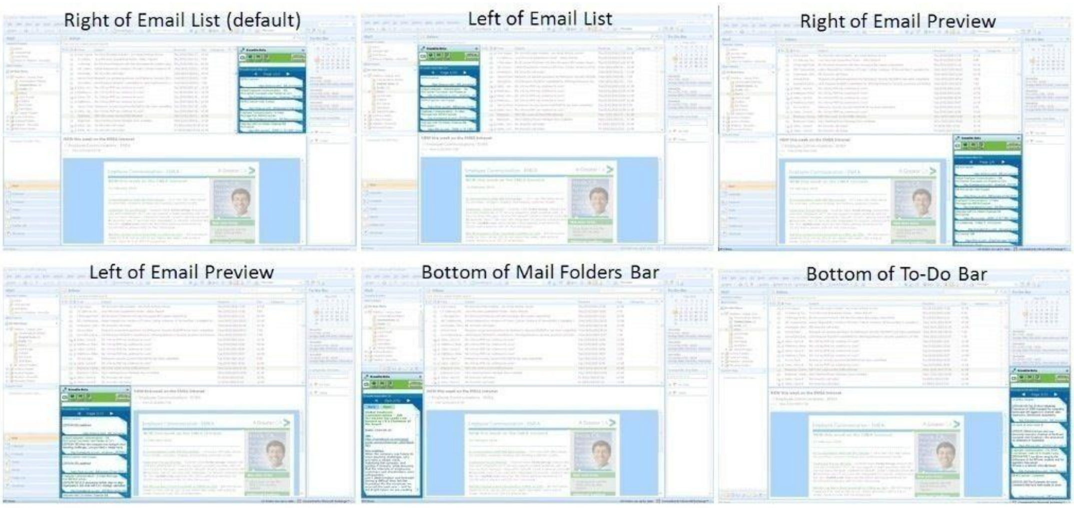


Figure 78. Flexible positioning of the KnowDis UI within Outlook.

7.9.1. The KnowDis User Interface

The recommendation items are the central component of the KnowDis UI. When the user selects a new email message and the email remains active beyond a short interval, the selected search domain is queried using keywords generated from the active email.

The search results are displayed as recommended items. Through experimentation, we have chosen 1.5sec as a reasonable interval, to compensate for quick continuous email selections, e.g. when browsing/navigating using a keyboard, before a query is being sent. If the user selects another email message before the search results have been displayed, a new query is launched and the recommended items from the first query are not displayed. The number of visible recommendation items is adjusted automatically based on the visual height available for KnowDis in Outlook and selected size of recommendation items (adjustable from 0-8 lines of description).

Figure 79. KnowDis recommendation - Title only recommendation

Figure 80. KnowDis recommendation - Title and 2 line description recommendation

Figure 81. KnowDis recommendation - Title and 4 line description recommendation

Figure 82. KnowDis recommendation - Title and 8 line description recommendation

The preview button allows switching to a more detailed view of a particular recommendation item. The preview mode uses the entire height available within KnowDis.

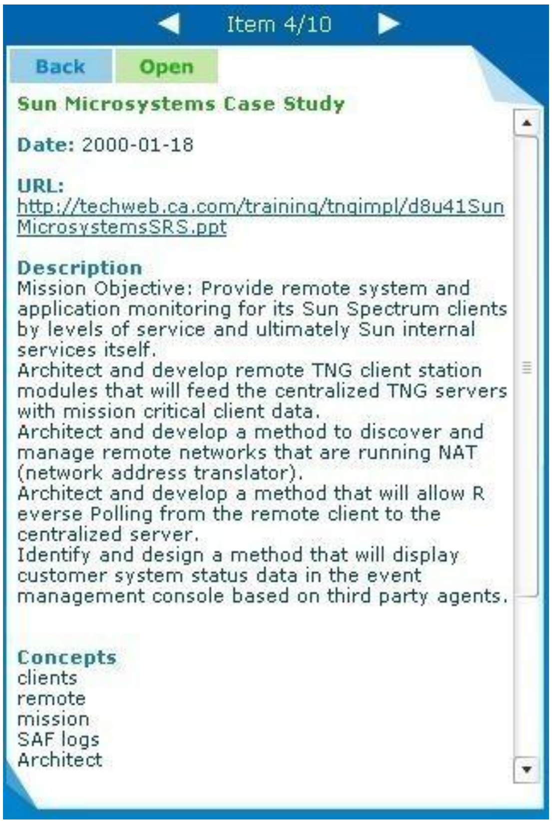


Figure 83. KnowDis preview mode

Search domain buttons in the top-left of the KnowDis UI allow the user to switch between Intranet, web, email, and file search (see Figure 84).

Figure 84. KnowDis header - search domain buttons

A search/keywords toggle button further allows switching between the display of a keyword list and the display of a search field for manual searches (using self-chosen keywords).

Figure 85. KnowDis header with search field (UCL version)

Figure 86. KnowDis header with keyword list (highlighted keywords are used in search)

A settings button in the top-right provides access to the user settings menu (see Figure 87), which allow customizing:

1. Stop-lists for automatic keyword generation (adding new stop-words, or removing existing ones),
2. Knowledge repositories to be included in Intranet search (e.g. SharePoint, Wikis, etc.),
3. File-types to be used for desktop file searches (e.g. Word, Excel, PowerPoint, PDF, HTML, etc.),
4. Sites to be used in custom site searches on the Web (e.g. acm.org, sigchi.org, bbc.co.uk, etc.).

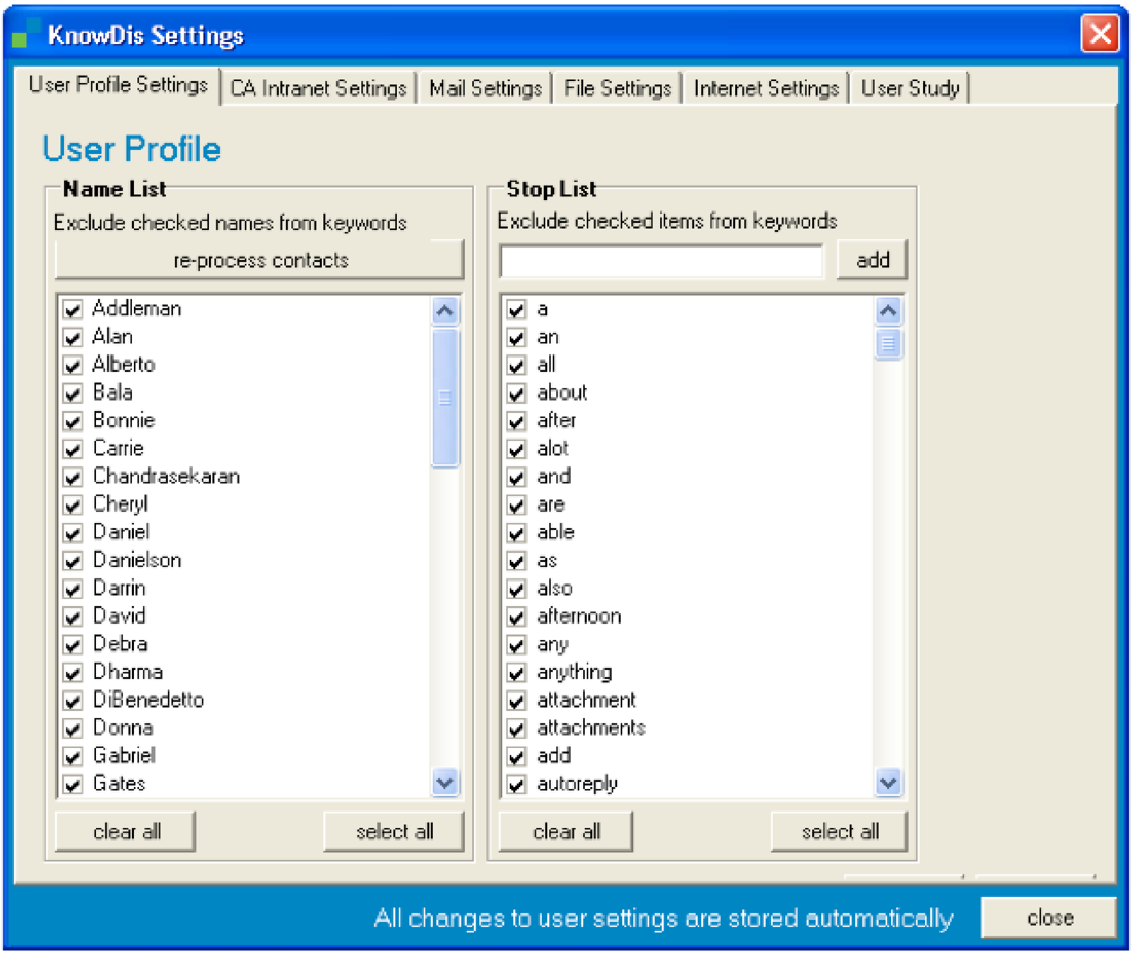


Figure 87. KnowDis Setting Menu

The keyword list displays the ten most relevant keywords extracted from the active email message. Based on experimentation and results from the pilot study in Phase 1, the first three keywords are used to automatically query the active search domain, and thus are highlighted in the keyword list. The user can select/deselect keywords as desired and re-query the desired domain, or train the keyword generation algorithm by adding undesired keywords to the existing stop-word list, in situ through a context menu.

7.9.2. Recommendations through Proactive Search

The KnowDis client processes a selected email, generates keywords from that email and then automatically initiates a call to one of the search providers for related information (see Figure 88). The specific algorithm used for the keyword generation can be found in Appendix C (see section 10.12). Although we are aware of potential concerns with regards to proactively providing recommendations (such as potentially being distractive), within this study we wanted to understand employees' perception of such an intervention, if done in an unobtrusive manner. The key benefits of displaying the results of the proactive search in this manner are:

1. No articulation of keywords is required.
2. No decision is required on whether to run a search for additional information or not – user can simply glance at initial results.

In a production version of KnowDis, employees would very likely be given the ability to switch between a pull and push-type recommendation mechanism. Although it needs to be noted that opting for a pull-type mechanism, will impact the user's ability to "stumble upon" good-to-know information.

1.9.3 Utilization of Existing Search Providers

The idea of KnowDis is to utilise existing search technologies. There are marginal disadvantages to this approach, mostly in the form of less flexibility in adjusting/configuring the search technologies. But those are easily outweighed considering the huge gains achieved by simply plugging existing search providers into the KnowDis architecture. The search providers available in the current prototype are:

1. Coveo Enterprise Search web service to enable querying corporate knowledge repositories
2. Microsoft Bing web service to enable querying the entire web, a specific web site, or a group of web sites
3. Google Desktop Search service to enable querying local files and email (implemented as separate search domains)

A KnowDis server component provides centralised activity logs of users' interactions with KnowDis and basic interaction activity within Outlook (see Figure 88).

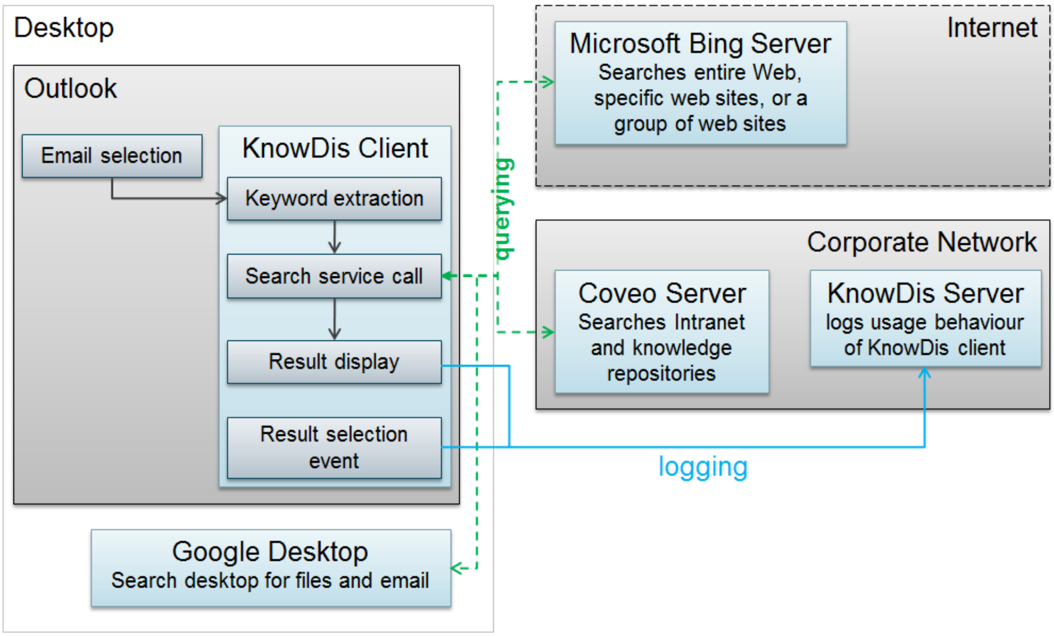


Figure 88. KnowDis Architecture Overview

7.10. Evaluation 2

Phase 2 was conducted over a seven week period. During that time, according to the company's IT department, no major upgrades of search engine or content taxonomies occurred.

7.10.1. Field Study Method

47 employees filled in the initial pre-study questionnaire. 36 of those employees installed and used KnowDis to some degree. Twenty-eight out of those 36 were classified as active users (more than one week of usage), with around thirty days of average active usage in a seven week period. After the official completion of the user study, we asked users to fill in a second post-study questionnaire, which was completed by 24 out of those 28 active users. All reported analyses of usage logs and questionnaire responses are based on those 24 active users – henceforth referred to as the study participants.

The 24 study participants (20 male, 4 female) were distributed globally - 15 in the US, five in India, and one each in Japan, Australia, Germany and the UK. Out of the 15 US participants, five were based in the company headquarter, the rest distributed across offices in different states. 17 of the study participants reported mostly working from an office, four participants reported mostly working from home, and three reported a mix of home and office working. Eight participants were between 25 and 34 years old, five participants were between 35 and 44 years old, nine participants were between 45 and 54 years old and two participants were between 55 and 64 years old. The study participants were from a range of technology groups in the company, covering various software and technology, sales, management and R&D roles. A further 13 users who installed the prototype did not complete the study. They showed little usage (less than a week), did not fill in the post-study questionnaire, or left the company during the study.

We also conducted 25 semi-structured interviews (most active KnowDis users, plus some less active) to understand employees' experience of KnowDis in more detail. Participants interviewed covered support engineers (P1, P12, P24, P25), software engineers (P3, P10, P17, P20), software architects (P7, P11, P16, P22), user experience architects (P4, P23), user interface designers (P13, P18), instructors (P6, P15), as well as sales (P14), management (P2, P5, P8) and R&D roles (P9, P19, P21).

To support the adoption of the second KnowDis prototype, a wiki page was created documenting the new prototype and all of its functionality in sufficient detail. The wiki contained a link to the KnowDis prototype download, installation instructions, an overview of the prototype's functionality, an FAQ section, as well as information about the participant's role and privacy implications (similar to the first study). In addition, ten short how-to videos were created and shared on the wiki. Those videos covered the following topics:

1. How to customise the KnowDis UI
2. How to customise keywords
3. How to move KnowDis within the Outlook UI
4. How to open recommendations
5. How to send feedback emails
6. How to switch search providers
7. How to use context menus
8. How to use desktop mail and file recommendations
9. How to search within KnowDis by drag and drop
10. How to use recommendations from the Web

7.10.2. Post study interview method

After the seven week field trial, all participants were asked to participate in post-study interviews. 25 participants agreed to and took part in those interviews, which lasted approximately 45 minutes each. For each of the interviews, participants' pre-questionnaire information was analysed to identify any particular issues that may have been raised by the participant and would benefit from further discussion within the interview.

The interviews were semi-structured and split into three parts. In the first part, participants were asked a few more general questions as a way to warm up:

1. Please talk briefly about your work at CA. What kind of tasks do you have to do on a typical day?
2. What were your expectations of KnowDis? How were those expectations met?
3. In particular, how would you describe your strategy to retrieve and manage information necessary for your work? (e.g. search usage, archival strategies)
4. Do you think KnowDis has impacted your work-strategy in any way (with regards to information retrieval and management)? If not, do you think it could have an impact on your work-strategy, if KnowDis would better meet your needs?

In the second part, the conversation drilled-down onto specific aspects about KnowDis, particularly the various customisation options within the new prototype:

1. Did you find the user settings in KnowDis useful? If yes, why? If not, why not?(e.g. adjust CA collections, web sites to search, use context search, etc.)
2. Did you find the keyword feature useful? If yes, why? If not, why not? How did you use it (e.g. adjust keywords, adjust stop-lists)?
3. Did you find the (manual) search functionality useful? If yes, why? If not, why not?
4. Did you find the UI customization features useful? If yes, why? If not, why not? (e.g. placement of KnowDis in Outlook, size of recommendation items, font size)
5. Did you minimize KnowDis at some point? (Why?)
6. Which changes would you like to see in KnowDis?

In the third part, the conversation concluded with some reflecting questions about the general usefulness of KnowDis in participants' work:

1. Do you remember an instance where KnowDis recommended some useful information to you? *(Use sent out summary email information for interviewee, if no example presented)* What was that information? Why was it relevant?
2. Do you remember an instance where you were reading an email and thought KnowDis should have some relevant information but it didn't?
3. Do you have any other comments or questions about the user study?

At the end of each interview, the participant was asked for any other comments or questions about the study.

7.10.3. Results

Overall, participants kept KnowDis visible in Outlook 74% of the time (calculated based on a participant's interactions with email messages). This number varied vastly between participants, with some having KnowDis visible all the time, while others would hide KnowDis 99% of time, and make it visible when they wanted to use it.

The search domain by default was the corporate knowledge repository and this domain was used 78% of the time; the public web was used as the search domain 21% of the time. Email and the local file system were rarely chosen as the search domain (combined usage was 1%). Two factors may have contributed to this distribution: 1) Corporate search was the default search domain, and 2) email and file search required the presence of Google Desktop on the participant's machine, which contrary to our findings in Phase 1 was not the case for most participants. Although KnowDis usage is largely passive because it provides recommendations unobtrusively, interaction with KnowDis was high in the first week, when participants familiarised themselves with KnowDis, and dropped to low levels in subsequent weeks. This is not surprising because click-through rates tend to be low for unobtrusive advertisements (Drèze & Hussherr, 2003).

A workload construct was developed by combining a variety of factors: The level of human interaction (with co-workers, or customers), email handling (reading, writing, archiving, re-finding, etc.) and information handling (creation, retrieval, sharing). Based on a total of 34 measures, workload was calculated for each participant. On a 0-1 scale, the average workload was 0.65 (min = 0.53, max = 0.92). Participants were split into a high workload (10 participants with above average load) and a low workload group (14 participants with below average load). We then analysed interactions with email and KnowDis separately for the two groups. A breakdown by week (see Figure 89) illustrates that high KnowDis usage in the first week was caused primarily by participants from the high workload group (red bars). Usage for that group drops of sharply from the second week. For participants of the low workload group (green bars) usage of KnowDis continuously increases until week 5, when it drops to a lower level. This drop is positively correlated with a drop in email usage (0.43) for the low workload group (green line). The drop in KnowDis usage for the high workload group is marginally negatively correlated with a rise in email usage (-0.18, red line).

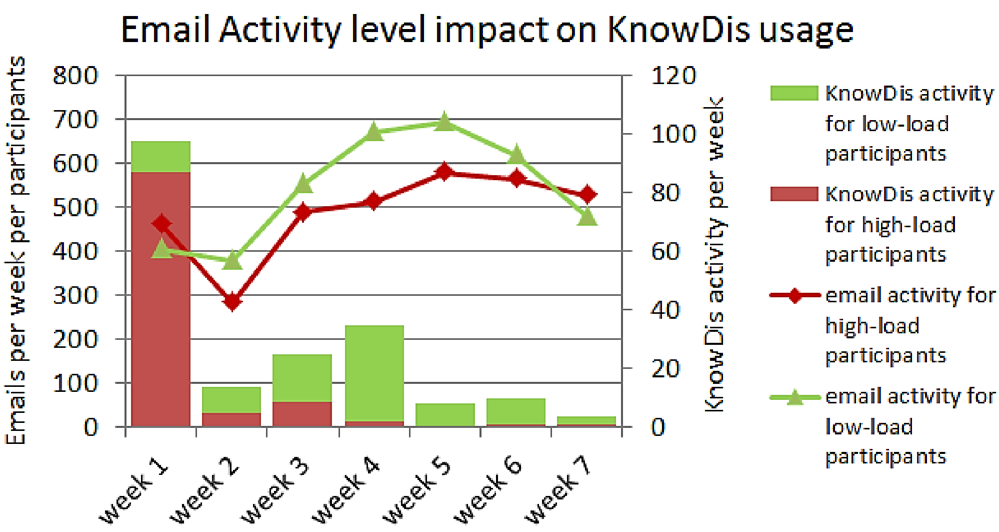


Figure 89. Email vs. KnowDis interactions for average participants by week of high vs. low load group

A breakdown by hour of day (see Figure 90) illustrates the pattern of email usage for the high workload group (red line) and low workload group (green line) and the contrasting styles of use for KnowDis. Whereas the low workload group makes use of KnowDis more consistently throughout the day (green bars) with a peak in the morning and another one in the afternoon, the high workload group shows significant peaks of use just before and after lunchtime (red bars).

