Post-web cognition: evolving knowledge strategies for global information environments

Alan Dix[1], Andrew Howes[2], Stephen Payne[2]

alan@hcibook.com, howesa@cardiff.ac.uk, paynes@cardiff.ac.uk

http://www.hcibook.com/alan/topics/web/

1. Computing Department, Lancaster University, Lancaster, LA1 4YR, UK
2. School of Psychology, Cardiff University, Cardiff, UK.

abstract

This paper considers the changing cognitive demands of the web as a representative of emerging pervasive and virtually instantaneous global information access environments. Because information retrieval time on the web approaches that of human memory the appropriate knowledge strategies for seeking and remembering information begin to change. There is strong existing theoretical work on search itself; however, the complete knowledge acquisition cycle ends in memorisation ready for future needs including meta-knowledge of where and how to find information. We present preliminary results in this area but argue that this requires new empirical work and cognitive models.

keywords: knowledge strategies, cognitive models, world-wide web, internet, searching, memory

Introduction

Human intelligence is based on the capacity to process, store, and retrieve information that is relevant to social, emotional, and cognitive needs. This capacity has developed and exists through interaction with an information-bearing environment, which itself is created and evolving. Human intelligence both shapes and is shaped by the information processing tools that it has created.

From this perspective the World-Wide Web is the latest addition to a line of information processing technologies that has a synergistic relationship with human intelligence. The web provides access to an unparalleled volume of information at time costs that are orders of magnitude lower than those required for traditional print media. It has recently been claimed that the web will soon be large enough to store all of the information ever codified in human history [[C02]]. Crucially the time cost of information access is now approaching comparison with the time cost of retrieval from human memory.

So the web is not just a repository of information but offers the opportunity for people to adopt radically new *knowledge strategies*. It has the potential to force people into a new relationship with knowledge by tapping into powerful, evolved desires to satisfy individual knowledge needs.

There is a pressing need to develop theoretical models of the trade-off space of knowledge strategies and knowledge needs that are facilitated by information technology (both current and proposed), and to explore their impact on the intelligence of the individuals who adopt them. This understanding of the rich interaction between humans and information will be a crucial element informing the design of future knowledge support tools.

In this paper we will examine the current state of research and practice in this area and propose a set of interrelated facets that are required for an understanding of these knowledge strategies and the cognitive processes that underlie them.

In the next section we will use a simplified model of information seeking and use to lay out the main activities of knowledge acquisition and to form a 'map' into which following discussion can be placed. We will then look at existing solutions and models: tools for aiding and visualising search; models of search, in particular information foraging theory; and models of external information use. Finally, we will examine in detail the properties of externally focused knowledge strategies including the issues of meta-knowledge (e.g. knowing where information is and how to find it). This analysis acts both as a design resource in its own right and also as a means to highlight areas needing further research.