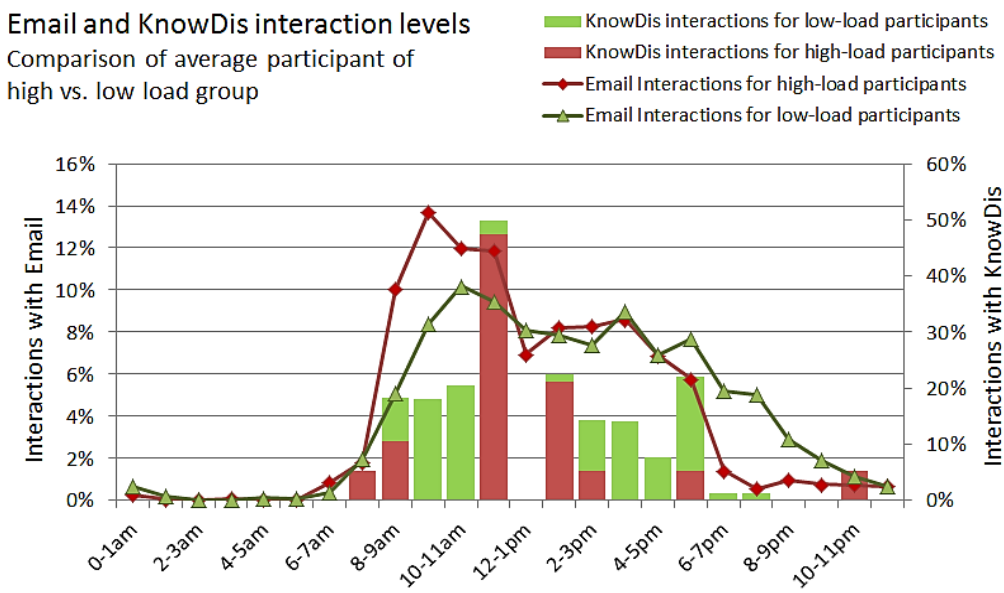


Figure 90. Email vs. KnowDis interactions for average participants by time of day of high vs. low load group (percentages relate to a day's use)

Notably, email usage within the high workload group is split evenly between morning

(50% before 12pm) and afternoon (50% after 12pm), while email usage within the low workload group is more concentrated in the afternoon (36% before 12pm, 64% after 12pm).

7.10.3.1. Questionnaires

As part of the study design, pre- and post-study questionnaires were used to evaluate the impact of KnowDis (as a Microsoft Outlook add-in). The statistical analysis is based on the 24 out of 28 active participants who filled in both pre- and post-study questionnaires. Significance is evaluated through repeated measures t-tests.

The questionnaires used a combination of open-response and closed-response questions. The Likert scales had either five or six alternatives, e.g., strongly disagree, disagree, somewhat disagree, somewhat agree, agree, and strongly agree. To facilitate comparisons across the different scales, all responses were normalised to a 0 to 1 range.

Both before and after using KnowDis, participants considered email essential to their work environment. Prior to using KnowDis, the average rating of the importance of email among active participants was 0.83 (on 0 to 1 range; SD = 0.18), and after using KnowDis the average was 0.79 (SD = 0.22). The difference between the two scores is not significant. Likewise, no statistical differences were found between pre-and post-study questions regarding their efficiency in using mail (0.54 and 0.51), nor in their agreement that they spend a lot of time trying to figure out where (i.e., wikis, blogs, public web, internal collections, previous email, etc.) to find relevant information when responding to emails (0.56 and 0.53). Thus, the overall use of email, its efficiency or the need for finding related information while reading email did not vary significantly between the pre- and post-questionnaires.

The most commonly used feature of KnowDis has been the automatic recommendation of related corporate information. Participants used this feature 78% of the time. There was not enough activity using the web, email or desktop search functions of KnowDis to warrant separate statistical analyses for those domains. Most of the KnowDis users agreed that email responses depending on information search are completed more efficiently and more effectively when related information is automatically available.

a) Participants found searching for information within the corporate knowledge repositories significantly more efficient, when having KnowDis available within Outlook. Prior to using KnowDis, the mean efficiency rating for search for corporate information while acting upon an email was 0.27. However, when asked in the context of their KnowDis experience, the rated efficiency increased to 0.40. This difference was significant in a paired t-test (t_{23}= -2.145, $p < 0.05$). This question made explicit reference to KnowDis and Outlook, and clearly suggests a benefit of KnowDis. Nonetheless the low values suggest much room for improvement beyond the current prototype. This topic will be further explored in the qualitative section below.

Two additional questions examined the impact of KnowDis on more general attitudes toward information searching, without explicit reference to KnowDis or email.

b) Participants found it significantly easier to find relevant information from another business unit, as the rating on an agreement scale decreased significantly from 0.79 to 0.72 for the statement: *"Relevant information is often difficult to find when it is created in another business unit in [the large IT organization]"* (t-test: $t_{23} = 2.23$, $p < 0.05$).

c) Participants felt less overwhelmed by the amount of information they needed to search to do their work, when having KnowDis available within Outlook. Ratings on an agreement scale decreased from 0.56 to 0.49 for the statement: *"I feel overwhelmed by the amount of information I need to search for in order to perform my job"*, but this difference was not significant. However, a post-hoc exploration of the data revealed an age effect. The eight participants that were under 35, showed no improvement in self-ratings of feeling overwhelmed, with the mean scores for this group increasing from 0.50 to 0.53. However, the 16 participants who were over 35, showed a significant decrease in feelings of being overwhelmed, from 0.575 to 0.475 ($t_{15} = 2.52$, $p < 0.05$).

A re-examination of the responses to the two previous questions showed a similar but far less dramatic difference. In those responses both young and old showed the same predicted change in behaviour.

Finally, we asked whether participants would like to continue using KnowDis after the study, 14 of 24 strongly or weakly agreed (weighted agreement score is 0.56). 16 of 24 participants strongly or weakly agreed with the statement: *"I would recommend KnowDis to a colleague"* (weighted agreement score is 0.58). Of the participants who would not like to continue using KnowDis (in its current state), 9 of 14 strongly or weakly agreed that they would use KnowDis more, if it were easier to use (weighted agreement score is 0.63). Notably, high ratings of KnowDis were uncorrelated with the increases in perceived efficiency noted above. This discrepancy is explained by some of the negative comments regarding the KnowDis design, in particular the decision to use Google desktop search to support the desktop and email search, by concerns about the internal search engine, and by whether or not training materials were used.

The post-questionnaire also revealed that a large proportion of participants (13 of 24) did not look at any of the supporting material for KnowDis– a user guide wiki and video demos of key features. We analyzed the impact of the study material on participants' wish to continue using KnowDis and found a significant difference ($t_{21} = 2.34$, $p < 0.05$) between the 2 groups:

- 8 of 11 participants who used at least some of the supporting material would like to continue using KnowDis (weighted agreement score is 0.67), vs.
- 6 of 13 participants who did not look at any of the materials would like to continue using KnowDis (weighted agreement score is 0.46)

Those numbers suggest that making use of at least some supporting material significantly increased the perceived usefulness of KnowDis. Considering that KnowDis is a prototype, we also asked whether participants would use KnowDis more, under certain conditions:

- 23 of 24 participants would use KnowDis more if search worked better (weighted agreement score is 0.96)
- 20 of 24 participants would use KnowDis more if results from various search domains were presented in a combined view (weighted agreement score is 0.83) *(out of participants who would like to continue using KnowDis, 100% said they would use it more to some degree, if it provided a combined view of results)*
- 16 of 24 participants would use KnowDis more if they could break their (information retrieval) habits (weighted agreement score is 0.67)

7.10.4. Results of interviews with 25 participants

7.10.4.1. Information Retrieval Practices

Web Search

The vast majority of participants makes heavy use of Google to deal with information retrieval tasks with one participant reportedly using Google "*100 times a day*" (P3) as he prefers re-searching information, rather than downloading it for later re-use. Web search is being used by participants to look for technical information (P2) such as windows issues (P16) or windows errors (P25), academic papers via Google Scholar (P19) or information on "*how to set up experiments*" (P23). While one participant reported using Yahoo for web searches in addition to Google (P1), none reported the use of any other search engine either additionally or exclusively for web searches. Some participants report using specific web sites for their web-based information retrieval. P1 uses Wikipedia and answers.yahoo.com to look up definitions in Japanese, whereas P15 uses Amazon and Oracle websites to stay on top of his primary teaching topic 'cloud computing'.

Desktop Search

Only a small number of participants still report using Google Desktop Search to find information locally on their computers (P1, P22, P23, P25). Some, such as P11 have switched to using Windows Desktop Search instead since upgrading to Windows 7. Others may not use any desktop search tool, or rely on specialised tools such as Mendeley (P3).

Intranet Search

The effectiveness of searching internal resources is particularly crucial given that some some participants report relying on it for 60% (P2) or 70% (P12) of their information retrieval tasks. However, participants generally acknowledge that retrieving information from the organisation's intranet is challenging. While some participants specifically pinpoint the intranet search engine as generally working badly (P7, P8) or more specifically having "*too many marketing results*" (P16) but not enough useful technical information, a few find it to work well (P11, P12), with one of the latter group referring to himself as an intranet search 'power user' particularly appreciating and making heavy use of the search engines vast filtering options (P14). Similarly, opinion about the organisation's SharePoint is divided, with some participants relying on it heavily for finding information (P2, P4, P22), and others perceiving it to work badly (P7). A few participants report primarily using specific internal knowledge repositories, such as knowledge bases (P6), online project management tools (P13), or software issue databases (P12) for their work. One participant working from India reported that their local intranet does not seem to be indexed by the organisation's intranet search engine (P13). A hint at more general problems with the organisation's intranet and consequently any search engine's ability to deliver meaningful recommendations is P7's request to group knowledge repositories more logically.

Email Search

For many participants, email archives represent a crucial repository of potentially relevant information. While some organise their email in folders by people who contact them frequently (P21), others prefer organising their folders by projects or topics (P22, P23). Some participants use a range of separate email archives to manage their 'email knowledge bases' (P3, P12), and one reports having archived all email of the last eight years (P8). Participants who report using desktop search to re-find email generally seem find this to work quite well (P8, P21, P22, P23). One participant reportedly does not use search much because his email is well-organised (P10) and another participant makes heavy use of an Outlook add-in called Xobni to help him stay organised (P22). P21 reports struggling with making sense of information emailed to his personal email account by external research collaborators.

Search Strategies

Some participants report using dedicated tools to support their information retrieval processes, such as delicious.com for online bookmarking (P4, P7) or mind-mapping tools for very important information (P3). Other participants acknowledge that they do not bookmark useful results from web searches (P5), or that bookmark management is getting too much (P20). Two participants report in more detail on their strategies to escalate the retrieval of information. For P6, this information retrieval strategy is:

1. Personal memory
2. Product documentation
3. Email archives
4. Intranet knowledge base
5. Google

For P14, a common information retrieval strategy to deal with customers' or partners' information requests via email is:

1. Local folder with relevant PDFs
2. Web search to check for more current versions of locally archived documents
3. Search of organisation's public-facing website for links to relevant white papers

For others, such as P11, talking to colleagues is a crucial strategy to retrieve relevant information.

Information Requests

The reported level of information requests received via email varies hugely between participants. Whereas P6 reports receiving one or two emails per week with a question for information, P20 estimates receiving around 250 emails per day from developers, a lot of which require "*some research*". Other participants, struggle less with the frequency of information requests, but with their vagueness. P7 reportedly struggles with people emailing for help or advice on a particular aspect of work without referencing the correct place to look at or without sharing the relevant documents. In a similar vain, P16 reports having to loo up or research information based on a "*very vague email and a 100 page attachment*".

7.10.4.2. Feedback On KnowDis

Motivation

Some of the participants reported taking part in the field study as they "*wanted to try something new*" (P2), or were simply curious (P8, P10). Others more concretely argued that they liked the concept of helping users find new information associated with corporate email (P2, P4, P21).

Relevance of Recommendations

Participants' opinions about the general usefulness of KnowDis were divided. Some liked the idea and hoped it would be useful but did not find KnowDis to provide 'real value' to them (P7, P9, P11). Other participants felt that KnowDis was useful to them (P20, P25), and provided "*a lot of value*" (P17). P20 acknowledged that KnowDis was both "*helpful and distracting*" and that KnowDis would require "*a learning curve to not click too often on information*" (P20), but that after getting used to it, he felt KnowDis was "*very useful*". P19 in particular noted that the way KnowDis' recommendations were being displayed meant that she could derive the organisation's repository each recommendation was coming from through the displayed URL. As the participant was fairly new to the company, this in turn allowed her to learn which repositories within the organisation are particularly useful and relevant. P19 even argued that she would receive better results through KnowDis than if she would manually search on the intranet search engine website. In a similar vain, P3 argued that a side-effect of KnowDis is "*useful hints*" to interesting information. Such information may not be immediately relevant to the particular email in question, but may help to "*complete understanding [and provide] better background info*" (P3). Some participants felt that the generation of keywords was not good enough to realise the potential value of KnowDis. They argued that the keyword generation "*algorithms need to be better*" (P11) to improve keyword quality (P2) and be more flexible when determining an appropriate number of keywords (P8).

A significant number of participants argued that the poor quality of the intranet search engine got in the way of allowing KnowDis to provide more relevant recommendations - "*Idea is great, but search too bad because of [intranet search engine]*" (P7). P10 noted that there is too much marketing and not enough technical information on the intranet, whereas P22 suggested that too many old files or multiple versions of files meant recommendations provided by KnowDis are less relevant. Another participant argued that the intranet search technology may simply be 'hopeless' when trying to retrieve meaningful results from more than one or two keywords (P21). Another rather unique problem raised by P15 was that as this participant worked remotely and was not connected to the company intranet via VPN for most of the time, KnowDis simply did not work as it could not connect to the intranet search engine to provide recommendations.

Performance

A few participants observed performance problems when using KnowDis, in particular that it made Outlook load slower (P12). However these problems seemed to be related to reportedly old machines some participants were using. P18 argued that if he "*had a newer, faster machine, would like to keep using KnowDis*". While P20 reported being initially concerned about performance issues when installing KnowDis, but did not notice any problems, P25 felt that "*[KnowDis works better later in the day, very slow in the mornings*". Only one participant reported an error within Outlook after "*8 to 10* *days*" of using KnowDis, but also noted that seemed to be related to a printer driver and possibly not caused by or related to KnowDis (P13).

KnowDis UI

Participants generally felt that KnowDis "*looks pretty good*" (P3), is not intrusive (P19), liked its user interface (P9, P12, P25) or its usability - "*usability-wise I like it a lot*" (P14). While P1 commented on liking KnowDis' usage of space, and how it provides "*gist of content in a small space*", P14 and P19 in particular liked and used KnowDis' preview mode quite a bit. Two participants however found some of KnowDis' functionality hard to use (P21, P23). One participant further expressed concern about KnowDis grabbing too much of his attention (P7), when he does not want it to. While P7 reportedly found the generation of keywords from email useful as a means to utilise these terms for the retrieval of related content, this participant felt that resulting recommendations should not be displayed within Outlook, but elsewhere such as on a dedicated web site. Participants made good use of the ability to flexibly position the KnowDis add-in within Outlook or minimising it. A few participants preferred positioning KnowDis in the bottom-left of the Outlook window (P1, P7, P22), others preferred having KnowDis on the right side (P6, P11, P24). The absolute position of KnowDis within Outlook partly depends on how participants have set up the rest of the Outlook UI, primarily displaying an email preview window or not, and whether below the email list or to the right side of the email list. Interestingly, two participants who had Outlook set-up such that email previews display to the right side of the email list, opted for displaying KnowDis right in the middle of Outlook, between the email list and the email preview area (P19, P23). Both participants reported having a large screen and preferring this layout and did not mind the screen real-estate KnowDis occupies in the Outlook window. In contrast, participants with reportedly small screens felt that KnowDis took too much space (P18) and consequently kept it minimised most of the time, only bringing up KnowDis when needed (P12, P15, P17, P22).

Use of Additional Functionality

Many participants reported not making any or barely any use of customisation features within KnowDis. The reasons given for the lack of customisation usage varied from being unaware of the options (P22), or not wanting to customise (P7) to being "*happy with settings*" (P24) and not making adjustments because recommendations were generally useful already (P19). In some other cases, motivations were more complex. P1 did not make use of customisation features as all recommendations were in English and he was looking for recommendations in Japanese. Another participant felt that making customisations manually would not be sufficiently effective as his requirements constantly change, thus requiring a more flexible 'profiling' of recommendations based on chaining requirements (P11). Some participants reported making use of the customisation features (P2, P12, P14, P25), P14 even made active use of the stop list customisation by adding terms that were not relevant keywords in his opinion to improve recommendations.

Search Domain Customization

The majority of participants used KnowDis to receive recommendations from the organisation's intranet - the default search domain setting - either primarily or all of the time. While P21 did not notice the ability to switch between different domains to receive recommendations from, P20 consciously preferred internal recommendations as he "*doesn't need web search much*". The intranet domain setting allowed further customisation of specific internal repositories to be included or excluded from recommendations. While P25 did experiment with adjusting these repository settings, he reportedly re-selected all repositories later on as "*it might be all relevant*". In contrast, P17 found the ability to only select and search specific repositories very useful.

The web domain setting was used by a minority of users. This can be attributed partially to the habitual reliance on Google as first choice for any web-based information retrieval as expressed by P16, who is "*used to rely on Google*". However, the web domain setting did allow for further customisation to retrieve recommendations from a selection of specific web sites. Some participants appreciated this ability such as P14, who did use it during the study, and P9, who planned on making use of it after discovering the functionality during the post-study interview. The email domain setting was used only by a small number of participants. While a few participants specifically expressed that receiving email recommendations through KnowDis "*would have been useful*" (P5) or "*would help more*" (P17), this functionality required Google Desktop software to be installed, which did not seem to be the case for many participants during this second field study. For some participants who could and did make use of email recommendations, it was the most used (P22) or most important (P23) feature.

Keyword Adjustments

KnowDis allowed participants to customise keywords to improve recommendations. However, a large number of participants did not use this feature. While some reportedly never tried (P1, P2, P4, P12, P13), others used it "*once in a while*" (P24), found it too much effort (P5, P21). Participants who made use of keyword adjustments generally seemed to report getting more useful recommendations (P14, P15, P17, P19). One participant reported not noticing an improvement in recommendations (P22).

Manual Search

Approximately a third of participants did not use the manual search functionality at all, and around half of participants tried manual search, but did not use it much. The lack of adoption of this feature for the majority of participants is primarily because participants are "*so used to do searches the used way*" (P12) and because the dedicated intranet search pages are deemed more effective (P2). However, a minority of participants did report habitually using the manual search feature - to retrieve recommendations for "*keywords that didn't show up in automatically generated keyword list*" (P14), to conduct general intranet and web searches (P24), and because it worked well (P19).

Adoption

A number of participants expressed a strong desire for continuing to use KnowDis as part of their email and information retrieval workflows. One participant thought it is "*pretty cool*" how KnowDis "*helps people to know where to look for information … from within Outlook*" (P12), while another one was "*very impressed*" with how KnowDis provided recommendations from email and web domains and "*integrated very seamlessly*" (P15) into his workflow. P15 more generally commented that "*it's encouraging that project like KnowDis exist, it's a great tool*". P20 argued that he was curious about what information KnowDis would be recommending in relation to email he was receiving. As he received approximately 250 emails per day from developers that mostly required some additional research this was particularly relevant to him. P20 reported that "*before KnowDis, used to go a lot to [intranet search website] to search*" to retrieve necessary information to respond to developer emails but now thinks that KnowDis will "*make life easier hopefully*" (P20). P11 has recently upgraded Outlook (from version 2007 to 2010), since then KnowDis doesn't appear anymore, and he didn't try to re-install – but if KnowDis would work properly in Outlook 2010, he "*wants to continue using it*".

One participant (P19) in particular made heavy use of KnowDis to get a better understanding of the inner workings of the organisation she joined less than a year ago. Before her participation in this study, she struggled to make sense of organisational structures in general, and the way knowledge repositories are organised. She said that she "*couldn't find information needed*" and that KnowDis "*has made search [within the organisation] easier*". She found the keywords generated by KnowDis "*better than the ones I created*" and that recommendations helped her "*understand what's going on in the organisation*" (P19).

<u>Adoption Challenges</u>

At the other end of the spectrum, some participants argued that their work habits are "*very engrained*" (P4) so KnowDis would have to work very well to have an impact on work-strategy. P4 also felt that "*It's hard to decide whether something is relevant*" (P4). However, P15 acknowledge that sometimes irrelevant recommendations are provided "*but that's how search works*" (P15). While P23 more generally commented that "*[KnowDis] doesn't offer what I need very well*" (P23), P18 argued that Outlook is already "*really full [with information]*", as he got a lot of archived email, and felt that KnowDis adding more information within Outlook was distracting.

7.10.4.3. Participants' wish list for KnowDis

When asked about any desirable changes or improvements to future versions of KnowDis, participants suggested a plethora of potential new functionality. The majority of the feedback can be grouped into three themes, namely user experience, personalisation, as well as extended integration with additional search providers, sources or tools.

<u>User Experience Improvements</u>

Some participants proposed additional layout functionality, such as displaying recommendations in a horizontal manner (P3), due to Outlook already having too many vertical columns (of folders, email lists, email preview, to-do bar, etc.). Another participant suggested to show recommendations no in a side bar, but as a second tab that would allow toggling between email and recommendations much like one would tab between different web pages in a browser (P6). P6 also suggested a tree-like hierarchical organisation of recommendations that would group recommendations by topic and improve the user's ability to get an overview of recommendations and then allow drilling down into a particular facet. A number of participants argued that recommendations should be displayed in the browser as "*that's what everyone is using*" (P7). As a potential solution to this approach, three participants independently suggested that a search for recommendations could still be triggered through KnowDis in Outlook, but would then allow users to take a search state from KnowDis to a browser-based UI, where more functionality could be exposed, more screen real-estate could be utilised and the user could explore recommendations more effectively (P11, P12, P14). P11 suggested the ability to augment the list of automatically generated keywords with additional manually selected ones to improve recommendations. In the current prototype, users could only choose a selection of automatically generated keywords or conduct a manual search from scratch. P7 wanted the behavior of KnowDis changed such that searches for recommendations would only be triggered manually via a dedicated button. P22 would like to see the implementation of infinite scrolling to allow for more recommendations to be automatically retrieved when scrolling down, rather than having to manually between pages. P23 suggested to replace the use of context menus, which require a right-button click, with the use of split button drop downs that allow accessing additional functionality via the left mouse button. P24 proposed to change the colour scheme to match Outlook. Finally, two participants suggested the integration of tooltips (P23) or other forms of tutorials (P10) that would improve the onboarding experience of new users. P10 proposed the use of tutorials tailored to specific roles in the organisation such as tutorials for software developers, for support engineers, and for sales or marketing staff.

Personalisation

Some participants requested tailoring recommendations, either automatically based on their company role (P4, P24), or more manually by allowing the user to add profile information in a similar manner like Linked-in (P4, P20). Other participants more generally proposed the integration of "*my site*" (P1) or "*my search*" (P3) functionality as a potential separate channel to retrieve tailored recommendations. Two participants suggested tailoring recommendations not based on a user profile but rather on the topic of an email (P6) or the nature of a task (P12). As a potential solution to this approach, P6 suggested allowing the user to associate specific keywords with appropriate knowledge repositories manually.

Integration with Search Providers

A number of participants wanted KnowDis to integrate with Windows Desktop search for recommendations of local files and email (P3, P17, P20). For web-based recommendations, P7 and P25 requested the ability to use Google search, rather than Bing. P25 in particular felt that in the current prototype, web search was not very useful because of its reliance on Bing. Another participant suggested the ability to discover "*other people with similar email*" (P9) to identify relevant people within the organisation.

Integration with Search Sources

A range of participants also requested the integration of additional search domains (P13, P18, P20, P23), the ability to manually add additional domains (P17, P20), or to allow for more specific targeting of individual sources of information (P16, P23, P24). Specifically, P20 and P23 expressed an interest in being able to receive recommendations of related bookmarks based on their bookmark collections stored locally. P13 and P18 would like to see the integration of KnowDis with a collaboration tool called versionone. This agile project management tool is crucial to both participants' workflow and being able to see "*what happened to my projects, when coming to work ... inside Outlook*" (P18) would be very useful. P17 and P20 requested the ability to add their own web sites (something already supported, but unknown to most participants), as well as shared folders or drives on the intranet. Other participants wanted to be able to pinpoint particular SharePoint project sites (P23, P24), or particular intranet wiki pages (P16).

Integration with Other Tools

Some participants argued that the integration of KnowDis with other tools relevant to their work flow would be desirable. For P5, having KnowDis integrated into Eclipse (a programming IDE) would mean he could receive programming-related recommendations, based on the code he is writing. P5 also suggested that integrating KnowDis with the to-do list within Outlook "*might be useful*". P15 thought that the integration of KnowDis into MS PowerPoint "*would be intriguing*". And P25 proposed integrating KnowDis with the organisation's service desk software so that employees who are writing a ticket would automatically receive recommendations of related issues.

Other Proposed Improvements

A few participants requested unified search functionality - the ability to see recommendations from different search domains, such as web and intranet, integrated in one view (P6, P14, P24). Two participants specifically asked for internationalization support within KnowDis such that recommendations in languages other than English (specifically Japanese) could be provided (P1, P24). P3's request for KnowDis to be able provide a link to an answer related to a question received in an email seems outlandish at first. However in the context of online services such as Yahoo Answers, or Quora, this idea may be a lot more feasible and potentially useful. P18 proposed the ability for KnowDis to provide recommendations about similar products from competitors. This could allow easy access to product information from competitors and enable product designers to be a step ahead of the competition, e.g. By coming up with better features. While this functionality is somewhat supported already by the current prototype via site-specific recommendations in the web domain, P18 was not aware of this particular customisation feature. However, this request illustrates an interesting use case of KnowDis tutorials tailored to specific roles within the organisation.

7.10.4.4. Usefulness of KnowDis

The third part of the post-study interviews concluded with questions about the general usefulness of KnowDis and whether participants remembered particular instances of successful or unsuccessful recommendations. Given the length of the field study plus the time it took to conduct 25 debriefing interviews, participants generally struggled to recall specific instances of good or bad recommendations in much detail. However, due to the time given to use KnowDis in their real-world work context, participants demonstrated strong opinions about its usefulness. Overall, participants' feedback on the perceived usefulness of KnowDis varied from "*isn't useful for me*" (P8) to "*changes the way I will work in the future*" (P3), with a number of participants somewhere in between those two extremes. From the feedback participants provided, we identified two main groups. The first group (14/25) – we call them adopters – embraces KnowDis to varying degrees. For some, KnowDis worked well, the way it was provided: "*this is great […] it really saved me time.*" (P25). Others embraced the concept of proactive recommendations, but thought that "*algorithms need to be better*" (P11). The second group (11/25) – we call them sceptics – was more or less cautious about the benefits KnowDis could provide them with. P5 and P8 expressed that KnowDis is not useful for them. P17 argued KnowDis is a "*wonderful tool, didn't adopt it, because used to work in a different way*".

Sceptics raised a number of reasons why KnowDis did not work well enough for them. In line with the results from the post-questionnaire, the largest proportion of scepticism derived from the perceived quality of recommendations, attributed to the search engines used. P9 said recommendations are "*not good enough*", and P7 argued that the "*idea is great, but search too bad because of [corporate search engine used]*". P8 thought too many keywords were used for the proactive search, and P17 suggested generating keywords from email subjects only. P21 summed up his impression, saying that the "*concept of KnowDis is a really good idea*", but that for KnowDis to work, a "*decent search engine is needed, which doesn't exist within [the organization]*". P7 and P25 thought web search was not good enough because of KnowDis' reliance on Microsoft Bing, rather than Google search. Proactive retrieval of recommendations caused some scepticism, as P7 perceived KnowDis "*a bit distracting*". P7 described his general work strategy as: "*I'm in research mode, design mode, analysis mode*", and perceived KnowDis proactive recommendations intrusive, trying to "*push him*" into research mode. P20 was rather torn, saying KnowDis has been "*helpful and distracting*", and that it "*requires a learning curve to not click too often on information*". P3 also preferred working in blocks (dealing with email, doing some 'actual' work, then dealing with more email), explicitly mentioning that he "*tries to avoid context switching*". Nevertheless he acknowledged that KnowDis provides "*useful hints*" - such as "*Wiki pages with interesting information*", and argued that KnowDis helped "*complete [his] understanding [of a matter]*". Sometimes, mundane reasons - such as an old, slow machine with a very small screen (P18) - made KnowDis distracting, due to performance problems and too much occupied screen real-estate. When asked how to change KnowDis, both P7 and P18 said they preferred getting related information not within Outlook but on a separate web site. P6 suggested displaying recommendations in a separate tab in Outlook, but not as a side bar.

'Not useful for me'-type Feedback

A few participants simply reported not discovering anything particularly useful among the recommendations received during the field study without being able to give a particular reason (P4, P8, P11). For some others, a range of reasons impacted the general usefulness of KnowDis, such as receiving bad recommendations to email being 'junk' (P5), a poor search engine being used to retrieve recommendations (P7), or "*too much sales-related information*" (P21) as part of recommendations. In addition, some other factors hint at more general adoption challenges, such as the onboarding process more specifically - "*didn't really take the time to take [KnowDis] seriously as a tool*" (P10) - as well as some participants' expectations more generally. While P1 was hoping KnowDis would be multi-lingual and work in Japanese, P9 was expecting valuable recommendations related to a random email that was not actually important to him. Finally, P5 suggested that he does not actually need more information. One participant, who reported only testing KnowDis to see how it works, argued that KnowDis is a "*wonderful tool, [but he] didn't adopt it, because used to work in a different way*" (P17).

Adopters generally appreciated receiving proactive recommendations within Outlook, integrating this functionality into their process of email handling. P1 used KnowDis' contextual information as means to better understand an email: "*gist of content in a small space*" (P1). P14 in particular "*liked the preview mode with more detailed information*", although he "*had to mess a lot with the keywords to make it more useful*", P14 argued that he now "*always go to [KnowDis] first [to look for relevant information]*". P24 "*definitely wants to continue using [KnowDis]*", arguing that the "*nice thing about it [is], it's just there*" and he "*doesn't need to start up browser first*". Despite concerns about the search algorithm, P15 acknowledged that "*sometimes irrelevant things [were shown], but that's how search works*" and that he was "*very impressed with what [KnowDis] does*". P15 suggested that KnowDis works particularly well for cross-referencing additional information and that it impacted his work-strategy, as it "*reminds that there is more information than what's in the email*" (P15).

'Very useful for me'-type Feedback

Some participants recalled specific bits of interesting information such as expense policies (P6) or performance review documentation (P15) upon receiving related email. P18 reported finding useful information related to development standards during an email discussion on a particular product. P3 suggested that when receiving technology newsletters such as from Microsoft or IBM, KnowDis might surface information on new technologies in its recommendations. However, others only recalled finding interesting information without remembering exactly what (P16, P20, P25). P19 suggested that "*[KnowDis] has really helped me to get relevant information […] be* *more aware of where information is*" and that she was able to find information with KnowDis that she could not find on the intranet itself. P20 argued that KnowDis helped him getting a better understanding as it "*brought up things I didn't know*" and found new internal repositories he did not think would be useful before. P20 in particular recalls an instance where he was working on his development machine (without KnowDis) searching for information on the intranet but could not find it. He then switched to another machine, where KnowDis was able to trigger the right information upon him opening an email related to the information he was looking for. P20 suggests that he is now using KnowDis for 25% of his information searches. Another participant reported that KnowDis "*is great … really helps [and] really saved me time*" (P25) and wants to keep using it as it changed the way she works.

7.11. Discussion

The KnowDis study investigated *how users respond to embedded proactive search in an email application* (RQ 4) as part of a two-year research collaboration with CA Labs of CA Technologies. Two field studies were conducted focussing on a more qualitative subjective and longitudinal investigation into the impact of KnowDis on knowledge workers in an IT organisation. via usage monitoring, pre- and post-study questionnaires and interviews. The KnowDis prototype was iteratively developed as an add-in for Outlook, proactively recommending documents related to the current context as defined by the email being read and to serve as a mechanism to embed just-in-time information discovery into knowledge workers email-related activities. Recommended documents were primarily available via a corporate intranet; however the capability for both desktop and web search was also incorporated in the tool. This investigation of RQ 4 was addressed using three specific research questions, which are reported in the following sections.

7.11.1. Is KnowDis Useful?

RQ 4.1 Do users find having a proactive search tool (KnowDis) embedded in their email application useful or not? If not, why not? If yes, how is it useful and how does it integrate with their day-to-day work?

The post-study interviews revealed that users' reactions about the usefulness of KnowDis were divided. While some perceived KnowDis not to be useful, others felt that it would change the way they will work in future. The differences between those two groups of users, referred to in the study as adopters and sceptics are discussed in detail in section 7.10.4.4. It is worth noting though that within the group of sceptics, the main concern was not the underlying concept of embedding just-in-time information discovery through the proactive search tool. Instead, the main problem was the perceived quality of (and thus trust in) the underlying search engine used to serve information from the corporate knowledge repositories. Within the adopter group, one example that stood out was a new employee, who perceived KnowDis to be particularly useful. This participant was quite concerned towards the end of the study about loosing the ability to make use of KnowDis (again this is discussed in more detail in section 7.10.4.4).

Results demonstrate that the majority of participants preferred having information provided in the context of their work (as opposed to having to stop reading their email in order to search for information) - the biggest challenge being quality of search and keyword generation.

7.11.2. Is KnowDis Distracting?

Do users find proactive search features distracting? If yes, how is it distracting? What can be done differently to make it less distracting?

One group of participants – referred to as adopters – generally appreciated receiving proactive recommendations within Outlook, and integrated this functionality into their process of email handling. On the other hand, the other group of participants – referred to as sceptics – perceived KnowDis to be somewhat distracting or intrusive.

As the potential intrusiveness of KnowDis to some users was considered during its development, the ability to minimise the KnowDis sidebar was included. Through the interaction data collected, we observed that some participants decided to close the KnowDis sidebar for most of the time. Those participants would describe their work strategy as a working in different modes, which they actively switch between. As a result, one participant perceived the proactive search features as nudging him into research mode.

Based on all interviewees' feedback on work-style (e.g. dealing with email, and information search), we observed that a rather large proportion of sceptics seems to prefer working in a more structured and well-planned way. For this group, proactive search might not be a useful feature; rather an optional "button to trigger search manually" might be the default option. This group is reminiscent of the *non-encounterers* identified by Erdelez (Erdelez, 1995). Non-encounterers are people who report that they seldom acquire important information through accidental or incidental counters.

In contrast, participants demonstrating a more flexible work-style, constantly switching between tasks, email and information search fell mostly in the group of adopters. Those participants in particular embraced the serendipitous nature of KnowDis, as it "*reminds that there is more information than what's in the email*" (P15) and one "*can already see list of what's out there*" (P25). These participants are reminiscent of the *encounterers* or *super-encounterers* identified by Erdelez (Erdelez, 1995). Encounterers recognize that they often "bump" into information. Super-encounterers recognize these accidental encounters as an important form of information acquisition.

Some adopters utilised the related information provided by KnowDis to evaluate the context and relevance of an email before actually shifting their attention to the actual content of that email. Those participants used KnowDis to aid them in categorising and prioritising email, and even preferred KnowDis centrally within Outlook in-between the email list and the actual email message preview.

7.11.3. Is KnowDis Improving Efficiency?

Do users think their work-related tasks that depend upon information search become more efficient and more effective when proactive search tools are available?

As the analysis of the pre- and post-questionnaires found (see section 7.10.3.1), participants found searching for information within the corporate knowledge repositories significantly more efficient, when having KnowDis available within Outlook. They also found it significantly easier to find relevant information from another business unit. The post-hoc exploration of the responses to *"I feel overwhelmed by the amount of information I need to search for in order to perform my job"* revealed significant age-related differences, where participants over 35 years old did feel significantly less overwhelmed by the amount of information needed to search when having KnowDis available. More importantly, this difference between groups of users is another reminder – as surfaced in Main Study 1 for novice vs. expert users of an FMI prototype – that not all users are alike and that it is important to explore why particular users like or dislike, find useful or useless, a particular product feature or service.

While not all participants found the specific KnowDis implementation of embedding just-in-time information discovery into their real-world tasks useful or unobtrusive, this was to be expected. Any tool, particularly prototype implementations, will encounter users, who for one reason or another do not like certain features, work in a different way, or simply have tasks that do not require a particular tool. However, looking at the bigger picture of all the collected and analysed user feedback, from the interaction data, questionnaires and the interviews, the KnowDis study clearly demonstrates that the developed proactive knowledge discovery tool can improve the email experience and reduce the sense of information overload of knowledge workers.

Some participants mentioned other potential uses of KnowDis during the interviews, such as using KnowDis to proactively monitor an organisation's competitors. One participant highlighted potential benefits of KnowDis to developers by *"allowing easy access to product information from competitors"* and to *"enable designers to be a step ahead of competition, e.g. come up with better features"* (P18).

Finally, a few participants showed genuine concern about losing the benefits provided to them through KnowDis after the study, which demonstrates some real-life impact on employees' work-strategies.

7.11.4. Limitations of Main Study 2

7.11.4.1. Perceived Ability of Search Providers

Participants experience with corporate search engines, and familiarity with using Google on a daily basis created preconceptions that impeded a neutral assessment in many cases. Participants' reliance on Google for much of their information search shaped their expectation for ease of search, relevance, number of results, etc. In their experience, internal search engines perform poorly compared to Google; because

KnowDis used an internal search engine, these participants did not expect KnowDis to perform well.

For example, the web search functionality integrated into KnowDis using Bing allowed manually searching or retrieving recommendations from a range of web sites simultaneously (e.g. using a list of favourite web sites). Although this type of search is potentially very useful, its concept was unfamiliar to participants, presumably because such searches are not supported in Google. After explaining this functionality to interviewees after the study, a number of them wanted more time with KnowDis to try out this feature.

7.11.4.2. Perceived Usefulness of Keywords

On the one hand, some participants did not expect automatically generated keywords to be useful. P2 explained KnowDis' perceived failure to generate useful keywords with his email, because his email is "not keyword rich". P5 perceived much of received email to be 'junk', thus he did not expect KnowDis to provide much useful information.

On the other hand, effective keyword generation from email is challenging. A large amount of corporate email is rather process focused, general communications or work coordination, and not information rich. Future keyword extraction algorithms used in KnowDis or similar systems might need to firstly categorize an incoming email (similarly to how Google Mail now distinguishes between important, social, promotional, updates or forum emails) to apply the best keyword extraction strategy. This goes in hand with the need to better understand the task an email is related to. In contrast, P19's very positive perception of KnowDis' automatically generated keywords is probably rooted in her being new to the company and not having adopted corporate terminology yet.

7.11.4.3. Ability of Personalisation

To make personalisation in a tool such as KnowDis useful, it needs to support dynamic adjustment of search domains and provide federated search results. It will further require vast connectivity to an array of search providers that capture the essence of employees work environment. For example, P5 expressed the need for KnowDis to integrate more tightly with "his tools" to make it more useful – such as Outlook's to-do list, Eclipse IDE, and Windows Desktop Search. Participants also argued that the display of combined results (from more than one search provider) would be highly desirable.

8. Conclusions

8.1. Introduction

Each study chapter contains a discussion section interpreting the results, drawing conclusions and considering limitations of the respective study. This chapter will not repeat those discussions in detail, but rather offer summative reflections on the overarching research question: *How to design just-in-time information services to improve the user experience of goal-driven interactions with information*?

8.2. The Research Findings

8.2.1. How to design just-in-time information services to improve the user experience of goal-driven interactions with information?

In this thesis, the iterative design, development and evaluation of two prototype systems – FMI and KnowDis – explored how the design of just-in-time information services might improve the user experience for goal-driven interactions with information.

Through the studies reported in this thesis, both prototype systems have demonstrated improvements to the user experience for the respective domains and contexts of use they were designed for and evaluated in. The FMI prototype demonstrated that it can be mapped onto reasonably large information spaces, and through dynamic contextualisation of its navigational elements provide significantly better task performance for information search tasks than a traditional blog-based web interface (see section 0). The KnowDis prototype demonstrated that it can be successfully used by knowledge workers as alongside their day-to-day work and make work-related tasks that depend upon information search more efficient (see section 7.11.3).

These findings are promising indicators that designing information services using the just-in-time information paradigm can improve users' information experiences. However, the prototypes built for this research and the studies conducted are very specific instances of the application of the just-in-time information paradigm.

Substantial future work will be needed to further explore applications of the paradigm in different domains, using different prototype implementations and different contexts of use.

In the following two sections, the specific contributions of the FMI prototype design (see section 8.2.2), and the KnowDis prototype design (see section 8.2.3) will be discussed. Subsequently, section 0 will discuss some more general contributions to the HCI community. And finally, section 0 will explore a range of directions for future work – some of which have already commenced.

8.2.2. Did the FMI prototype design improve the user experience of goal-driven interactions with information?

The Focus Metaphor Interface (FMI) prototype iteratively evolved throughout the preliminary studies as well as Main Study 1. With Main Study 1, the FMI prototype reached a sufficient enough representation of a just-in-time information interface that could investigate *the effects of interaction-driven dynamic updating of contextual elements in a Focus-Metaphor Interface (FMI) on task performance and user preference; and how does user interaction behaviour differ* (RQ 3).

At the point of Main Study 1, the FMI prototype enabled the user to interaction with information for the purposes of satisfying an information goal by maximizing the amount of information relevant to the specific information goal, as well as minimizing the amount of information extraneous to the specific information goal. The FMI prototype achieved this by providing the user with relevant information in a pro-active and contextual manner as per definition of the just-in-time information paradigm.

<u>In short, the FMI prototype design did improve the performance aspect of the user experience for all participants and improved the usability aspect of the user experience for novice participants. However, the FMI prototype design seemed to be less effective and usable for expert participants.</u>

More specifically, participants using the FMI prototype made significantly fewer errors and completed information search tasks in significantly less time, than participants using the traditional blog interface. Consequently, from a performance perspective only, the FMI prototype design did significantly improve the user experience of users' goal-driven interactions with information.

The results for user preference were somewhat less conclusive. Overall, significant differences could neither be found between the FMI prototype and the traditional blog interface through a usability questionnaire (see section 6.5.2.1) nor in the direct comparison (see section 6.5.2.2). However, the contrasting user feedback on the most positive and most negative aspects of each UI, as well as a more detailed analysis of the user feedback by computer expertise, suggest that the FMI prototype design may be more appropriate for some users than for others. It may also be more appropriate for some contexts of use than for others.

Specifically, as observations lead to the assumption that the FMI is more beneficial to novice users than it is to experts, a contributing factor could be the fact that experts have more strongly formed mental models of how to engage in tasks of information search or information exploration on the web, and that they require an interface to more strongly meet their expectations. However, any lab-based testing of a new UI concept, using a somewhat abstract context is likely to meet the same challenge when comparing novice and expert participants.

It seemed thus desirable to evaluate the use of just-in-time concepts in a more real-world context and in a more longitudinal evaluation that removes any bias due to novelty. Main Study 2 (Chapter 7) addressed this issue, by evaluating the just-in-time concept as part of a field study with (expert) knowledge workers of an IT organisation over the course of several weeks using a the custom-built KnowDis prototype.

8.2.3. Did the KnowDis prototype design improve the user experience of goal-driven interactions with information?

Main Study 2 investigated *how users respond to embedded proactive search in an email application* (RQ 4). This study was conducted as part of a two-year research collaboration with CA Labs of CA Technologies, in which the ideas and concepts developed earlier in this research (and reported in the previous studies) were applied to a corporate environment and tested on knowledge-workers and their real-world information tasks. As part of this collaboration, field studies were conducted with a focus on a more qualitative subjective and longitudinal investigation as to how users perceive and react to the integration of just-in-time information interfaces into their primary (work) activities. For this investigation, the KnowDis prototypes were developed to serve as the mechanism to embed just-in-time information discovery into knowledge workers email-related activities.

In short, the KnowDis prototype design did improve the user experience for participants overall by making work-related information search more efficient. However, while the KnowDis prototype design was useful for some knowledge workers and even integrated seamlessly into their day-to-day work, it appeared to be less useful and even distracting to others.

This apparent conflict in knowledge workers perceptions of KnowDis should be seen less as a weakness in the design of the KnowDis prototype – for not appealing to all participants – but rather as a strength of the conducted field study – for surfacing the naturally existing differences in people's work styles. Yes, a more advanced KnowDis prototype or future system inspired by KnowDis should take into account different work styles and as much as possible cater for them. For example, differences in work styles across knowledge workers suggest that real-world deployments of proactive search features should be user-controllable – e.g. offering an on/off switch – or even more advanced customizations.