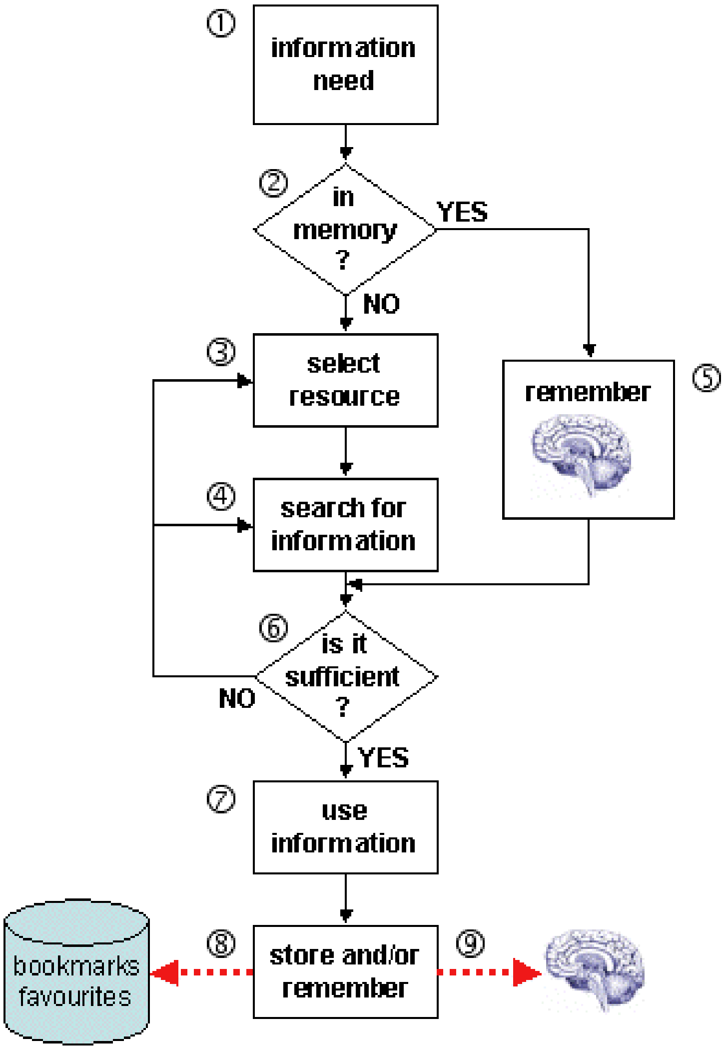

Figure 1. Knowledge acquisition

The Knowledge Acquisition Cycle

We will begin by looking at the typical activities that occur when you consult an external knowledge source. In reality any such consultation will be very context specific, depending on the nature of the problem, the nature of the knowledge sources being accessed, etc. However, we shall use a 'rationalised' model to exemplify the main features. Figure 1 shows these main activities. First of all some previous activity or chain of thought leads to a need for new knowledge (1). This may be very specific: "who won the World Cup in 1952?", or quite open: "I'd like to know more about Byron's self-imposed exile in Greece".

Sometimes the information is already known, in which case it may be retrieved from memory (5). If that retrieval is effortless enough and sufficient for the activity, the information user may never become aware of the information need and this is then perhaps a degenerate example of knowledge acquisition. At other times one does become aware of the need and has to deliberately 'think' about remembering: "I'm sure we met before ... ah yes, at the HCI conference last year". In this case there is a more explicit choice to try and remember the fact (2).

Where the information is not already known (or forgotten), then some form of explicit search becomes necessary, beginning with the choice of what resource to access (3): books, colleagues, the web. The word 'resource' here is deliberately used in quite a loose sense, as there are many potential resources available. Even focusing on the web there are 'sub-resources' available: particular web sites, bookmark lists, browser history.

Having decided on a resource, one then searches the resource in a manner suitable for it (4): asking a colleague, looking up the index of a book, using bookmarks on the web. Typically this may involve some selection of tactics: use the book index versus skimming the pages.

It may well be that this process itself prompts memories so that the two branches in figure 1 interact.

As information is retrieved, whether from memory or from external sources, there is some form of checking whether the knowledge gained reliable and sufficient for the purpose (6). Sometimes this checking happens before the actual use of the information (7) and sometimes the process of checking is implicit in the use.

Finally, there may be some incidental and/or deliberate storing of some aspect of the information, whether externally in physical or electronic form in notes, web bookmarks, etc. (8), and/or internally in memory (9). This has been drawn at the end of the process, but is likely to occur as part of the other activities; for example a bookmark is likely to be recorded at the point a useful web page is found (4) and the very act of using information (7) implicitly memorises it (the heart of constructive learning).

Notice that this act of storing or memorising itself makes information and knowledge available for future needs. We can thus see the act of knowledge acquisition as part of a wider knowledge cycle (figure 2).

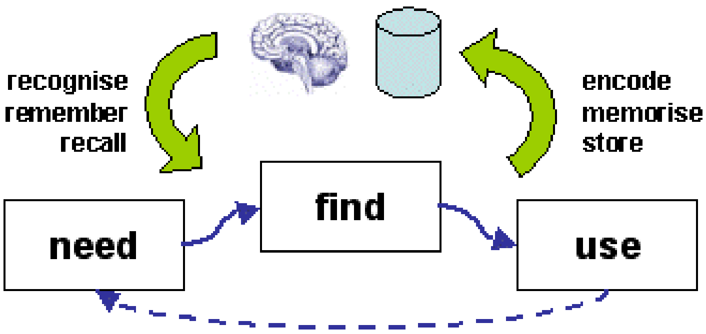

Figure 2. Knowledge cycle

The three main boxes in figure 2 represent the main phases of the knowledge acquisition process in figure 1. Note that the arrows for "encode / memorise / store" and "recognise / remember / recall" are deliberately drawn between the boxes in order to emphasise that memorisation may happen as a result of the finding or use, as well as being a separate phase. Also the recall may either be part of the 'finding' stage (e.g. if the storing is as a bookmark that needs to be accessed) or may bypass it (if all the necessary information is known). Note too that the line linking 'use' to 'need' is dashed, as this may be causal (the use of information generates new information needs) or simply temporal (a new need at a later time).