This research also confirms and extends the conclusion reached by Gupta et al. (Gupta, Sharda, Ducheneaut, Zhao, 2006) and Billsus et al. (Billsus, Hilbert & Maynes-Aminzade, 2005) that email overload and resulting work stress is more dependent on the complexity of tasks linked to the processing of particular emails, rather than the number of email being processed (see section 7.10.3). As outlined above, KnowDis provided an approach to lessening feelings of overload by enriching email with local, enterprise and internet context through proactive information search.

8.3. Contributions to HCI

The research in this thesis has demonstrated that the just-in-time information paradigm can inform the design of information systems and services, which improve the user experience of goal-driven information interactions as outlined in section 8.2. The following sections discuss some further contributions of this thesis to the HCI community.

8.3.1. The Just-in-time Information Paradigm Contribution

The just-in-time information paradigm has been formulated as part of this thesis to help articulate and frame the theoretical and conceptual work that went into the design of the FMI and KnowDis prototype systems. As a theoretical contribution, the just-in-time information paradigm is intended not as a rigid framework of rules to design future information systems, but rather as an artefact contributed to HCI research to encourage more outside the box thinking – much like it is promoted in design research focused on solving design problems. This artefact intends to encourage us to challenge the status quo – that is how we perceive keyword-based search as the primary, and often only viable mechanism to cope with our ever increasing and ever more complex information needs.

In Main Study 1, comments by participants, such as *"can't browse"* are revealing in the sense that the interaction paradigm underlying the FMI is a different mode of interacting with information, it's not keyword-based search, nor is it 'traditional' browsing as such, but rather a way to guide the user through a topic. The FMI was conceptualised and designed to aid users with complex information tasks, which seems to be reflected in some of the other comments such as *"no time to lose interest"*, *"not in control"* or *"very precise, more information for a topic"*. Many existing studies around web-based search focus on optimising the use of traditional search engine result pages (SERP) (e.g. Huang & White, 2011).

But the shortcomings of SERP-based information retrieval are increasingly well-known and accepted (e.g. Chierichetti, Kumar & Raghavan, 2011). The reliance on user-created keywords and optimisation of search algorithms for simple fact-based information retrieval means that much of the research on SERP interaction typically utilises very narrow and simple factual search tasks, which does not reflect the true range of information needs of internet users. As a result, a growing line of research into sensemaking (e.g. Russell, Stefik & Pirolli, 1993; Weick, 1995; Stefik, 2004) both acknowledges and tries to tackle more complex information needs.

The instantiations of the just-in-time information paradigm designed, developed and evaluated as part of this research – the FMI prototype and the KnowDis prototype – are just modest initial steps towards an exploration of information systems and services that more effectively support information foraging and eventually sensemaking processes in our day-to-day struggles with information overload. It is worth noting again that the just-in-time information paradigm does not intend to invent anything new really, but rather – in the spirit of a traditional design process – takes existing concepts, methods and technologies to combine them into something else, something that eventually becomes more than the sum of its parts.

8.3.1.1. Proactive Information Delivery

The results for evaluations of the FMI and KnowDis prototypes both revealed strong individual differences between users. These differences are likely linked to different preferences in work style – partially influenced by factors such as computer expertise, age, workload, etc. In hindsight, it thus seems obvious that the pro-active provisioning of information works well for some but less so for others. It highlights that more research is needed into refining mechanisms for pro-active information delivery.

The fundamental idea of the just-in-time information paradigm is that information should be made available at the right time, when the user will perceive it as potentially relevant. In the KnowDis studies, this was the case for participants belonging to the group of adopters. But for the group of sceptics who work in modes, just-in-time simply means something else than was provided by the specific KnowDis prototype tested. Based on the observations and user feedback gathered in this thesis and considering Erdelez' work on non-encounterers, encounterers, and super encounterers of information (Erdelez, 1995) it may be worthwhile to explore three types of just-in-time information provisioning:

Instant push: Much like has been the focus of this research, instant push is aimed at super encounterers and provides just-in-time information immediately alongside a given user context.

Scheduled push: The concept of scheduled push could follow the notion that some users (e.g. encounterers) prefer receiving recommendations in a more scheduled fashion – say at the beginning of the day, after lunch, or at other times when they intend to be in 'research mode'. Information would then be provided just-in-time for the user entering 'research mode' and utilize a larger set of previously attended to information to provide a recommendations based on a broader context.

Pull-only: The concept of pull-only information recommendations is aimed at users who always want to be in control (e.g. non-encounterers) and prefer attending to information only when they desire. For such users, recommendations could be made available on a service such as a web application (e.g. personalised news apps work in a similar fashion), where recommendations are again based on a continuously larger set of previously attended to information.

8.3.2. The GazeSpace Contribution

As technology advances, and input modalities change, the just-in-time information paradigm should support the design of interfaces that can cope with those advances. The surge in multi-touch devices over the last few years is an indicator for the speed at which fundamental changes in human-information interaction can and will occur. Small children that grow up today surrounded by tablet devices and smartphones experience touch as the most natural way of interacting with digital information. The use of a keyboard and mouse is a much more abstract concept to adapt to for this generation. Augmented reality devices like Google Glass, or virtual reality headsets like Oculus Rift are the most recent examples of how the interface between humans and technology is evolving. The effective use of those technologies for more than email notifications or immersive gaming requires new paradigms for human-information interaction using those modalities.

Since this study, both algorithms have been included in a comparison study on gaze-to-object mapping algorithms (Špakov, 2011). GazeSpace is also referenced in the chapter *"Computer Control by Gaze"* (Skovsgaard, Räihä & Tall, 2011) of the book *"Gaze Interaction and Applications of Eye Tracking"*(Majaranta, Aoki, Donegan, *et al.*, 2011). This study did not compare the GazeSpace system with traditional mouse-based input, as the main aim has been to investigate initial user responses to the GazeSpace prototype and the developed algorithms for dynamic activation times. That said, as eye-gaze technology is becoming more mainstream, comparing subsequent iterations of the GazeSpace system to alternative input modalities may be desirable. As the most common complaint mentioned has been the speed of interaction, future work should test lower activation thresholds, as one participant noted that *"if faster [GazeSpace] would be fun and relaxing"*. This also provides some insight into why more participants preferred the SIA algorithm, it simply resulted in faster activations (as it did not have a interest decay function).

For the reported GazeSpace study, longer activation thresholds have been consciously chosen to prevent high error rates. These could have negatively biased participant's general evaluation of this novel interaction technique. When using static dwell times for selection, speed of interaction and error rates will always influence each other. To compete with traditional input techniques, future eye gaze systems might require adaptive selection algorithms, which can be tailored to individual users and task complexity. It also appears crucial to experiment with different settings for gaze-interest-thresholds and timing values for the decay function in future studies.

8.3.3. The Knowledge Worker Research Contribution

In their work at PARC, Convertino et al. (Convertino, Kairam, Hong, et al., 2010) argue that "there is no silver bullet for addressing information management problems for knowledge workers", and that instead, a range of classes of knowledge workers exist - with distinct informational needs. Their work on a "Cross-channel Information Management Tool For Workers in Enterprise Task Forces" (Convertino, Kairam, Hong, et al., 2010) is based on earlier work concerned with the "Design of a Task Force Workspace" (Convertino, Hong, Nelson, et al., 2009). While their early work on supporting expert knowledge workers through facilitating 'activity awareness and social sensemaking' (Convertino, Hong, Nelson, et al., 2009) is primarily focused around a web-based workspace, Convertino et al.'s later work pays much more attention to email (with 29 references to 'email') as a crucial content management tool for knowledge workers (Convertino, Kairam, Hong, et al., 2010): *"Email emerged as the central tool for managing and exchanging information, transferring documents, and coordinating; as such, it is critical that any new system either incorporates email or at least offloads some information normally transmitted via email".*

Convertino et al. reference early work on KnowDis (Laqua, Sasse, Gates, *et al.*, 2009) as an *"email plug-in that aggregates relevant information from diverse corporate sources"* (Convertino, Kairam, Hong, *et al.*, 2010). The authors go on positioning their 'Workstreams System' prototype as a *"possible solution to alleviate the email overload problem, which arises from making email the hub of every transaction"* (Convertino, Kairam, Hong, *et al.*, 2010) and explicitly note that *"this solution is different from those that have attempted extending functionalities of email"* - those being (Laqua, Sasse, Gates, *et al.*, 2009).

The influence of the KnowDis project on the design of the Workstreams System at PARC has its roots at the 2009 Sensemaking workshop at ACM CHI, co-organised by Ed H. Chi, co-author in (Convertino, Hong, Nelson, *et al.*, 2009) and (Convertino, Kairam, Hong, *et al.*, 2010), and at which (Laqua, Sasse, Gates, *et al.*, 2009) was presented.

The key difference between KnowDis and the work on the Workstreams System seems to be PARC's attempt to motivate change in the general usage of email, which is extremely hard to achieve. Google failed at attempting exactly that when introducing the now defunct Google Wave.

The main motivation within the KnowDis project was never to replace or redesign email. While this may be a valid design goal, when attempting to more literally apply the just-in-time information paradigm, it would have simply been out of scope of the project, both in time and in resources.

8.3.4. Research Process Contributions - Advice for Future PhD Students

This research applied an interwoven methodology combining lab-based experiments and longitudinal field studies, which collected a rich mixture of quantitative and qualitative data (see methodology matrix in section 3.3.1). This process integrated summative evaluations into a larger context of formative evaluations, in which each study created new insights that informed the on-going cycle of concept development, iterative design and further studies.

8.3.4.1. Benefits of the Research Process

The main benefit of the applied research methodology – the combination of lab-based and field studies – has been a vital a part to develop both a richer and more tangible understanding of how to tackle today's problem of information overload and ultimately led to the just-in-time information paradigm as outlined in section 1.3.3.

The conduction of a series of lab-based studies (chapter 4 to 6) has been a vital part in iteratively and formatively designing, developing and validating a tangible concept for minimalist just-in-time information interfaces. Without these studies, and the related prototype development, at best a set of design recommendations could have been provided that are grounded in thorough qualitative research. But those design recommendations would have to be quite abstract and would not be tangible to the extend that a prototype has been built, tested, and actually validated the design guidelines.

However, if only lab-based studies would have been conducted using the FMI prototype implementations, a lack of ecological validity would have impacted the framing of the just-in-time paradigm. Most likely, a different kind of conceptual paradigm based primarily on the minimalist visualisation techniques used in the FMI would have been formulated. Such a paradigm would in its own right have described a method for designing adaptive interfaces, which demonstrate superior performance for certain information tasks. However, this paradigm would have lacked the understanding on how to apply such interfaces in the real world and thus may not have found much application in the future.

As the field studies reported in this thesis are fundamentally based on the insights of the lab-based studies conducted earlier, it is not possible to truthfully assess the value of the field studies in their own right. The conception of research strand two (KnowDis prototypes), which lead to the field studies is a continuation from research strand one (FMI prototypes) and represents the application of the insights from research strand one to real life scenarios. The field studies contributed a much richer understanding of how just-in-time information interfaces can be put to work in real-world scenarios.

Prior to the conduction of these field studies, the primary focus of the research investigating the viability of the just-in-time information interface prototypes has been on comparing it with the existing modes of human-information interaction on the web - browsing and keyword-based search. This comparison infers that interacting with just-in-time information interfaces would be either like browsing or like keyword-based search, and potentially more effective in certain task contexts. Only after the conduction of the field studies, the insight emerged that the paradigm underlying just-in-time information interfaces is neither like browsing or keyword-based search, but rather a different type of human-information interaction.

8.3.4.2. Problems of the Research Process

The main problem of the applied research methodology has been one of complexity. It has been challenging to integrate design thinking, development effort, research study design and execution into the timeframe of a single PhD thesis. It also should be noted that the time taken to conduct both strands of this research - the lab-based studies around the FMI prototypes, and the field studies around the KnowDis prototypes - has exceeded the typical time for a PhD thesis. Besides the overall complexity of the process taken within this research, there have been some concrete challenges that I faced when deploying the KnowDis prototypes across a large IT organisation.

The aim was to evaluate KnowDis in field studies with actual knowledge workers, in real-life work contexts and over an extended period of time. This approach proved challenging in that it required to pass a number of technological and practical hurdles.

First, the prototype had to be high-fidelity so it could be deployed on employees' machines. A typical high-fidelity prototype is created such that it runs sufficiently well on a test environment either for lab-based testing, or on a test machine that is being used in the field for evaluation with an experimenter being present. But neither of these approaches would have allowed for a longitudinal evaluation in which the prototype to be assessed was actually becoming part of participants' daily work routine. Moreover, in the case of this study, participants were distributed globally - some in the US or Canada, some in Europe, others in India, Japan and Australia.

Second, deployment of the prototype to participants of the field study required liaising with the organisation's IT security and networking departments, to ensure the application would be safe and not incur unreasonable loads on the network infrastructure. Given that the organisation has over 10,000 employees and primarily develops enterprise software, the study approval process was expectedly thorough and time-consuming.

Third, the prototype had to run stably on a variety of combinations of MS Windows operating system and MS Outlook versions. While this was one of the key challenges with an early version of the prototype, the utilisation of Add-In-Express, a framework for building Microsoft Office extension, helped to significantly boost the stability of the prototype across participants' computer environments for the second field study.

The severeness of these challenges, but also the inherent value in tackling them may be best summarised by quoting reviewer three of the paper published at CHI 2011 (Laqua, Sasse, Greenspan, *et al.*, 2011):

> *"A piece of software that is more than demoware, but is something that can be used to support authentic tasks in real work environments. The fact that the authors deployed their tool demonstrated the overall feasibility of these tools to make their way into enterprise use. As a research also at a corporate lab, this is hard to accomplish and deserves high recognition."*
>
> *"A great study of the system's use in practice. This, to me, is the biggest contribution of the work. The design of the field study hit a lot of important requirements for me: 1) the experiment was conducted over a long enough period for users to really appraise the value of the tool and settle into a realistic pattern of use; 2) the users used the tools to support their real life work tasks; 3) no expectation of use was placed on users. The study produces some nice insights into how tools like this should be further refined, as well as insight about the overall utility in supporting knowledge discovery tasks."*

8.3.4.3. A Call for Studies with More Participants

Normally, eye tracking studies are conducted with 20, maybe 24 participants - or more precisely 10-12 participants per condition tested. This number is typically the threshold at which significant differences in terms of the eye tracking data can be found. Many eye tracking studies also focus purely or at least primarily on the analysis of the gaze data, at the expense of a detailed collection of user feedback through questionnaires or interviews.

Main Study 1 was conducted with 60 participants, or 30 participants per condition and also included fairly detailed questionnaires collecting usability feedback, demographic and computer expertise data. The detailed post-hoc analysis of Main Study 1 is therefore a novel contribution. It provides a much richer picture than typical lab-based evaluations of new software prototypes, which often report on user feedback in a one-dimensional manner and utilise a very homogenous group of participants. While it may not be feasible to always test new software prototypes with larger groups of a diverse range of participants, the findings in Main Study 1 do suggest that user satisfaction evaluations in studies with small and homogenous groups of participants do provide little value.

Another facet setting apart the participant group in Main Study 1 from many other eye tracking studies is the diversity of participants used. Participants were recruited via the psychology pool, a web site run by the University College London's Department of Psychology to advertise research studies to a large group of registered users. The resulting mix of participants was very diverse, both in age range and computer literacy, whereas in typical eye tracking studies, participants may be solely drawn from a student population or employees of a technology company.

8.3.4.4. A Call for More Field Studies

It is one thing to build a system and test it in the lab in a very controlled environment, using self-selected tasks and on a typically homogenous audience. It is an entirely different thing to build a system and let it out into the wild to evaluate its use in an uncontrolled environment. If HCI research does not tackle this problem, it is left to the companies building new products, like start-ups that are typically ill-equipped regarding knowledge of HCI best practices, methods and so on. At best, this kind of approach might be tackled by R&D departments of large IT organisations, but how many of these exist? Maybe a potential alternative could be a more thriving culture of collaboration between HCI researchers and software start-ups to support the creation of new tools following user-centred design best practices, evaluated on actual pain points of users in the real-world such that the tools being conceived, designed and built will eventually make a difference.

Another aspect noticed when conducting the KnowDis field studies was that interviews with knowledge workers indicated that information overload is a barrier to adopting new tools that might help reduce their overload. This is a vicious cycle, where employees continue to be overwhelmed, yet cannot take the time to learn new tools or techniques that would help improve their situation. Some participants found the tool useful to their work, selecting the recommended links. However some participants were so overloaded with work, that they found it challenging to adapt a new tool into their workflow. Future versions of KnowDis and other tools intended to knowledge management must address this challenge.

How to Support Field Studies

The workload of participants should be considered for an effective evaluation of work-related tools in the field. Furthermore, it is essential to ensure all participants have the necessary understanding of the tool to be evaluated. The provision of user guides (PDF/Wiki), or video demos might not be enough, and the use of 1 to 1 on-site or remote meetings might be advisable.

8.4. Future Research

The main goal of this thesis was to evaluate *how to design just-in-time information services to improve the user experience of goal-driven interactions with information*? The developed FMI prototype represents a conceptually close interpretation of the just-in-time information paradigm. Its minimalist design may require augmentation for more widespread real-world use. Such augmentation can be classified along a range of future avenues for research. Within human-computer interaction, the use of context could be studied in much greater detail, to investigate research questions such as:

1. How does the number of context items being displayed relate to a) different types of users, and b) different types of information search and exploration tasks?
2. How does the layout of context items being displayed affect performance and user satisfaction, again related to different types of users and task types?
3. How can the level of detail of information displayed in the context items be augmented such that it is optimal to different types of users and task types?

Within machine learning, the specific just-in-time information interface prototypes as well as the general paradigm provide a suitable platform for an in-depth investigation of more computational research questions, such as:

1. Which models of personalisation are most suitable to just-in-time interfaces?
2. How can user modelling become an integral part of a continuous personalised information experience using just-in-time interfaces?
3. How can automated tagging be improved through technologies such as concept modelling, or related methods?
4. How does search technology need to be altered or optimised differently to work effectively as part of the just-in-time paradigm?

Within the context of educational sciences, the potential benefit of just-in-time information interfaces as a tool for focused learning could be explored. The support for such interfaces may be particularly beneficial for learning activities in the context of activities that require the consultation of a lot of background information, source material, etc. As learning materials are typically domain specific and may even contain rich meta data that could be re-used for effective personalisation, exploring the potential of future educational research utilising just-in-time information paradigm for learning or the creating of learning environments could be promising. Given the recent trends to massive open online learning courses such as the Khan academy (http://www.khanacademy.org), or Udacity (http://www.udacity.com/), the need to for more effective online learning experiences should grow significantly over the next few years.

Within computer forensics, information discovery plays vital part, as typically vast amounts of data need to be analysed. While much of this analysis may be done semi-automatically by intelligent computer programs, there will always also be episodes of human inspection. Such inspections might still require the evaluation of large amounts of potentially relevant files. Future research could explore whether this process may benefit from the application of the just-in-time information interfaces.

While the FMI prototypes were built using Adobe Flash and related technologies, which represented the only effective means in the time 2005 - 2008 for such endeavour, today any new iterations should be built using HTML5, CSS3 and Javascript. Moreover, cloud-based computing will allow for much more effective backend setups, than the single old server used to power the various FMI iterations.

The following is an example for the use of modern languages to support the development of future iterations of just-in-time information interfaces. The application of the just-in-time information paradigm to intelligent in-car services is area of future research, which has already commenced during the process of writing up of this thesis. As part of the SafeTrip EU project, I worked on developing and evaluating an information discovery interface for an in-car tablet-based map application. The initial concept is described in (Beeharee, Laqua & Sasse, 2011). The first prototype is a touch-enabled web-app for tablet devices (see Figure 91), which has been tested as part of an in-car field study. Results have been published in (Beeharee & Vaccaro, 2011).

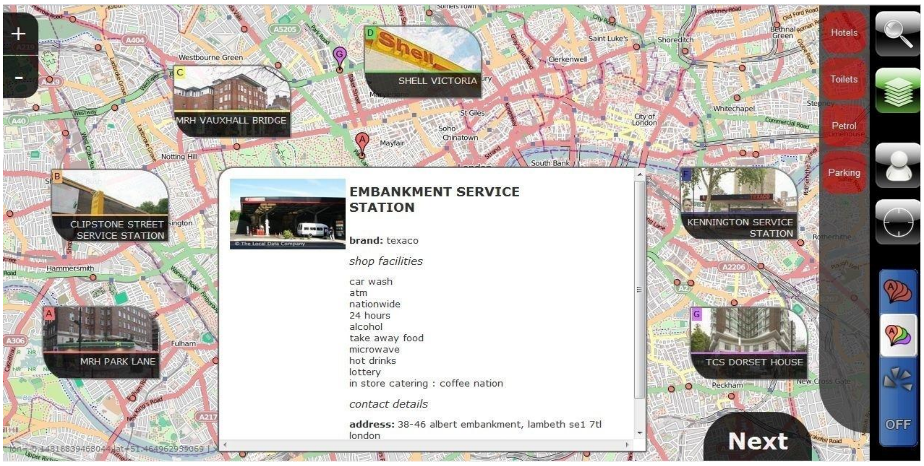


Figure 91: FMI prototype application for in-car information discovery

The first, most obvious scenario for follow-up research on the KnowDis prototypes in particular would be a study specifically designed to evaluate its usefulness with employees that are new to an organisation only. It is a well-known problem, that employees typically take months to familiarise themselves with the organisational structures of a large organisation, and more generally to find their way about how things are done. Future studies should investigate how embedding just-in-time information discovery into the evolving workflows of new employees could get them more productive in less time.

Future iterations of KnowDis could be evaluated in the context of web-based email clients such as Gmail for example through the development of browser extensions. The benefit of this approach would be the ability to recruit participants for field studies from a much larger audience. The development effort of a browser extension should also be significantly lower than the effort it took to build the KnowDis prototype. This approach would also allow reaching out to much more diverse range of knowledge workers that are part of technology start-ups, artists, creatives in agencies, etc. which typically rely on web-based email solutions such as Google Apps for Business (http://www.google.com/enterprise/apps/business/).

The concept of KnowDis further could and should be evolved into a general purpose context-providing mechanism that reaches beyond just the email context. As already discussed during interviews with participants of the KnowDis field trials, the concept of KnowDis may work well for work contexts involving general office applications such as text editors, spreadsheet program's, presentation tools, or even for programming tasks. Those usage contexts will offer a wide range of opportunities for future field studies to explore how the just-in-time information paradigm may benefit knowledge workers writing reports, analysing data, preparing presentations or writing code.

Finally, the other obvious next step to take is the integration of the KnowDis strand of just-in-time discovery interfaces with some of the more fundamental changes as derived from the FMI strand. The premise of this integration would be to connect the pro-active information discovery aspects embedded into specific work contexts with the more fundamental information exploration aspects of the minimalist just-in-time interfaces as embodied by the FMI prototypes. This concept was touched up on in the 2009 sensemaking workshop paper (Laqua, Sasse, Gates, *et al.*, 2009) through a design concept (see Figure 92), but was out of scope of the project in the end.

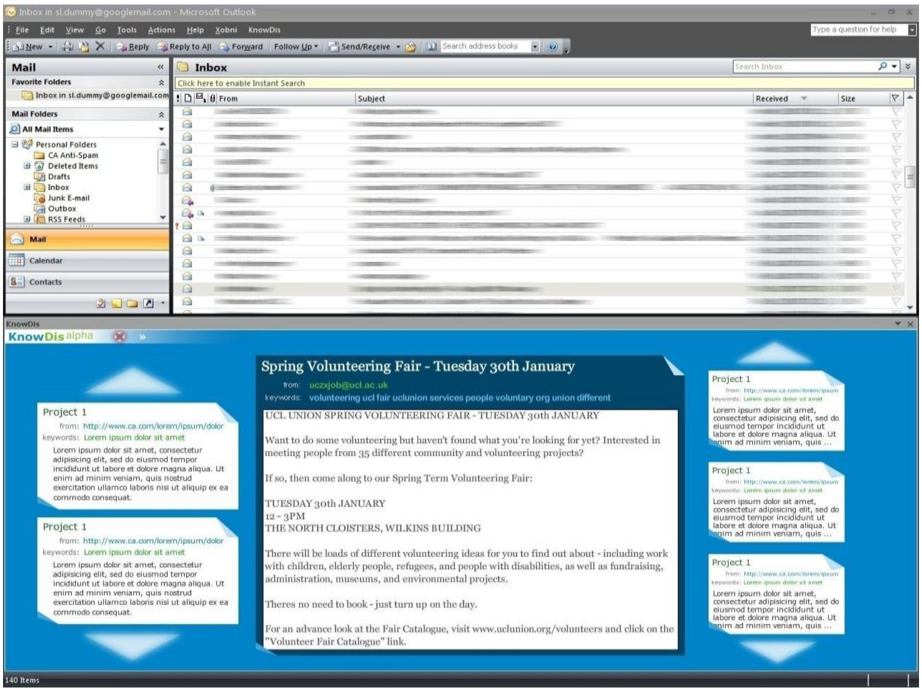


Figure 92: Alternative KnowDis concept for more integrated information exploration support

9. Glossary

Content Element (in an FMI)

The content element displays the currently selected article in detail and is displayed in the centre of the screen. As the user interacts with one of the contextual elements, the currently visible content element in the centre of the screen disappears, the contextual element selected changes its state to become a content element and then moves into the centre of the screen (see section 4.3).

Contextual (Navigation) Element (in an FMI)

Contextual interface elements form part of a Focus-Metaphor Interface and are arranged around the primary content element, which displays a selected article. The contextual elements function as navigation (activated through clicking) and provide previews onto the underlying content much like snippets on search engine result pages (SERP) (see section 4.3).

Dynamic Interest Decay (DID)

Dynamic interest decay (DID) is the description of a gaze-activation algorithm designed as part of this thesis. The DID algorithm collects gaze over each target area (navigational elements) until a predefined threshold is reached. The DID algorithm is a variation of the SIA algorithm. However, if the target threshold is not reached and attention turns away from the area, then the DID algorithm decreases the collected interest for each element based on a predefined timing value. This approach intends to model the user losing interest in the particular element (see section 5.4.3).

Focus-Metaphor Interface (FMI)

The FMI prototypes are the first strand of design, development and evaluation of the just-in-time information paradigm (the second being the KnowDis prototypes). In essence, an FMI is a minimalist user interface only containing snippets of information represented at different levels of detail. A traditional FMI would not provide any navigation that allows browsing all information contained within the UI but rather requires the user to explore information through interaction with contextual elements. Various FMI prototypes have been evaluated as part of this research (see chapters 4, 5, and 6).

Information Scent

Information scent is used to describe how well a user will be able to judge the value of information they will encounter when choosing a particular navigational option or path (see section 1.3.1.1).

Information Goal Continuum

Information Goal Continuum is term coined in this thesis. The user 'navigates' an information goal continuum when the iterative process of information foraging (awareness > understanding > decision) is sustained until an information goal is reached. When the user fails the reach their information goal by making wrong or ineffective navigation decisions – the user is leaving the information goal continuum. For a more detailed discussion, see section 1.3.2.

Just-in-time Information Paradigm

At the point of interaction with information through a user interface for the purposes of satisfying an information goal (1) the amount of information relevant to an information goal should be maximised, and (2) the amount of information extraneous to an information goal should be minimised. Further, the user should be provided with relevant information through the just-in-time information interface in a pro-active and contextual manner that does not require the user to manually articulate (e.g. through keywords) what they are looking for. Instead, the just-in-time information interface should facilitate serendipitous encounters with useful information in an anticipatory manner

KnowDis

The KnowDis prototypes are the second strand of design, development and evaluation of the just-in-time information paradigm (the first being the FMI prototypes). In particular, KnowDis describes a series of prototype add-ins for Microsoft Outlook to aid knowledge discovery in organisations that has been evaluated as part of a two year collaborative research project (see chapter 7).

Serendipity

Defined by the Oxford dictionary as *"the occurrence and development of events by chance in a happy or beneficial way"*, in this thesis, serendipity primarily relates to the chance encounter of beneficial information within an information system.

Static Interest Accumulation (SIA)

Static interest accumulation (SIA) is the description of a gaze-activation algorithm designed as part of this thesis. The SIA algorithm collects gaze over each target area (navigational elements) until a predefined threshold is reached. If the target threshold is not reached and attention turns away from the area, then the SIA algorithm will remember the collected interest for each element and continue increasing attention once the user returns to the element. This approach intends to support gaze-based interaction on elements that require the user to attend to them for longer periods of time and make sense of them before selection, for example by reading text contained in the element (see section 5.4.3).

10. References

A

Adams, A., & Cox, A. L. (2008). Questionnaires, in-depth interviews and focus groups. In: Cairns, P., & Cox, A. L. (Eds.), *Research Methods for Human Computer Interaction* (1st ed., pp. 17–34). Cambridge, UK: Cambridge University Press.

Adams, A., Lunt, P., & Cairns, P. (2008). A qualitative approach to HCI research. In: Cairns, P., & Cox, A. (Eds.). *Research Methods for Human-Computer Interaction* (1st ed., pp. 138–157). Cambridge, UK: Cambridge University Press.

Ahn, J.-W., Brusilovsky, P., He, D., Grady, J., & Li, Q. (2008). Personalized web exploration with task models. *Proceedings of the 17th international conference on World Wide Web (WWW '08)*. ACM, New York, NY, USA, 1-10.

Allen, B. (1994). Perceptual speed, learning and information retrieval performance. *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR '94).* Springer-Verlag New York, Inc., New York, NY, USA, 71-80.

Attfield, S., & Blandford, A. (2009). Improving the Cost Structure of Sensemaking Tasks: Analysing User Concepts to Inform Information System Design. *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2009).* Uppsala, Sweden, 532–545.

B

Babcock, J., Li, D., Parkhurst, D., & Winfield, D. (2012). openEyes. Retrieved from: http://thirtysixthspan.com/openEyes/

Bagnoli, C. (2011). Constructivism in Metaethics. In Zalta, E. N. (Ed.). *Stanford Encyclopedia of Philosophy.* Retrieved from: http://plato.stanford.edu/entries/constructivism-metaethics/

Bannon, L., Cypher, A., Greenspan, S., & Monty, M. L. (1983). Evaluation and Analysis of Users' Activity Organization. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '83)*. ACM Press. New York, NY, USA, 54–57.

Bargas-Avila, J. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM Press. New York, NY, USA, 2689-2698.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13 (5), 407 - 424.

Baudisch, P., Lee, B., & Hanna, L. (2004). Fishnet, a fisheye web browser with search term popouts: a comparative evaluation with overview and linear view. *Proceedings of the working conference on Advanced visual interfaces (AVI '04).* ACM, New York, NY, USA, 133-140.

Bederson, B. B. (2000). Fisheye Menus. *Proceedings of ACM Conference on User Interface Software and Technology (UIST 2000)*, ACM Press, New York, NY, USA, 217-226.

Beeharee, A. K., & Vaccaro, A. (2011). In-Car Contextualised Search on Tablets. *Adjunct Proceedings of 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2011)*, Salzburg, Austria, 171 – 175.

Beeharee, A. K., Laqua, S., & Sasse, M. A. (2011). Navigating Haystacks at 70 mph: Intelligent Search for Intelligent In-Car Services. *Proceedings of the 3rd International Multimodal Interfaces for*

Automobile Applications (MIAA), workshop at IUI 2011. Palo Alto, CA, USA, 25–28.

Benway, J. P., & Lane, D. M. (1998). Banner Blindness. *Internetworking.* Retrieved January 15, 2012, from http://www.internettg.org/dec98/banner_blindness.html

Beyer, H., & Holtzblatt, K. (1997). *Contextual Design: Defining Customer-Centered Systems (1st ed.).* San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Billsus, D., Brunk, C.A., Evans, C., Gladish, B., et al. (2002). Adaptive interfaces for ubiquitous web access. *Communications of the ACM*, 45 (5), 34–38.

Billsus, D., Hilbert, D. M., & Maynes-Aminzade, D. (2005). Improving proactive information systems. *Proceedings of the 10th international conference on Intelligent user interfaces (IUI '05).* ACM, New York, NY, USA, 159–166.

Boardman, R., & Sasse, M. A. (2004). Stuff goes into the computer and doesn't come out: a cross-tool study of personal information management. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04).* ACM, New York, NY, USA, 583-590.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Proceedings of the seventh international conference on World Wide Web (WWW '98)*. Brisbane, Australia, 107–117.

Brooks, C., & Montanez, N. (2006). Improved annotation of the blogosphere via autotagging and hierarchical clustering. *Proceedings of the 15th International Conference on World Wide Web (WWW '06).* ACM, New York, NY, 625-632.

Brown, B., & Reeves, S. (2011). Into the wild: challenges and opportunities for field trial methods. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, New York, NY, USA, 1657-1666.

Brown, M. H., Marais, H., Najork, M. A., & Weihl, W. E. (1997). Focus+Context Display of Web Pages: Implementation Alternatives. *DEC Systems Research Center (SRC) Technical Note 1997-010, 1997.* Retrieved March 15, 2012, from http://apotheca.hpl.hp.com/ftp/gatekeeper/pub/DEC/SRC/technical-notes/SRC-1997-010-html

Brown, P. J., Burleson, W., Lamming, M., Rahlff, O.-W., et al. (2000). Context-awareness: some compelling applications. In: *Proceedings the CHI 2000 Workshop on The What, Who, Where, When, Why and How of Context-Awareness* (The Hague, Netherlands, Apr 1 - 6, 2000). CHI' 00. ACM Press, New York, NY.

Brown, P. J., & Jones, G. J. F. (2001). Context-aware Retrieval: Exploring a New Environment for Information Retrieval and Information Filtering. *Personal and Ubiquitous Computing*, *5*(4), 253–263. doi:10.1007/s007790170004

Bruner, J. S. (1966). *Toward a Theory of Instruction* (1st ed.). Cambridge, MA, USA: Harvard University Press.

Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction,* 11 (1-2), 87 - 110.

Brusilovsky, P. (2007). Adaptive Navigation Support. In Brusilovsky, P., Kobsa, A., & Nejdl, W. (Eds.), *Lecture Notes In Computer Science: The adaptive web* (Vol. 4321, pp. 263–290). Berlin: Springer-Verlag.

Brusilovsky, P., & Maybury, M. T. (2002). From adaptive hypermedia to the adaptive web. *Communications of the ACM*, 45 (5), 30-33.

Brusilovsky, P., Kobsa, A., & Nejdl, W. (Eds.) (2007). *The Adaptive Web*. Berlin: Springer-Verlag.

Brzozowski, M. J. (2009). WaterCooler: exploring an organization through enterprise social media.

Proceedings of the ACM 2009 international conference on Supporting group work (GROUP '09). ACM, New York, NY, USA, 219–228.

Bunt, A., Carenini, G., & Conati, C. (2007). Adaptive content presentation for the web. In Brusilovsky, P., Kobsa, A., & Nejdl, W. (Eds.), *Lecture Notes In Computer Science: The adaptive web* (Vol. 4321, pp. 409-432). Berlin: Springer-Verlag.

Bush, V. (1945). As We May Think. *The Atlantic*. Retrieved November 4, 2012 from http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/

Bystrom, K., & Jarvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing and Management*, 31 (2), 191-213.

C

Card, S.K., & Nation, D. (2002). Degree-of-interest trees: A component of an attention-reactive user interface. *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '02)*. ACM, New York, NY, USA, 231–245

Carroll, J. M., & Campbell, R. L. (1986). Softening Up Hard Science: Reply to Newell and Card. *Human-Computer Interaction*, 2 (3), 227–249.

Carroll, J. M. (2002). Making use is more than a matter of task analysis. *Interacting with Computers*, 14 (5), 619–627.

Carroll, J. M., & Kellogg, W. A. (1989). Artifact as theory-nexus: hermeneutics meets theory-based design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '89)*. ACM, New York, NY, USA, 7-14.

Chi, E. H., Hong, L., Heiser, J., Card, S. K., et al. (2007). ScentIndex and ScentHighlights: productive reading techniques for conceptually reorganizing subject indexes and highlighting passages. *Information Visualization*, 6 (1), 32-47.

Chi, E. H., Pirolli, P., Chen, K., & Pitkow, J. (2001). Using information scent to model user information needs and actions and the Web. *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '01)*. ACM, New York, NY, USA, 490-497.

Chierichetti, F., Kumar, R., & Raghavan, P. (2011). Optimizing two-dimensional search results presentation. *Proceedings of the fourth ACM international conference on Web search and data mining (WSDM '11).* Hong Kong, China, 257-266.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '88).* ACM, New York, NY, USA, 213-218.

Chun, M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*. 4 (5), 170–178.

Claburn, T. (2009). Military Grapples With Information Overload. *Information Week*. Retrieved November 20, 2011 from http://www.informationweek.com/news/government/enterprise-architecture/218401332

Convertino, G., Hong, L., Nelson, L., Pirolli, P., & Chi, E. H. (2009). Activity Awareness and Social Sensemaking 2.0: Design of a Task Force Workspace. In Schmorrow, D. D., Estabrooke, I. V., & Grootjen, M. (Eds.), *Lecture Notes In Computer Science: Foundations of Augmented Cognition, Neuroergonomics and Operational Neuroscience* (Vol. 5638, pp. 128-137). Berlin: Springer-Verlag.

Convertino, G., Kairam, S., Hong, L., Suh, B., Chi, E. H. (2010). Designing a cross-channel information management tool for workers in enterprise task forces. *Proceedings of the International Conference on Advanced Visual Interfaces (AVI '10)*. ACM, New York, NY, USA, 103–110.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24 (1), 87–114.

Cowan, N. (2010). The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, 19 (1), 51–57.

Crosby, M.E., Iding, M.K., & Chin, D.N. (2003). Research on Task Complexity as a Foundation for Augmented Cognition. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*. IEEE, Big Island, HI, USA.

Cutrell, E., & Guan, Z. (2007). An eye-tracking study of information usage in Web search: Variations in target position and contextual snippet length. *Microsoft Research TechReport MSR-TR-2007-01*. Retrieved August 04, 2011 from http://research.microsoft.com/apps/pubs/default.aspx?id=70395

Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). A Diary Study of Task Switching and Interruptions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04).* ACM, New York, NY, USA, 175–182.

D

Danielson, D. (2002). Web navigation and the behavioral effects of constantly visible site maps. *Interacting with Computers*, 14 (5), 601–618.

Davenport, T., & Prusak, L. (2000). *Working knowledge: how organizations manage what they know* (1st ed.). Boston, MS, USA: Harvard Business School Press.

de Vignemont, F. (2011). Bodily Awareness. In Zalta, E. N. (Ed.). *Stanford Encyclopedia of Philosophy*. Retrieved from: http://plato.stanford.edu/entries/constructivism-metaethics/

Debner, J. A., & Jacoby, L. L. (1994). Unconscious perception: Attention, awareness, and control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20 (2), 304-317.

Demaine, J. (1996). Information Visualization. *Network Notes #35, Information Technology Services, National Library of Canada.* Retrieved November 8, 2012, from http://dnthong.blogspot.co.uk/2005/03/information-visualization.html

DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., et al. (2008). Motivations for social networking at work. *Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW '08).* ACM, New York, NY, USA, 711–720

Dow, S., Fortuna, J., Schwartz, D., Altringer, B., et al. (2011). Prototyping dynamics: sharing multiple designs improves exploration, group rapport, and results. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, New York, NY, USA, 2807-2816.

Drèze, X., & Hussherr, F.-X. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing*, 17 (4), 8–23.

Drury, C. G., & Clement, M. R. (1978). The effect of area, density, and number of background characters on visual search. *Human Factors*. 20 (5), 597–602.

Dubie, D. (2006). Time spent searching cuts into company productivity. *Networkworld*. Retrieved August 29, 2011, from: http://www.networkworld.com/news/2006/102006-search-cuts-productivity.html

Duchowski, A. (2007). *Eye tracking methodology* (2nd ed.). London, UK: Springer-Verlag London Limited.

Dumais, S., & Cutrell, E. (2001). Optimizing search by showing results in context. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01).* ACM, New York, NY, USA, 277-284.

E

Eddy, S. R. (2005). 'Antedisciplinary' Science. *PLoS Computational Biology*, 1 (1), e6. doi:10.1371/journal.pcbi.0010006.

Ehrlich, K., & Shami, N. S. (2008). Searching for Expertise. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08).* ACM, New York, NY, USA, 1093–1096.

Ehrlich, K., Lin, C.-Y., & Griffiths-Fisher, V. (2007). Searching for experts in the enterprise: combining text and social network analysis. *Proceedings of the 2007 international ACM conference on Supporting group work (GROUP '07).* ACM, New York, NY, USA, 117-126.

Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). Manual for the Kit of Factor-Referenced Cognitive Tests. *Educational Testing Service*. Retrieved August 14, 2008, from http://www.ets.org/Media/Research/pdf/Manual_for_Kit_of_Factor-Referenced_Cognitive_Tests.pdf

Eppler, M., & Mengis, J. (2004). The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*, 20 (5), 325–344.

Ercikan, K. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher*, 35 (5), 14-23.

Erdelez, S. (1995). *Information encountering: An exploration beyond information seeking*. Unpublished information and library science doctorol thesis, Department of Information Transfer, Syracuse University.

F

Fear, J. (2008). Choice Overload: Australians coping with financial decisions. *The Australia Institute*. Retrieved June 17, 2011, from http://www.tai.org.au/documents/dp_fulltext/DP99.pdf

Floridi, L. (2004). Outline of a Theory of Strongly Semantic Information. *Minds and Machines*, 14 (2), 197–221.

Floridi, L. (2011). Semantic Conceptions of Information. In Zalta, E. N. (Ed.). *Stanford Encyclopedia of Philosophy*. Retrieved from: http://plato.stanford.edu/entries/information-semantic/

Furnas, G. (1981). The FISHEYE view: A new look at structured files. *Bell Laboratories*, Technical Memorandum 82-11221-22. Retrieved March 25, 2011, from http://furnas.people.si.umich.edu/Papers/FisheyeOriginalTM.pdf

Furnas, G. (1986). Generalized fisheye views. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '86).* ACM, New York, NY, USA, 16-23.

Furnas, G. (2006). A fisheye follow-up: further reflections on focus+ context. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06).* ACM, New York, NY, USA, 999-1008.

Furniss, D., Blandford, A., & Curzon, P. (2011). Confessions from a grounded theory PhD: experiences and lessons learnt. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, New York, NY, USA, 113-122.

Fuster, J. M. (1998). Linkage at the top. *Neuron*. 21 (6), 1223–1229.

G

Gabbard, J. L., Hix, D., & Swan, J. E. (1999). User-centered design and evaluation of virtual environments. *Computer Graphics and Applications*, 19 (6), 51–59.

Gallagher, S., & Zahavi, D. (2010). Phenomenological Approaches to Self-Consciousness. In Zalta, E. N. (Ed.). *Stanford Encyclopedia of Philosophy.* Retrieved from: http://plato.stanford.edu/entries/self-consciousness-phenomenological/

Gamble, T., & May, J. (2003). Task switching and switch cost between task sequences. *Proceedings of the Cognitive Section Annual Conference of the British Psychological Society*. Reading, UK.

Gamst, G. C., Liang, C. T. H., & Der-Karabetian, A. (2011). *Handbook of Multicultural Measures* (1st ed.). Thousand Oaks, CA, USA: Sage Publications, Inc.

Garcia-Molina, H., Koutrika, G., & Parameswaran, A. (2011). Information seeking. *Communications of the ACM*, 54 (11), 121–130.

Garstka, J. (2003). Network Centric Operations Conceptual Framework Version 1.0. *Defense Technical Information Center (DTIC).* Retrieved on August 16, 2011, from http://www.dtic.mil/dtic/tr/fulltext/u2/a457620.pdf

Geiger, S., & Reeves, B. (1993). The Effects of Scene Changes and Semantic Relatedness on Attention to Television. *Communication Research*, 20 (2), 155–175.