The left hand arrow has been labelled "recognise" as well as "remember / recall". This is because remembered or stored information is only useful if its existence is recognised (in the sense of familiarity) – "yes I've wanted that before, I bookmarked it". If there is no such recognition then it is likely that the information seeker will simply repeat the same finding activities as before. Of course, recognising familiarity is not sufficient on its own; the encoded information must also be retrieved: either by remembering it or by accessing an external resource such as a bookmark.

It is also important to note, and a point we will return to later, that the use of existing knowledge to inform the knowledge acquisition process goes beyond the simple recall of facts. Decisions as to which resource to use, the tactics to use in accessing the resource and whether it is worth searching at all, are each influenced by previous searches for related material.

As we look at the existing literature and models in the next section we shall see that the majority are focused on the finding phase of acquisition (the central activities of figure 1). We shall then argue that perhaps the most exciting issues, which are still largely an open area for study, are those related to the cycle in figure 2.

Where we are – current tools and models

Tools for search and visualisation

There has been extensive work on tools for aiding in the search and visualisation of information spaces in general and the web in particular. Indeed, if we look at the broader area of information spaces we would need to consider whole disciplines such as knowledge management, information retrieval and information visualisation. Even the web literature is massive and this emphasises our contention that this phase is already well studied.

The web has three coextensive geometries [[D99]] which are used for visualisation and search.

The first and most obvious geometry is that of the link structure. There are many visualisations based on drawing web pages as graph nodes in two and three dimensions. Some use simple geometric rules, perhaps breadth first traversal of pages with flower-petal layouts of daughter pages around their parents. Other visualisations use more self-organising layouts based on some form of relaxation algorithm such as ball and spring models.

All such visualisations have two problems. The first is fundamentally insoluble. The web is enormous: at the time of writing Google indexes over 2.5 billion pages. Visualisations therefore limit themselves to subsets of the web: a single site, pages 'close' to a given page, recently visited pages. The second problem is that the link structure does not lend itself well to 2D (or even 3D) representations. This is because with any link structure the number of nodes at a distance N from a node increases exponentially with N, whereas the amount of available space only increases with the square of the distance in 3D space or linearly in 2D space. The self organising layout algorithms to some extent 'solve' this implicitly, but at the expense of increasingly tortuous layout towards the periphery. One solution to this is the Hyperbolic Tree [[LR96,IX02]] that lays out a graph in hyperbolic space (where available space conveniently increases exponentially with distance). Of course, this hyperbolic space must eventually be projected into the two dimensions if the physical screen (see figure 3) giving a form of fish-eye display.

The link structure is also used for search, both in that users chase links as they browse, and also in the way Google uses it to build its automatic search facility. Google uses the pattern of links to decide which pages are the most important resources for any topic. So if you search for the term "web visualisation" Google offers first, not just those pages that mention "web visualisation", but pages that are linked from or linked to other pages on the same topic [[BP98]].

The second form of geometry is that given implicitly by content – two pages are related if they share common words. This has again been used both for visualisation and search. For visualisation this borrows pre-existing methods for layout of non-web documents including clustering-based layouts [[CC98]] and placement based on proximity to target documents [[PCMT01]]. For search this is the closest area to traditional information retrieval mechanisms and the way that most search engines work, as well as more proactive "these pages are like the one you are viewing" mechanisms such as Alexa [[L98]]. The scatter-gather browser is both a (textual) visualisation and a search interface, using word co-occurrence for clustering of documents and choice of 'typical' documents from each cluster as representatives of the content of the cluster [[CKPT92,PSHD96]].

Figure 3. Inxight Hyperbolic Tree [[IX02]]

The third type of geometry is that defined by people themselves. Two pages are close if they are frequently viewed by the same person or in the same session. This is the basis of 'recommender' systems that say "other people who liked XXX also liked YYY" [[AAAI98,RV97]], and of course is also used heavily in e-commerce for product suggestions.