Gena, C., & Weibelzahl, S. (2007). Usability Engineering for the Adaptive Web. In Brusilovsky, P., Kobsa, A., & Nejdl, W. (Eds.), *Lecture Notes In Computer Science: The adaptive web* (Vol. 4321, pp. 720–762). Berlin: Springer-Verlag.

Gerrig, R.J., & Zimbardo, P.G. (2007). *Psychology and Life.* (18th ed.). Harlow: Pearson.

Gibson, J. J. (1966). *The senses considered as perceptual systems* (1st ed.). Oxford, England: Houghton Mifflin.

Gibson, E. J. (1969). *Principles of perceptual learning and development* (1st ed.). East Norwalk, CT, United States: Appleton-Century-Crofs.

Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research* (1st ed.). New Brunswick, USA: Transaction Publishers.

Glasersfeld, von, E. (1995). *Radical Constructivism: A Way of Knowing and Learning* (1st ed.). Bristol, PA, USA: Falmer Press, Taylor & Francis Inc.

Goulding, C. (1998). Grounded theory: the missing methodology on the interpretivist agenda. *Qualitative Market Research: An International Journal*, 1 (1), 50–57.

Greenberg, S., & Cockburn, A. (1999). Getting Back to Back: Alternate Behaviors for a Web Browser's Back Buttons. *Proceedings of the 5th Annual Human Factors and the Web Conference*, Gaithersburg, Maryland.

Gupta, A., Sharda, R., Ducheneaut, N., & Zhao, J. (2006). E-Mail Management: A Techno-Managerial Research Perspective. *Communications of the Association for Information Systems*, 17, 941-961.

Gutwin, C. (2002). Improving Focus Targeting in Interactive Fisheye Views. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '02).* ACM, New York, NY, USA, 267–274.

H

Hailpern, J., Jitkoff, N., Warr, A., Karahalios, K., et al. (2011). YouPivot: Improving Recall with Contextual Search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, New York, NY, USA, 1521–1530.

Hansen, J. P., Johansen, A. S., Hansen, D. W., Itoh, K., et al. (2003). Command without a click: Dwell time typing by mouse and gaze selections. *Proceedings of the 9th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2009)*. Cape Town, South Africa, 121–128.

Harper, B. D., Slaughter, L. A., & Norman, K. L. (1997). Questionnaire administration via the WWW: A validation & reliability study for a user satisfaction questionnaire. *Proceedings of WebNet 97 – World Conference on the WWW, Internet & Intranet.* Toronto, Canada.

He, D., Brusilovsky, P., Ahn, J., Grady, J., et al. (2008). An evaluation of adaptive filtering in the context of realistic task-based information exploration. *Information Processing & Management*, 44 (2), 511–533.

Hearst, M. (2009). *Search user interfaces* (1st ed.). New York, USA: Cambridge University Press.

Hoeber, O., & Yang, X.D. (2008). Evaluating WordBars in exploratory Web search scenarios. *Information Processing & Management*, 44 (2), 485–510.

Hollnagel, E. (1992). Coping, coupling and control: the modelling of muddling through. *Mental models and everyday activities, 2nd interdisciplinasy workshop on mental models,* Robinson College, Cambridge, UK.

Holtzblatt, K., Wendell, J. B., & Wood, S. (2005). *Rapid contextual design: a how-to guide to key techniques for user-centered design* (1st ed.). San Francisco, CA, USA: Elsevier, The Morgan Kaufmann Series in Interactive Technologies.

Hornof, A. J., & Halverson, T. (2003). Cognitive Strategies and Eye Movements for Searching Hierarchical Computer Displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03).* ACM, New York, NY, USA, 1–8.

Huang, J., & White, R. (2011). No clicks, no problem: using cursor movements to understand and improve search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, New York, NY, USA, 1225-1234.

Hyönä, J., Radach, R., & Deubel, H. (Eds.) (2003). *The mind's eye: cognitive and applied aspects of eye movement research* (1st ed.). Amsterdam, Netherlands: Elsevier, North-Holland.

I

Indratmo, Vassileva, J., & Gutwin, C. (2008). Exploring blog archives with interactive visualization. *Proceedings of the working conference on Advanced visual interfaces (AVI '08).* ACM, New York, NY, USA, 39-46.

Ingwersen, P., & Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context* (1st ed.). Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Inkpen, K., Whittaker, S., Czerwinski, M., Fernandez, R., & Wallace, J. (2009). GroupBanter: Supporting Serendipitous Group Conversations with IM. In Bertino, E., & Joshi, J. B. D. (Eds.), *Lecture Notes of the Institute of Computer Sciences, Social Informatics and Telecommunications Engineering: Collaborative Computing: Networking, Applications and Worksharing* (Vol. 10, pp. 485–498) Berlin: Springer-Verlag.

Iyengar, S. S., & Kamenica, E. (2007). *Choice overload and simplicity seeking.* Unpublished working paper, Graduate School of Business, University of Chicago.

J

Jones, W., Pirolli, P., Card, S., Fidel, R., et al. (2006). It's about the information stupid!: why we need a separate field of human-information interaction. *Extended Abstracts on Human Factors in Computing Systems (CHI EA '06).* ACM, New York, NY, USA, 65-68.

Junker, A. M., & Hansen, J. P. (2006). Gaze pointing and facial EMG clicking. *Proceedings of the 2nd COGAIN Annual Conference on Communication by Gaze Interaction.* Turin, Italy.

K

Klein, G., Phillips, J. K., Raill, E. L., & Peluso, D. A. (2007). A Data-Frame Theory of Sensemaking. In Hoffman, R. R. (Ed.), *Expertise Out of Context. Proceedings of Sixth International Conference on Naturalistic Decision Making* (pp. 113–155). Pensacola Beach, Florida, USA: Lawrence Erlbaum Assoc. Inc.

Kobsa, A., Koenemann, J., & Pohl, W. (2001). Personalised hypermedia presentation techniques for improving online customer relationships. *The Knowledge Engineering Review,* 16 (2), 111-155.

Kolb, D.A. (1984). *Experiential learning: Expierence as the Source of Learning and Development* (1st ed.). Upper Saddle River, New Jersey, USA: Prentice Hall, Inc.

Kosara, R., Miksch, S., & Hauser, H. (2002). Focus+context taken literally. *IEEE Computer Graphics,* 22 (1), 22-29.

Kraft, R., Chang, C. C., Maghoul, F., & Kumar, R. (2006). Searching with context. *Proceedings of the 15th international conference on World Wide Web (WWW '06).* ACM, New York, NY, USA, 477-486.

Kraft, R., Maghoul, F., & Chang, C. C. (2005). Y!Q: contextual search at the point of inspiration. *Proceedings of the 14th ACM international conference on Information and knowledge management (CIKM '05),* ACM New York, NY, USA, 816-823.

Kuhlen, R. (1991). *Hypertext: ein nicht-lineares Medium zwischen Buch und Wissensbank* (1st ed). Berlin, Germany: Springer.

Kumar, M., Paepcke, A., & Winograd, T. (2007). Eyepoint: practical pointing and selection using gaze and keyboard. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07).* ACM, New York, NY, USA, 421–430.

L

Lamping, J., Rao, R., & Pirolli, P. (1995). A focus+ context technique based on hyperbolic geometry for visualizing large hierarchies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '95).* ACM, New York, NY, USA, 401-408.

Laqua, S., & Brna, P. (2005). The focus-metaphor approach: A novel concept for the design of adaptive and user-centric interfaces. *Proceedings of the 7th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2005)*. Rome, Italy, 295-308.

Laqua, S., Sasse, M., Gates, C., & Greenspan, S. (2009). Making Sense of the Unknown: Knowledge Dissemination in Organizations. *2nd Sensemaking Workshop at CHI 2009.* New York, NY, USA.

Laqua, S., Sasse, M.A., Greenspan, S., & Gates, C. (2011). Do you KnowDis? A user study of a knowledge discovery tool for organizations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11).* ACM, New York, NY, USA, 2887-2896.

Lawrence, S. (2000). Context in web search. *IEEE Data Engineering Bulletin,* 23, 25-32.

Leshed, G., Haber, E.M., Matthews, T., & Lau, T. (2008). CoScripter: automating & sharing how-to knowledge in the enterprise. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08).* ACM, New York, NY, USA, 1719-1728.

Levy, M., Hadar, I., & Greenspan, S., Hadar, E. (2010). Uncovering cultural perceptions and barriers during knowledge audit. *Journal of Knowledge Management,* 14 (1), 114-127.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7 (1), 57–78.

Lewis, J. R. (2002). Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human-Computer Interaction*, 14 (3-4), 463–488.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22 (140), 1–55.

Lin, C. (2006). Optimal Web site reorganization considering information overload and search depth. *European Journal of Operational Research,* 173 (3), 839–848.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*. 7 (1), 76–80.

Lynch, C. (2001). Personalization and recommender systems in the larger context: New directions and research questions. *Proccedings of the second DELOS Network of Excellence Workshop: Personalisation and Recommender Systems in Digital Libraries*. Dublin, Ireland, 84-88.

M

Ma, Z., Pant, G., & Sheng, O. R. L. (2007). Interest-based personalized search. *ACM Transactions on Information Systems*, 25 (1), article no.5.

Mack, A., & Rock, I. (1998). *Inattentional blindness* (1st ed.). Cambridge, MA, USA: MIT Press.

Mackinlay, J. D., Robertson, G. G., & Card, S. K. (1991). The perspective wall: Detail and context smoothly integrated. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91).* ACM, New York, NY, USA, 173–176.

Majaranta, P., Aoki, H., Donegan, M., Hansen, D.W., et al. (2011). *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies* (1st ed.). Hershey, PA, USA: IGI Global.

Majaranta, P., Aula, A., & Räihä, K.-J. (2004). Effects of feedback on eye typing with a short dwell time. *Proceedings of the 2004 symposium on Eye tracking research & applications (ETRA '04).* ACM, New York, NY, USA, 139–146.

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49 (4), 41–46.

Marchionini, G., & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer,* 21 (1), 70–80.

Marsh, R. (2008). *Similar Posts Plugin for Wordpress.* Retrieved from http://rmarsh.com/plugins/similar-posts/

Mayring, P. (2000). Qualitative Content Analysis. *Forum: Qualitative Social Research*, 1 (2), article no.20.

McCloskey, M. (2007). Intranet Information Architecture (IA) Trends. *Nielsen Norman Group*. Retrieved January 15, 2012. from: http://www.nngroup.com/articles/intranet-information-architecture-ia/

McFarlane, D., & Latorella, K. (2002). The Scope and Importance of Human Interruption in Human-Computer Interaction Design. *Human-Computer Interaction*, 17 (1), 1–61.

Merikle, P. M., & Joordens, S. (1997). Parallels between Perception without Attention and Perception without Awareness. *Consciousness and cognition*, 6 (2-3), 219-236.

Micarelli, A., Gasparetti, F., Sciarrone, F., & Gauch, S. (2007). Personalized Search on the World Wide Web. In Brusilovsky, P., Kobsa, A., & Nejdl, W. (Eds.), *Lecture Notes In Computer Science: The adaptive web* (Vol. 4321, pp. 195–230). Berlin: Springer-Verlag.

Millen, D. R., & Fontaine, M. A. (2003). Improving Individual and Organizational Performance through Communities of Practice. *Proceedings of the ACM 2003 international conference on Supporting group work (GROUP '03).* ACM, New York, NY, USA, 205–211.

Millen, D. R., Feinberg, J., & Kerr, B. (2006). Dogear: Social bookmarking in the enterprise. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06).* ACM, New York, NY, USA, 111-120.

Miniotas, D., Špakov, O., & MacKenzie, I. S. (2004). Eye gaze interaction with expanding targets. *Extended Abstracts on Human Factors in Computing Systems (CHI EA '04).* ACM, New York, NY, USA, 1255–1258.

Miniotas, D., Špakov, O., Tugoy, I., & MacKenzie, I. S. (2005). Extending the limits for gaze pointing through the use of speech. *Information Technology and Control*, 34 (3), 225–230.

Miyake, A., & Shah, P. (Eds.) (1999). *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge, UK: Cambridge University Press.

Monsell, S. (2003). Task switching. *Trends in Cognitive Science*s, 7 (3), 134–140.

Mooers, C. N. (1959). Some mathematical fundamentals of the use of symbols in information retrieval. *Proceedings of IFIP Congress*, 315–320.

Munzner, T., Guimbretière, F., Tasiran, S., Zhang, L., & Zhou, Y. (2003). TreeJuxtaposer: scalable tree comparison using Focus+ Context with guaranteed visibility. *Proceedings of ACM SIGGRAPH 2003*, ACM, New York, NY, USA, 453-462.

N

Nation, D., Roberts, D., & Card, S.K. (2000). *Browse hierarchical data with the degree of interest tree.* Unpublished paper, National Security Agency & Xerox Palo Alto Research Center.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. In Prinz, W., & Sanders, A. F. (Eds.). *Cognition and motor processes* (1st ed., pp 255-293). Berlin: Springer-Verlag.

Nielsen, J. (1997). Why Advertising Doesn't Work on the Web. *Nielsen Norman Group*. Retrieved January 15, 2012, from: http://www.nngroup.com/articles/why-advertising-doesnt-work-on-the-web/

Nielsen, J. (2002). Intranet Usability: The Trillion-Dollar Question. *Nielsen Norman Group.* Retrieved November 11, 2012, from: http://www.nngroup.com/articles/intranet-usability-the-trillion-dollar-question/

Nielsen, J. (2005). Scrolling and Scrollbars. *Nielsen Norman Group.* Retrieved November 11, 2012, from http://www.nngroup.com/articles/scrolling-and-scrollbars/

Nielsen, J. (2007). Banner Blindness: Old and New Findings. *Nielsen Norman Group.* Retrieved January 15, 2012, from: http://www.nngroup.com/articles/banner-blindness-old-and-new-findings/

Noik, E.G. (1994). A Space of Presentation Emphasis Techniques for Visualizing Graphs. *Proceedings of Graphics Interface '94*. Banff, Alberta, Canada, 225-234.

Norman, D. A. (1992). *Turn signals are the facial expressions of automobiles* (1st ed.). Cambridge, MA, USA: Addison-Wesley.

Norman, D. A. (1999). Commentary-Banner Blindness. *Internetworking.* Retrieved January 15, 2012, from http://www.internettg.org/mar99/commentary.html

O

Olsen, W. (2004). Triangulation in Social Research: Qualitative and Quantitative Methods Can Really Be Mixed. In: M. Holborn (Ed.), *Developments in Sociology*. Ormskirk: Causeway Press.

Ortoleva, P. (2008). The price of flexibility: towards a theory of thinking aversion. *Journal of Economic Theory*, 148 (2013), 903-934.

Oulasvirta, A., & Saariluoma, P. (2004). Long-term working memory and interrupting messages in human-computer interaction. *Behaviour & Information Technology*, 23 (1), 53–64.

Outsell (2005). HotTopics: 2001 vs. 2005: Research Study Reveals Dramatic Changes Among Information Consumers. *Outsell Inc.* Retrieved November 29, 2012, from: http://www.outsellinc.com/store/products/219

P

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38 (1), 1–4.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The PageRank Citation Ranking: Bringing Order to the Web. *Stanford InfoLab, Technical Report.* Retrieved March 15, 2011, from http://ilpubs.stanford.edu:8090/422/

Palmer, C. R., Pesenti, J., Valdes-Perez, R. E., Christel, M. G., et al. (2001). Demonstration of hierarchical document clustering of digital library retrieval results. *Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries*. ACM, New York, NY, USA, 451.

Pirolli, P. (2007). *Information Foraging Theory: Adaptive Interaction with Information (1st ed.).* Oxford Series in Human-Technology Interaction, New York: Oxford University Press.

Pirolli, P., Card, S. K., & Van Der Wege, M. M. (2000). The effect of information scent on searching information: visualizations of large tree structures. *Proceedings of the working conference on Advanced visual interfaces (AVI '00).* ACM, New York, NY, USA, 161-172.

Pirolli, P., Schank, P., Hearst, M., & Diehl, C. (1996). Scatter/gather browsing communicates the topic structure of a very large text collection. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96).* ACM, New York, NY, USA, 213–220.

Pitkow, J., Schütze, H., Cass, T., Cooley, R., et al. (2002). Personalized search. *Communications of the ACM*, 45 (9), 50–55.

Pretschner, A., & Gauch, S. (1999). Ontology based personalized search. *Proc. 11th IEEE Intl. Conf. on Tools with Artificial Intelligence*. Chicago, IL, USA, 391–398.

Punch, K. F. (2005). *Introduction to social research: quantitative and qualitative approaches* (7th ed.). Trowbridge, Wiltshire: SAGE Publications Ltd.

Pylyshyn, Z. W. (2003). *Seeing and visualizing: it's not what you think* (1st ed.). Cambridge, MA: MIT Press.

Q

Qu, Y., & Furnas, G. W. (2008). Model-driven formative evaluation of exploratory search: A study under a sensemaking framework. *Information Processing & Management*, 44 (2), 534–555.

R

Reeves, B., & Nass, C. (2000). Perceptual Bandwidth. *Communications of the ACM*, 43 (3), 65–70.

Resnikoff, H. L. (1989). *The Illusion of Reality* (1st ed.). New York, NY: Springer-Verlag.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124 (2), 207–231.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27 (4), 763–797.

Russell, D. M., Jeffries, R., & Irani, L. (2008). Sensemaking for the rest of us. *Sensemaking Workshop at CHI 2008*. New York, NY, USA.

Russell, D. M., Stefik, M. J., & Pirolli, P., Card, S. K. (1993). The cost structure of sensemaking. *Proceedings of the INTERACT '93 nd CHI '93 Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 269-276.

Russell, D. M., Furnas, G., Stefik, M. J., Card, S. K., Pirolli, P. (2008). Sensemaking Workshop CHI 2008 Extend Abstract. *Extended Abstracts on Human Factors in Computing Systems (CHI EA '08).* ACM, New York, NY, USA, 3981–3984.

Russell, D. M., Pirolli, P., Furnas, G., Card, S. K., et al. (2009). Sensemaking Workshop CHI 2009 Extend Abstract. *Extended Abstracts on Human Factors in Computing Systems (CHI EA '00).* ACM, New York, NY, USA, 4751-4754.

Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance*, 27 (6), 1404-1419.

Ruthven, I., Baillie, M., Azzopardi, L., Bierig, R., et al. (2008). Contextual factors affecting the utility of surrogates within exploratory search. *Information Processing & Management*, 44 (2), 437–462.

S

Schilit, B., Adams, N., & Want, R. (1994). Context-Aware Computing Applications. *First Workshop on Mobile Computing Systems and Applications (WMCSA),* IEEE, 85–90.

Schrammel, J., Giller, V., Tscheligi, M., Kosara, R., et al. (2003). Experimental evaluation of semantic depth of field, a preattentive method for focus+ context visualization. *Proceedings of the 6th IFIP TC 13 International Conference on Human-Computer Interaction (INTERACT 2003)*. Zurich, Switzerland, 888–891.

Schulmeister, R. (2007). *Grundlagen hypermedialer Lernsysteme: Theory – Didaktik – Design* (4th ed.). Munich, Germany: Oldenbourg Publishers.

Schwartz, B. (2005). *The Paradox of Choice: Why More is Less*. 2005 (1st ed.). New York: Harper Collins.

Shanker, T., & Richtel, M. (2011). Military Struggles to Harness a Flood of Data. *The New York Times,* 16 January 2011. Retrieved on November 20, 2011 from http://www.nytimes.com/2011/01/17/technology/17brain.html

Shannon, C. (1948). The mathematical theory of communication. *Reprinted with corrections from The Bell System Technical Journal.* 27, 379–423, 623–656.

Sibert, L. E., & Jacob, R. J. K. (2000). Evaluation of eye gaze interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00).* ACM, New York, NY, USA, 281–288.

Sieck, W. R., Klein, G., Peluso, D. A., Smith, J. L., & Harris-Thompson, D. (2007). *FOCUS: A model of sensemaking.* Technical Report, Research Institute for the Behavioral and Social Sciences, United States Army.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*. 63 (2), 129–138.

Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), *Computers, communications, and the public interest* (1st ed., pp. 37–72). Baltimore: The Johns Hopkins Press.

Simon, H. A. (1972). Theories of bounded rationality. In McGuire, C. B., & Radner, R. (Eds.). *Decision and organization* (1st ed., pp. 161-176). Amsterdam: North-Holland Publishing Company.

Simon, H. A. (1979). Rational decision making in business organizations. *The American Economic Review,* 69 (4), 493-513.

Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind and Society: Cognitive Studies in Economics and Social Sciences,* 1 (1), 25–39.

Simons, D., & Levin, D. (1997). Change blindness. *Trends in Cognitive Sciences*, 1 (7), 261–167.

Skovsgaard, H., Räihä, K.-J., & Tall, M. (2011). Computer Control by Gaze. In Majaranta, P., Aoki, H., Donegan, M., Hansen, D. W., et al. (Eds.), *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies* (1st ed., pp. 78-102). Hershey, PA: IGI Global.

Smith, E. E., & Kosslyn, S. M. (2008). *Cognitive Psychology: Mind and Brain* (1st ed.). Harlow: Pearson Prentice Hall.

Sloan, R., & Thompson, M. (2004). *Epic 2014*. Retrieved June 14, 2006 from http://idorosen.com/mirrors/robinsloan.com/epic/

Soucek, R., & Moser, K. (2010). Coping with information overload in email communication: Evaluation of a training intervention. *Computers in Human Behavior*, 26 (6), 1458–1466.

Speretta, M., & Gauch, S. (2005). Personalized search based on user search histories. *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI '05).* Washington, DC, USA, 622–628.

Spira, J. B. (2008). Information Overload: Now $900 Billion – What is Your Organization's Exposure? *Basex – Management Science for the Knowledge Economy*. Retrieved January 14, 2012, from http://www.basexblog.com/2008/12/19/information-overload-now-900-billion-what-is-your-organizations-exposure/

Spira, J. B. (2011). Some New Year's Resolutions for the Information Overload-Minded. *Basex – Management Science for the Knowledge Economy*. Retrieved January 14, 2012, from http://www.basexblog.com/2011/01/06/new-years-resolutions/

Spira, J. B., & Goldes, D. (2007). *Information Overload: We have met the enemy and he is us*. Basex Report, pp.1–30.

Sproull, L., & Kiesler, S. (1991). *Connections: new ways of working in the networked organization* (1st ed.). Cambridge, MA: MIT Press.

Stamen (2006). *Digg Labs: Swarm*. Retrieved November 18, 2012 from http://www.visualcomplexity.com/vc/project.cfm?id=380

Stasko, J., Goerg, C., & Liu, Z. (2008). Sensemaking across Text Documents: Human-Centered, Visual Exploration with Jigsaw. *Sensemaking Workshop at CHI 2008*. New York, NY, USA.

Stefik, M. (2004). *The new sensemakers: what's the next thing beyond search?* sensemaking. Innovation Pipeline. 2004 December; 2 (10): 13.

Stolterman, E. (2008). The nature of design practice and implications for interaction design research. *International Journal of Design*, 2 (1), 55–65.

Stonier, T. (1991). Towards a new theory of information. *Journal of Information Science*, 17 (5), 257-263.

Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research* (1st ed). Thousand Oaks, CA: Sage Publications.

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12 (2), 257–285.

Špakov, O. (2011). Comparison of gaze-to-objects mapping algorithms. *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications (NGCA '11).* ACM, New York, NY, USA, article no. 6.

T

Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: a study of orienteering behavior in directed search. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04).* ACM, New York, NY, USA, 415-522.

Teevan, J., Cutrell, E., Fisher, D., Drucker, S. M., et al. (2009). Visual snippets: summarizing web pages for search and revisitation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09).* ACM, New York, NY, USA, 2023-2032.

Tobii (2011). *Eye tracking by Tobii Technology.* Retrieved November 27, 2012, from http://www.tobii.com/group/about-tobii/eye-tracking-by-tobii/the-basics-of-eye-tracking/

V

Villa, R., Gildea, N., & Jose, J. M. (2008). FacetBrowser: a user interface for complex search tasks. *Proceedings of the 16th ACM international conference on Multimedia (MM '08).* ACM, New York, NY, USA, 489-498.

W

Ware, C. (2004). *Information visualization: perception for design* (2nd ed.) San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review*, 104 (1), 90–122.

Weick, K. E. (1995). *Sensemaking in organizations* (1st ed.). Thousand Oaks, CA: Sage Publications, Inc.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. *Organization Science*, 16 (4), 409–421.

Weldon, M. (2000). Remembering as a social process. In Medin, D. L. (Ed.), *The psychology of learning and motivation. Vol. 40.* (pp. 67–120). San Diego: Academic Press.

Weskamp, M., & Albritton, D. (2004). *newsmap*. Retrieved November 2, 2012, from http://marumushi.com/projects/newsmap

White, R. W., & Drucker, S. M. (2007). Investigating behavioral variability in web search. *Proceedings of the 15th international conference on World Wide Web (WWW '07).* ACM, New York, NY, USA, 21–30.

White, R. W., Marchionini, G., & Muresan, G. (2008). Evaluating exploratory search systems. *Information Processing & Management*, 44 (2), 433–436.

Whittaker, S., Bellotti, V., & Gwizdka, J. (2006). Email in personal information management. *Communications of the ACM*, 49 (1), 68–73.

Wickens, C.D., & Hollands, J.G. (2000). *Engineering psychology and human performance* (3rd ed.). Harlow: Pearson Prentice Hall.

Wixon, D., Holtzblatt, K., & Knox, S. (1990). Contextual design: an emergent view of system design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90).* ACM, New York, NY, USA, 329-336.

Wilson, G. (2005). Info-overload harms concentration more than marijuana. *New Scientist.* (2497).

Wood, S., Cox, R., & Cheng, P. (2006). Attention design: Eight issues to consider. *Computers in Human Behavior*, 22 (4), 588–602.

Z

Zaltman, G. (2003). *How customers think: essential insights into the mind of the market* (1st ed.). Boston, MA: Harvard Business School Press.

Zavala, A. (1965). Development of the forced-choice rating scale technique. *Psychological Bulletin*, 63 (2), 117–124.

Appendix A: Preliminary Study 2

10.1. Questions Used in the Study

e-Passport Questions

1. 1) When will they be issues in the UK?
2. 2) Do these passports have an electronic chip in them?
3. 3) How long is the interview?
4. 4) How many passport applicants were there last year?
5. 5) In what year will the interviews be compulsory?
6. 6) Can you tell me a benefit of introducing ePassports?
7. 7) Are the ePassports reader machines used at air ports 100% secure?
8. 8) What kind of personal information does the chip hold?
9. 9) How much does an ePassport cost?
10. 10) How long is the warranty of the microchip?
11. 11) Are there any other countries that use ePassports?

ID Cards Questions

1. 1) How much does an ID card for retired people cost?
2. 2) Do under 16s need an ID card?
3. 3) What biometric data will be stored on the card?
4. 4) How could ID cards help to fight terrorism?
5. 5) When will the ID cards be issued?
6. 6) Will foreigners staying in the UK have to get an ID card?
7. 7) Will the personal information be stored in just one database?
8. 8) Will it be compulsory to carry an ID card?
9. 9) What percentage of the population in France carries an ID card?
10. 10) Is reducing identity theft a benefit of ID cards?
11. 11) Do the liberal democrats support ID cards?
12.

10.2. Post Study Questionnaire

Page 1

EYE TRACKING SYSTEM USER QUESTIONNAIRE

Please answer the following questions:

1) Name: ....

2) Ethnic background
   - ☐ White: British
   - ☐ White: Other
   - ☐ Asian or Asian British
   - ☐ Black or Black British
   - ☐ Chinese
   - ☐ Other ....

3) Age: ....years

4) Gender: ☐ Male ☐ Female

5) Status:
   - ☐ Student
   - ☐ Part-time employment
   - ☐ Full-time employment
   - ☐ Other

6) Do you wear any eye wear?

   ☐ Glasses ☐ Contact lenses ☐ None

7) What is your vocation?
   - ☐ Student
   - ☐ Part-time employment
   - ☐ Full-time employment
   - ☐ Unemployed

8) How long have you been using a computer?
   - ☐ Less than 1 year
   - ☐ 1 year – 3 years
   - ☐ 3 years – 6 years
   - ☐ Over 6 years

9) How often do you use a computer?
   - ☐ Everyday
   - ☐ More than once a week
   - ☐ Once a week or less
   - ☐ Once a month or less
   - ☐ Once a year

10) For what purposes do you use a computer?
   - ☐ Internet
   - ☐ Emails
   - ☐ Word processing
   - ☐ Games
   - ☐ Graphics
   - ☐ Other: ....

11) How would you rate your expertise with computers?

   ☐ expert ☐ above average ☐ average ☐ below average ☐ beginner

12) Are you aware of eye tracking technology?

   ☐ Yes No ☐

   If yes, have you interacted with any eye tracking systems prior to this experiment?

   ☐ Yes No ☐

Overall impressions of the system

	1	2	3	4	5	6	
Terrible							Wonderful
Difficult							Easy
Frustrating							Satisfying
Inaccurate							Accurate
Dull							Interesting
Rigid							Flexible

Page 2

EYE TRACKING SYSTEM USER QUESTIONNAIRE

			1	2	3	4	5	6	
1	I found the system easy to use	Strongly Disagree							Strongly Agree
2	Learning to operate this system with my eyes was easy for me								
3	Performing the given task using this system was straightforward for me								
4	It was easy to find the information I needed								
5	I find it easy to navigate the system only with my eyes								
6	I believe I became productive quickly using this system								
			1	2	3	4	5	6	
7	Correcting my mistakes (e.g. activating the wrong element) was quite easy for me	Strongly Disagree							Strongly Agree
8	I find the system to be flexible to interact with								
9	The coloured border of elements (which darkened when looking at an element) helped me to navigate the system Any comment on this feature: …………………………………………………………………								
10	I feel that the system is quick to respond to my intentions								
11	Whenever I made a mistake, I recovered easily & quickly								
12	I found unexpected situations while interacting with the system								
			1	2	3	4	5	6	
13	The coloured background (red or green to indicate working status) helped me to interact with the system Any comment on this feature: …………………………………………………………………	Strongly Disagree							Strongly Agree
14	I had problems with 'unintentional activations' of elements while using the system								
15	Information provided was effective in helping me complete the tasks								
16	I am satisfied with the amount of time it took me to complete the task								
17	I feel comfortable using this system								
			1	2	3	4	5	6	
18	The information provided for this system was easy to understand	Strongly Disagree							Strongly Agree
19	The text was easy to read and the choice of type & size of fonts was appropriate								
20	The organization of information on the interface is clear								
21	The interface of this system is pleasant								
22	I found the pictures relevant to each topic and interesting								
23	It was easy to learn to use this system								
24	Overall, I am satisfied with how easy it is to use this system								
25	I would like to use this interface for other purposes								
			1	2	3	4	5	6	

Page 3

EYE TRACKING SYSTEM USER QUESTIONNAIRE

Apart from the content, did you feel any differences between the two experiments?

☐ Yes, I am ………% sure that there was a difference

☐ No, I am ………% sure that there was **no** difference

If yes,

- Which experiment did you like the most?

  ☐ 1st Experiment ☐ 2nd Experiment ☐ both the same

- Which interface was the easier to use?

  ☐ 1st Experiment ☐ 2nd Experiment ☐ both the same

  Please explain your answer briefly (what differences did you find):

  ………

  ………

  ………

The most negative aspect(s) of this system were:

1) ……………………………………………………………………………………………………

2) ……………………………………………………………………………………………………

3) ……………………………………………………………………………………………………

The most positive aspect(s) of this system were:

1) ……………………………………………………………………………………………………

2) ……………………………………………………………………………………………………

3) ……………………………………………………………………………………………………

If this eye tracking system can be used in other applications, where/how do you think it could be used?

……

……

……

Do you have any general comments about this experiment?

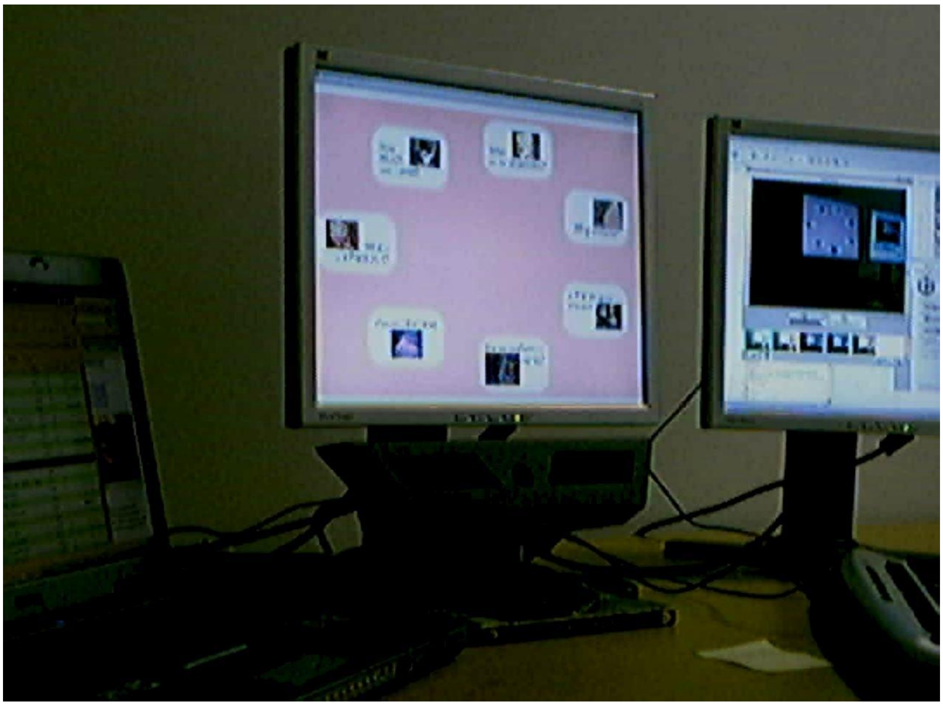
……

……

……

10.3. Images from the Study Setup


2007-03-19 14:28:36



Appendix B: Main Study 1

10.4. Ethical Guideline Form Used in the Study

001

IB / NF

[Reading on healthy lifestyles] eye-tracking experiment - Sven Laqua

Ethical Guidelines Form for research with human participants.

Name: **Sven Laqua**
Role: **Research Student**

Address: **Computer Science Department**
University College London

E-mail: **s.laqua@cs.ucl.ac.uk**

I am a research student in the Computer Science Department at University College London.
I am interested in **user testing for human computer interfaces (HCI).**

This research is consistent with the ethical guidelines set out by the British Psychological Society. These guidelines **include principles such as obtaining your informed consent before research starts, notifying you of your right to withdraw, and protection of your anonymity**. This sheet will hopefully provide you with enough information. However, if you have any questions or would like to discuss anything with me please let me know.

Experiment details:
You will be asked to read blog articles on a traditional and a novel user interface. The experiment will make use of non-intrusive eye-tracking (no head-mounted parts or other restrictions) to monitor your eye movements.
There will be a warm-up session to familiarize you with the interfaces. Your eyes will be calibrated and everything set up for the experiment.
The main experiment is split into two sessions. In the sessions you will be asked to complete small scenario-like tasks using web content on healthy lifestyles. After reading a scenario on the screen, you will access the lifestyles content and try to complete this scenario. Each session will consist of X small scenarios. After each session, you will be asked to complete a questionnaire on the conducted session. Between the two sessions, you can have a small break if desired. After completing both sessions, you will be debriefed and receive a short final questionnaire.
I will use the data obtained during the sessions, however you will remain anonymous as I will change any identifying name or place that is mentioned. This experiment will last between 45 minutes and 1 hour. You will be given more detailed instructions before the experiment.

When you have completed the experiment, you will receive a **compensation of £8** for your time.

The purpose of this form is to make sure that you are happy to take part in the research and you know what is involved.

Do you understand that you have the opportunity to ask questions and discuss the study? ✓ **YES – NO**

Do you understand that you are free to end the task required at anytime? ✓ **YES – NO**

Do you understand that you are free to choose not to do any task required? ✓ **YES – NO**

Do you agree to take part in this study? ✓ **YES – NO**

Do you grant permission for extracts from the records to be used in reports of the
research on the understanding that your anonymity will be maintained? ✓ **YES – NO**

NAME: [redacted]

DATE: 31 / 07 / **2007**

[Reading on healthy lifestyles] eye-tracking experiment - Sven Laqua

10.5. Information Search Tasks

Task 1

A colleague told you, that he just read an interesting article
on the pros and cons of drinking coffee every day.

As you worry about your coffee intake every day,
you want to **have a look at this article on coffee yourself...**

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

Task 2

After a few stressful weeks in the office you just successfully completed the latest project.

The weekend is coming up and instead of just being lazy in front of the TV for 2 days,
you decide to go hiking.

Find some information on what to take with you in your backpack.

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

Task 3

You have been working too much over the past 2 weeks
and did not get enough sleep.

Your friends start to make fun of the dark rings under your eyes.

You remember a colleague talking about helpful tips
on **how to treat tired eyes** in one of the articles,
you want to go find that information.

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

Task 4

You have heard from a friend about an article on a new nightclub in London.

It is special as it is the first "low energy" nightclub,
and this article talks in more detail about what sets this club apart
from its competition.

You are curious and want to find this article.

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

Task 5

Spring is here. There sun is shining,
people are happy, and you got hayfever...

You never really bothered to determine what causes
your symptoms, but finally decide to find out more.

A friend told you about a **great article on allergies** and you go **have a look for it.**

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

Task 6

You just moved into a nice property in the London suburbs.

The victorian house comes with a decent backyard but the grass looks terrible.

Having been at a colleagues barbecue party recently, you remember her wonderfully
green garden and people talking about this new method she used
after reading about it on this site.

You want to **find this article...**

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

Task 7

You like eating **salad** as part of your healthy diet,
but just 1 or 2 days after you got a new bag, the salad gets
slimy and soggy, and you end up throwing half of it away.

Talking with a colleague about your "salat troubles", she mentions that
she remembers seeing an article on **keeping salad fresh for longer**.

You a curious and want to find this article.

*(There is only one specific article, which this task refers to.
You need to find this one article.)*

10.6. Information Exploration Scenarios

Scenario 1

Living a busy lifestyle, your work occupies most of your time.

You think you basically have a healthy diet, but you noticed that you tire easily, and wonder if you are really eating right.

You decide to find a diet that improves your daily nutrition.

(You have 4 to 5 minutes to explore information that you believe provides useful insights concerning the given task. After completing the task, you will be asked to briefly reflect on the most useful information you found)

Scenario 2

You constantly hear news about global warming and that everybody should do their bit to help saving the environment.

The recent crazy weather finally convinces you that the climate is changing quickly, and you finally decide to do something yourself.

Living in a tiny flat in London and commuting with public transport every day, there is not much more you can do here.

But a friend told you that **ecological clothes** are becoming more and more popular.

You need some new cloths and think that might be a good way to do something for the environment.

Have a look at what has been written on this latest fashion trend.

(You have 4 to 5 minutes to explore information that you believe provides useful insights concerning the given task. After completing the task, you will be asked to briefly reflect on the most useful information you found)

Scenario 3

After the holidays, **you have gained a few pounds**. You are not happy and want to loose weight.

You have seen from your friends that diets don't work.

You believe that **excercise** is a much better way of getting back into shape.

There is a lot of of useful **information on sports and fitness** and you want to have a look and **find information that is suitable for you and fits into your lifestyle**.

(You have 4 to 5 minutes to explore information that you believe provides useful insights concerning the given task. After completing the task, you will be asked to briefly reflect on the most useful information you found)

Scenario 4

You go out a lot and drink too much alcohol on a regular basis.
You decide to cut your alcohol intake.

A friend told you that there is lots of information
about healthy drinks on this site.

You want to go and **find out more about the benefits of switching to water and other healthy drinks**.

(You have 4 to 5 minutes to explore information that you believe provides useful insights concerning the given task. After completing the task, you will be asked to briefly reflect on the most useful information you found)

Scenario 5

You are going to the gym regularly already,
but are not really satisfied with your progress.

You are slim enough, but want to gain some muscle.

Last time in the gym, you heard people talking about
how important the right diet is for muscle growth.

Remembering this site with lots of useful information
you want to find **out more about what other factors influence muscle building besides the actual workout.**

(You have 4 to 5 minutes to explore information that you believe provides useful insights concerning the given task. After completing the task, you will be asked to briefly reflect on the most useful information you found)

10.7. Questionnaire used

First Impression (6 point Likert scale used)

Terrible - Wonderful
Difficult - Easy
Frustrating - Satisfying
Inaccurate - Accurate
Dull - Interesting
Rigid - Flexible

General Usability Questions (6 point Likert scale used)

I found the system easy to use	Q01
It was easy to find the information I needed	Q02
Learning to navigate the system was easy	Q03
I believe I became productive quickly using the system	Q04
I find the system to be flexible to interact with	Q05
I feel the system is quick to respond to my intentions	Q06
When I made a mistake, I recovered easily and quickly	Q07
I am satisfied with the amount of time it took me to complete the task	Q08
I enjoyed reading the articles.	Q09
I feel comfortable using this system	Q10
I am satisfied with the amount of information provided	Q11
I am satisfied with the quality of information provided	Q12
The information provided for this system was easy to understand	Q13
The text was easy to read and the choice of type and size of fonts was appropriate	Q14
The organisation of information on the interface is clear	Q15
The interface of this system is pleasant	Q16
I would like to read more about these topics	Q17
Overall, I am satisfied with how easy it is to use the system	Q18

Appendix C: Main Study 2

10.8. Phase 1 - Interview Questions

Questions

Q1) How do you currently **use email**, and make use of CA **knowledge repositories**, either through search or other means?

a) Do you have problems organizing your work?

Q2) How would you categorize the **amount of email** you receive?

a) Do you have problems managing your email?

Q3) What **types of information** do you require for your everyday work?

a) Do you have problems getting to that information?

Q4) Do you attend (internal) **seminars, workshops** or other work related **events**?

a) Do you have problems finding out about those events?

Q5) How do you use tools, such as **SharePoint**, **Wikis** or other **collaborative work spaces**?

a) Are there any problems with integrating those tools into your daily work flow?

Q6) How often are tasks such as reading email and finding information on a knowledge repository entwined in your daily work?

a) Do you often interrupt email reading/responding to search for additional information that would help make sense of the email or is necessary for responding to the email?

b) What sort of files or info do you search for?

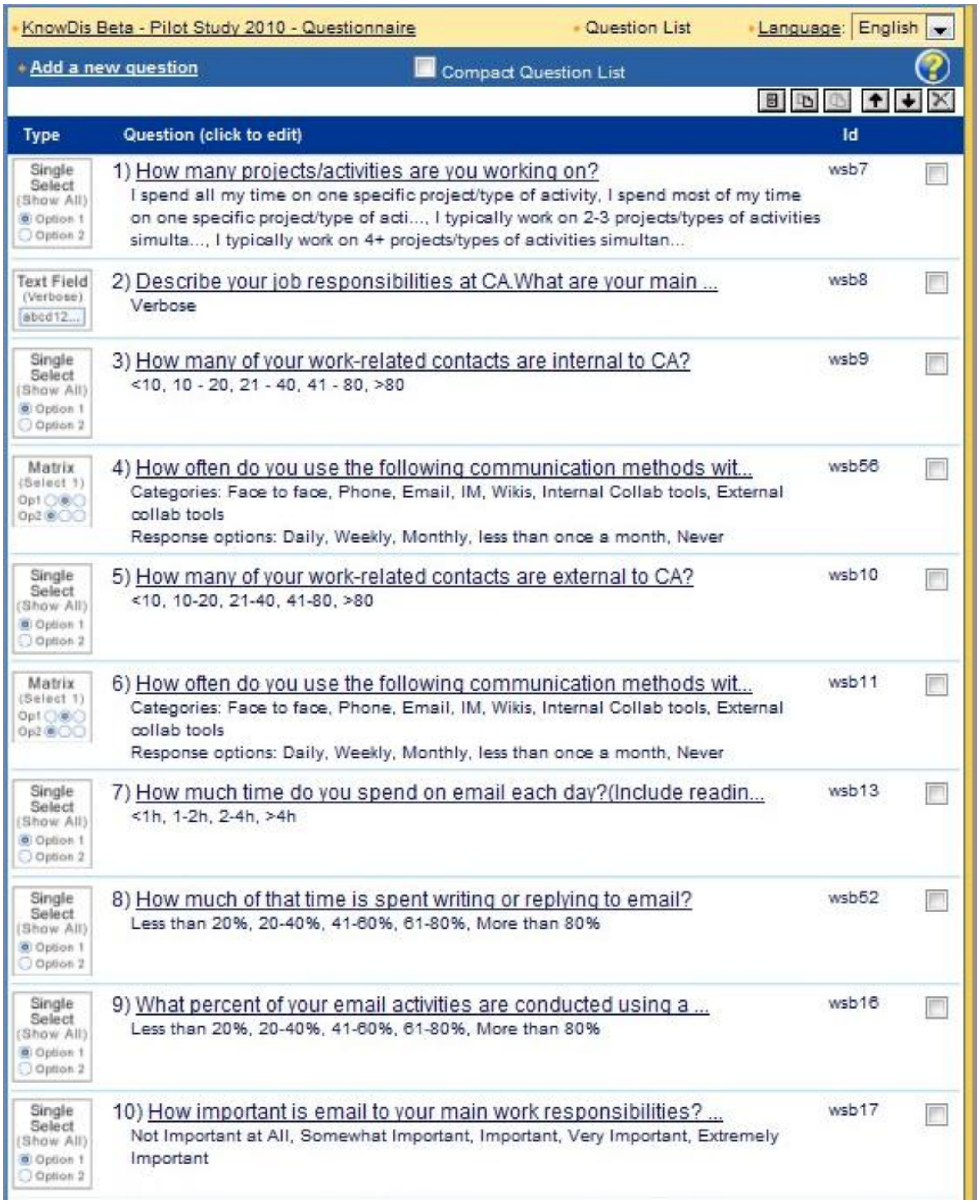
c) Do you search for that information on your PC, the CA Intranet, or the general Internet?

d) Are you trying to clarify something, or is it related to a task generated by, or associated with, the email?

Some more specific questions (maybe more useful for a separate questionnaire)

Do you use Outlook?
How do you read your email – in preview pane, or open each email in separate window?
What screen setup do you usually work with?
Do you work in the office, or remotely?
If you work remotely, are you usually connected to CA via VPN?

10.9. Phase 2 - Pre-study Questionnaire



KnowDis Beta - Pilot Study 2010 - Questionnaire · Question List · Language: English

Add a new question · Compact Question List

Type	Question (click to edit)	Id
Single Select (Show All)	1) How many projects/activities are you working on? I spend all my time on one specific project/type of activity, I spend most of my time on one specific project/type of acti..., I typically work on 2-3 projects/types of activities simulta..., I typically work on 4+ projects/types of activities simultan...	wsb7
Text Field (Verbose)	2) Describe your job responsibilities at CA.What are your main ... Verbose	wsb8
Single Select (Show All)	3) How many of your work-related contacts are internal to CA? <10, 10 - 20, 21 - 40, 41 - 80, >80	wsb9
Matrix (Select 1)	4) How often do you use the following communication methods wit... Categories: Face to face, Phone, Email, IM, Wikis, Internal Collab tools, External collab tools Response options: Daily, Weekly, Monthly, less than once a month, Never	wsb56
Single Select (Show All)	5) How many of your work-related contacts are external to CA? <10, 10-20, 21-40, 41-80, >80	wsb10
Matrix (Select 1)	6) How often do you use the following communication methods wit... Categories: Face to face, Phone, Email, IM, Wikis, Internal Collab tools, External collab tools Response options: Daily, Weekly, Monthly, less than once a month, Never	wsb11
Single Select (Show All)	7) How much time do you spend on email each day?(Include readin... <1h, 1-2h, 2-4h, >4h	wsb13
Single Select (Show All)	8) How much of that time is spent writing or replying to email? Less than 20%, 20-40%, 41-60%, 61-80%, More than 80%	wsb52
Single Select (Show All)	9) What percent of your email activities are conducted using a ... Less than 20%, 20-40%, 41-60%, 61-80%, More than 80%	wsb16
Single Select (Show All)	10) How important is email to your main work responsibilities? ... Not Important at All, Somewhat Important, Important, Very Important, Extremely Important	wsb17

Single Select (Show All)	11) How would you rate your efficiency in acting upon work-relat... not efficient , Somewhat efficient , Efficient , Very efficient , Extremely efficient	wsb18
Matrix (Select 1)	12) When acting upon an email in Outlook (replying, forwarding, ... Categories: relevant info from previous email, relevant info from local files , relevant info from Web , relevant info from CA Response options: Extremely efficient , Very efficient , Efficient , Somewhat efficient , not efficient	wsb53
Matrix (Select 1)	13) Please indicate your level of agreement with the statements ... Categories: I spend a lot of time just trying to figure out where to fin..., Relevant information is often difficult to find when it is c..., I feel overwhelmed by the amount of information I need to se..., It is easy to find the expert who has the necessary knowledg..., I feel I don't know what's happening in others parts of the ..., I sometimes ignore information that is broadcast by CA only ..., Relevant information is often found by coincidence when look..., Relevant information is often found by having a chance conve..., Whenever I'm conducting a search, I'm confident that I can f... Response options: Strongly Agree, Agree, Somewhat Agree, Somewhat Disagree, Disagree, Strongly Disagree	wsb55
Data Block	14) Please categorize what types of email you are receiving:(all... Fields: Company Business, Strategic, etc., Purely Personal, Personal but in professional context (e.g. it was good worki..., Logistic Arrangements (meeting scheduling, technical support..., Employment arrangements (job seeking, hiring, recommendation..., Document editing/checking	wsb20
Single Select (Show All)	15) How much of your work-time do you spend searching for inform... Less than 20%, 20-40%, 41-60%, 61-80%, More than 80%	wsb23
Data Block	16) Where do you find the information you are looking for?Please... Fields: In emails on my computer, In documents on my computer, On the CA network (e.g. Wiki, SharePoint, database), On the Web (e.g. developer forum, competitor's web site), In an individual (so you need to identify that person and co..., In a place other than the above.	wsb33
Single Select (Show All)	17) How many unique work-related documents do you "touch&qu... Less than 10 documents, 10 - 20 documents, 21 - 50 documents, 51 - 100 documents, More than 100 documents	wsb28
Multiple Select	18) When you share information with others, what format is used?... Summary notes, Links to relevant documents, White papers, Step by step instructions, Emails, Phone calls, IMs, Blogs, Wikis	wsb34
Single Select (Show All)	19) How well does the information you receive in emails (Word, P... Never meets my information needs, Occasionally meets my information needs, Meets my information needs about half the time, Mostly meets my information needs, Always meets my information needs	wsb37
Single Select (Show All)	20) How well does the information you find on CA's corporate web... Never meets my information needs, Occasionally meets my information needs, Meets my information needs about half the time, Mostly meets my information needs, Always meets my information needs	wsb58

Single Select (Show All) Option 1 Option 2	21) How often do you come across work-related information after ... Never, Once in a while, At least once a month, At least once a week, Every work day	wsb29	
Matrix (Select 1) Opt1 Op2	22) Please evaluate the following statements as it applies to yo... Categories: The number of projects/activities I'm involved in is:, The time I spend reading and sorting email overall is:, The time I spend writing or replying to email is:, The time I spend searching for relevant information in order..., The time I spend searching for work-related information on t..., The time I spend retrieving work-related information from Em..., The time I spend retrieving work-related information from Io..., The time I spend searching for work-related information on t..., The time I spend seeking work-related information/answers fr... Response options: Appropriate , 4, A bit too much , 2, Way too much	wsb60	
Text Field (Limited) abcd1234	23) Please enter your name (last name, first name): Text entry of 100 characters	wsb45	
Text Field (Limited) abcd1234	24) Please enter your email address: Text entry of 100 characters	wsb46	
Single Select (Show All) Option 1 Option 2	25) Please select your age range: Under 18, 18 - 24, 25 - 34, 35 - 44, 45 - 54, 55 - 64, 65 or older	wsb59	
Single Select (Horizontal) Op1 Op2	26) Please enter your gender: Female, Male	wsb48	
Single Select (Show All) Option 1 Option 2	27) How would you rate your expertise with computers? Expert, Above average, Average, Below average, Beginner	wsb49	
Single Select (Show All) Option 1 Option 2	28) How long have you been working for CA? Less than a year, 1-2 years, 2-5 years, More than 5 years	wsb50	
Single Select (Show All) Option 1 Option 2	29) Do you work from home or in one of CA's offices? Mostly from home, A bit of both, Mostly in a CA office	wsb51	
Multiple Select Option 1 Option 2	30) During the KnowDis user study, how would you prefer us to st... Email, Phone, Communicator / IM	wsb57	

10.10. Phase 2 - Post-study Questionnaire

KnowDis Beta - Pilot Study 2010 - Closing questionnaire | Question List | Language: English

Add a new question | Compact Question List

Type	Question (click to edit)	Id
Single Select (Show All)	1) How much time do you spend on email each day?(Include readin... <1h, 1-2h, 2-4h, >4h	wsb13
Single Select (Show All)	2) How much of that time is spent writing or replying to email? Less than 20%, 20-40%, 41-60%, 61-80%, More than 80%	wsb52
Single Select (Show All)	3) What percent of your email activities are conducted using a ... Less than 20%, 20-40%, 41-60%, 61-80%, More than 80%	wsb16
Single Select (Show All)	4) How important is email to your main work responsibilities? ... Not Important at All, Somewhat Important, Important, Very Important, Extremely Important	wsb17
Single Select (Show All)	5) How would you rate your efficiency in acting upon work-relat... not efficient , Somewhat efficient , Efficient , Very efficient , Extremely efficient	wsb18
Matrix (Select 1)	6) Please indicate your level of agreement with the statements ... Categories: I spend a lot of time just trying to figure out where to fin..., Relevant information is often difficult to find when it is c..., I feel overwhelmed by the amount of information I need to se..., It is easy to find the expert who has the necessary knowledg..., I feel I don't know what's happening in others parts of the ..., I sometimes ignore information that is broadcast by CA only ..., Relevant information is often found by coincidence when look..., Relevant information is often found by having a chance conve..., Whenever I'm conducting a search, I'm confident that I can f... Response options: Strongly Agree, Agree, Somewhat Agree, Somewhat Disagree, Disagree, Strongly Disagree	wsb55
Single Select (Show All)	7) How much of your work-time do you spend searching for inform... Less than 20%, 20-40%, 41-60%, 61-80%, More than 80%	wsb23
Data Block	8) Where do you find the information you are looking for?Please... Fields: In emails on my computer, In documents on my computer, On the CA network (e.g. Wiki, SharePoint, database), On the Web (e.g. developer forum, competitor's web site), In an individual (so you need to identify that person and co..., In a place other than the above.	wsb33
Single Select (Show All)	9) How many unique work-related documents do you "touch&qu... Less than 10 documents, 10 - 20 documents, 21 - 50 documents, 51 - 100 documents, More than 100 documents	wsb28
Single Select (Show All)	10) How well does the information you receive in emails (Word, P... Never meets my information needs, Occasionally meets my information needs, Meets my information needs about half the time, Mostly meets my information needs, Always meets my information needs	wsb37
Single Select (Show All)	11) How well does the information you find on CA's corporate web... Never meets my information needs, Occasionally meets my information needs, Meets my information needs about half the time, Mostly meets my information needs, Always meets my information needs	wsb58

Single Select (Show All)	12) How often do you come across work-related information after ... Never, Once in a while, At least once a month, At least once a week, Every work day	wsb29
Matrix (Select 1)	13) Please evaluate the following statements as it applies to yo... Categories: The number of projects/activities I'm involved in is:, The time I spend reading and sorting email overall is:, The time I spend writing or replying to email is:, The time I spend searching for relevant information in order..., The time I spend searching for work-related information on t..., The time I spend retrieving work-related information from Em..., The time I spend retrieving work-related information from lo..., The time I spend searching for work-related information on t..., The time I spend seeking work-related information/answers fr... Response options: Appropriate , 4, A bit too much , 2, Way too much	wsb60
Matrix (Select 1)	14) Please indicate your level of agreement with the following s... Categories: I would like to have KnowDis running on Outlook after this s..., I would recommend KnowDis to others at CA, I'm a frequent user of KnowDis, I think that KnowDis should be offered to CA customers as an..., KnowDis helps me be more productive, KnowDis is fun to have on my Outlook, Having KnowDis provide recommendations automatically without..., The amount of screen real-estate KnowDis occupied in Outlook..., The visual presentation of KnowDis is appealing/appropriate., I would use KnowDis more if it was better at finding relevan..., I would use KnowDis more if it was easier to use, I would use KnowDis more if it integrated web, intranet, des..., I would use KnowDis more, but is hard to break habits (of ho..., I frequently glance at the top recommendations presented by ..., I frequently select a recommendation presented by KnowDis, When using my browser (e.g. Internet Explorer), Outlook is h..., KnowDis should only provide recommendations upon the user's ..., KnowDis disrupts my workflow Response options: Strongly Agree, Agree, Somewhat Agree, Somewhat Disagree, Disagree, Strongly Disagree	wsb61
Matrix (Select 1)	15) Whilst having KnowDis available in Outlook, and acting upon ... Categories: relevant info from previous email, relevant info from local files , relevant info from Web , relevant info from CA Response options: Extremely efficient , Very efficient , Efficient , Somewhat efficient , not efficient	wsb53
Single Select (Show All)	16) Did the user guide on the KnowDis web site help you with usi... It helped a little bit, It helped quite a lot, I did not look at this information	wsb66
Single Select (Show All)	17) Did the video guides on the KnowDis web site help you with u... They helped a little bit., They helped quite a lot., I did not look at this information	wsb67
Text Field (Verbose)	18) If you think KnowDis is beneficial to your productivity, ple... Verbose	wsb62
Text Field (Verbose)	19) If you think KnowDis is NOT beneficial to you in its current... Verbose	wsb63
Text Field (Verbose)	20) Do you have any other suggestions on how to improve KnowDis? Verbose	wsb64
Text Field (Limited)	21) Please enter your name (last name, first name): Text entry of 100 characters	wsb45
Text Field (Limited)	22) Please enter your email address: Text entry of 100 characters	wsb46

10.11. Phase 2 - Post-study Interview Script

Prepare a copy of everyone's pre-questionnaire information prior to interview and identify particular issues that should be discussed…

Warm-up (5-10 min)

1) Please talk briefly about your work at CA. What kind of tasks do you have to do on a typical day?

2) What were your expectations of KnowDis? How were those expectations met?

3) In particular, how would you describe your strategy to retrieve and manage information necessary for your work? (e.g. search usage, archival strategies)

4) Do you think KnowDis has impacted your work-strategy in any way (with regards to information retrieval and management)? If not, do you think it could have an impact on your work-strategy, if KnowDis would better meet your needs?

Drill-down (15-20 min)

5) Did you find the user settings in KnowDis useful? If yes, why? If not, why not? *(e.g. adjust CA collections, web sites to search, use context search, etc.)*

6) Did you find the keyword feature useful? If yes, why? If not, why not? How did you use it *(e.g. adjust keywords, adjust stop-lists)?*

7) Did you find the (manual) search functionality useful? If yes, why? If not, why not?

8) Did you find the UI customization features useful? If yes, why? If not, why not? *(e.g. placement of KnowDis in Outlook, size of recommendation items, font size)*

9) Did you minimize KnowDis at some point? (Why?)

10) Which changes would you like to see in KnowDis?

Reflection (5-10 min)

11) Do you remember an instance where KnowDis recommended some useful information to you? *(Use sent out summary email information for interviewee, if no example presented)*
What was that information? Why was it relevant?

12) Do you remember an instance where you were reading an email and thought KnowDis should have some relevant information but it didn't?

Do you have any other comments or questions about the user study?

10.12. KnowDis Keyword Extraction Algorithm

```
internal String ExtractKeywordsV3(string rawText, int keyword_count) {
  int minimum_length = 2; //all words with 2 or 1 character are ignored

  //GET current email text and unique ID
  string keywordText = rawText;
  string emailID = _currentItemID;

  string finalKeywords = ""; //return string

  try {

    //----------------------------------------------
    // STEP 1: REMOVE newlines
    //----------------------------------------------
    keywordText = Regex.Replace(keywordText, Environment.NewLine, " ");
    keywordText = Regex.Replace(keywordText, "\\s", " ");

    //----------------------------------------------
    // STEP 2: REMOVE urls
    //----------------------------------------------
    string urlPattern = @"((https?|ftp|gopher|telnet|file|notes|ms-
help):((//)|(\\\\))+[\w\d:#@%/;$()~_?\+-=\\\.&]*)";

    keywordText = Regex.Replace(keywordText, urlPattern, " ");

    //----------------------------------------------
    // STEP 3: REMOVE emails
    //----------------------------------------------
    string emailPattern = @"\b[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-
Z]{2,4}\b";

    keywordText = Regex.Replace(keywordText, emailPattern, " ");

    //----------------------------------------------
    // STEP 4: REMOVE special chars
    //----------------------------------------------
    foreach (string s in _specialCharacters) {
      keywordText = keywordText.Replace(s, " ");
    }

    //----------------------------------------------
    // STEP 5: REMOVE whitespaces
    //----------------------------------------------
    keywordText = Regex.Replace(keywordText, "\\040+", " ");

    //----------------------------------------------
    // STEP 6: REMOVE stop words
    //----------------------------------------------

    Settings sett = new Settings();
    string userStopWords = sett.GetSetting("general/stoplist", "empty");
    string[] userStopWordArray = userStopWords.Split(new char[] { ';' });

    if (userStopWordArray.Length > 0) {
      foreach (string s in userStopWordArray) {
        // add spacing to prevent cutting parts of words - e.g. h(and) or
(can)dy or (lot)tery
        if (s.Substring(0, 1) == "1") {
          string spaced_s = " " + s.Substring(2) + " ";
          keywordText = keywordText.ToLower().Replace(spaced_s, " ");
        }
      }
    } else {
      foreach (string s in _overusedWords) {
```

```
        // add spacing to prevent cutting parts of words - e.g. h(and) or
(can)dy or (lot)tery
        string spaced_s = " " + s + " ";
        keywordText = keywordText.ToLower().Replace(spaced_s, " ");
      }
    }

    //------------------------------------------
    //STEP 7: NAME FILTERING
    //------------------------------------------

    //Get names from contacts
    Settings userSettings = new Settings();
    string[] activeNameArray =
userSettings.GetSetting("general/activenames", "empty").Split(new char[] {
';' });

    foreach (string s in activeNameArray) {
      string spaced_s = " " + s.ToLower() + " ";
      keywordText = keywordText.ToLower().Replace(spaced_s, " ");
    }

    //Get names from _currentMailItem
    string[] currItemNameArray = getNamesFromCurrentMail();

    foreach (string s in currItemNameArray) {
      string spaced_s = " " + s.ToLower() + " ";
      keywordText = keywordText.ToLower().Replace(spaced_s, " ");
    }

    //------------------------------------------
    // STEP 8: SORTING and GROUPING of words
    //------------------------------------------

    //Turn email text into array
    String[] textArray = keywordText.Split(new char[] { ' ' });

    // Create a dictionary of keywords as keys, with the value being the
count of its occurences
    Dictionary<String, Int32> keywordCount = new Dictionary<String,
Int32>();

    Dictionary<String, Int32> spamDict = new Dictionary<String, Int32>();

    // The list will be used to sort the key value pairs
    List<KeyValuePair<String, Int32>> myList = new List<KeyValuePair<String,
Int32>>();

    // For return of string
    StringBuilder sb = new StringBuilder();

    // filter out very short words
    foreach (string s in textArray) {
      if (s.Length > minimum_length) {
        try {
          // add item to dictionary, with initial count of 1
          keywordCount.Add(s, 1);
        } catch (ArgumentException) {
          // if item exists, key is not valid, count can be set +1
          keywordCount[s] += 1;
        }
      } else {
      }
    }

    //Populate list with dictionary - as dictionary cannot be sorted
properly
    myList.AddRange(keywordCount);
```

```
    //Sort List by keyword count
    myList.Sort(
      delegate(KeyValuePair<String, Int32> kvp1,
           KeyValuePair<String, Int32> kvp2) {
        return Comparer<int>.Default.Compare(kvp1.Value, kvp2.Value);
      }
    );

    //Revers sorting order, to have most found keyword on top of list
    myList.Reverse();


    //---------------------------------------------
    //HIGH-VALUE KEYWORD PROCESSING
    //---------------------------------------------

    //not implemented

    //---------------------------------------------
    //PROCESSING THE TOP KEYWORDS FOR A SPECIFIC EMAIL
    //---------------------------------------------
    for (int i = 0; i < keyword_count; i++) {
      //There might be less words in the email than the keyword_count
      try {
        sb.Append(myList[i].Key + " ");
      } catch (System.Exception ex) {
      }
    }
    finalKeywords = sb.ToString();
    finalKeywords = finalKeywords.Remove(finalKeywords.Length - 1);

    //---------------------------------------------
    //GENERATE KEYWORD CLOUD
    //---------------------------------------------
    _currentKeywordCloud = "";
    for (int i = 0; i < _keywordCloudCount; i++) {
      _currentKeywordCloud += myList[i].Key + " ";
    }
    _currentKeywordCloud =
_currentKeywordCloud.Remove(_currentKeywordCloud.Length - 1);
    cloudKeywordData = finalKeywords + ";" + _currentKeywordCloud;

  } catch (System.Exception) {
    //If keywords for email ID have been calculated before, they are not
calculated
    //again, but retrieved from the dictionary that archives keywords for a
session

    cloudKeywordData = finalKeywords + ";" + _currentKeywordCloud;
  }
  return finalKeywords;
}
```