As well as these explicit and implicit geometries of the web there are external structures given by numerous classified lists such as Yahoo [[Y02]] or the Open Directory project [[ODP02]].

These different geometries are not independent: links typically lie between material of a similar content and users visit pages that are similar; also pages that are ranked highly by search engines such as Google or Yahoo tend to be both visited more and visited more often together.

Understanding information seeking: information foraging theory and scent

One of the most developed theoretical frameworks in this area is the information foraging theory of Pirolli and Card [[PC99]]. This is focused principally on the information seeking part of the knowledge acquisition process. They use an analogy from computational biology and see the information user as a foraging animal seeking to browse (in the eating sense) the food source that is information. The animal has to constantly weigh up the advantages of continuing to eat and consume the food source in one place against seeking potentially richer sources elsewhere. Of course the act of going in search of further food, however much better it might be, will itself consume energy. Similarly the information browser needs to decide whether to stick with narrower sources of information or to seek better information. Of course the act of seeking further information, however much better it might be, will itself consume valuable time and effort.

Pirolli and Card use the same equations developed by biologists and apply them to information users to determine optimal browsing strategies, and find that this to some extent matches observed user behaviour.

An essential cue for animals is the scent of food that enables them to predict abundance of food in nearby areas. Similarly the efficiency of information seeking as an exploratory, goal directed activity is improved if the information systems give the users some 'scent' or indication of the utility or value of taking particular paths: for example, more meaningful links and better page organisation. This mirrors long-standing work in AI which emphasises the importance of heuristics in search: measures of the predicted value of making various choices [[JD96]]. Without appropriate heuristics larger search spaces (for example chess playing programs) are impossible.

History and revisitation

Studies have shown that there is considerable revisiting of the same web page during browsing [[TG97]]. Some of this is clearly due to backing up after following mistaken links or hub-and-spoke behaviour; indeed 30% of all navigation is the use of the 'back' button [[CP95]]. However, there are also a considerable residual number of pages that are 'really' revisited because the user wants to see the content again. Browsers support this behaviour both for short-term revisitation (back button and visit stack) and for the long term (history, bookmarks, favourites). In a formal analysis of several hypertext and web browsers it was found that the history and back mechanisms were subtly different in them all [[DM97]]. This emphasises results found in other studies that users find back and history confusing and this is reflected in behaviour with comparatively little use of history or multi-step back. Bookmarks are more heavily used, but still are known to have many problems. There have been some more radical interfaces proposed and used at an experimental level including the data mountain, which allows users to arrange thumbnails of bookmarked pages in a 2D landscape [[RC98]], and Kaasten and Greenberg's interface unifying history and bookmarks [[KG01,KGE02]].

Externalised knowledge

To some extent the web can be seen as part of a continuous externalising of human knowledge. Simple cave drawings date back tens of thousands of years and counting sticks are near universal. Language itself is a form of externalising, by sharing knowledge amongst the social group, but it is with written language that we see the ability to off-load memory into artefacts in an extensible manner. In oral societies all knowledge must be in human heads, although often the role of group memory is focused on key individuals: the bard, medicine man or scholar. This externalisation is still evident when an executive relies on a personal assistant (human!) to remember dates, or a meeting entrusts the secretary to take minutes. However, it is when this externalisation is in the form of artefacts (physical or electronic) that the external representation can be used as a resource for memory augmentation beyond human memory.

Ethnographic studies have shown that artefacts are central to the smooth running (or running at all) of many work environments [[H95]]. Physical artefacts fulfil many roles. Some of these roles are ephemeral: communication between team members and the organisation and coordination of ongoing work. However, they also function as part of the organisation's 'memory' of what is happening, what needs to happen and who is doing what [[D98,DRROSM02]].

Many of the models of cognition used in HCI, such as GOMS and many uses of task analysis (see reviews in [[DFAB98]]), have been, and still are, very 'output' oriented: users' goals and intentions are translated into actions, with little focus on the reaction of users to the environment. Distributed cognition and situated action were two responses to this bias that have become very influential in HCI. Situated action positions itself against the idea of users planning and then executing their actions, and instead sees users as acting and re-acting dynamically in a changing physical and social context [[S87]]. Distributed cognition is perhaps a more radical paradigm shift. It challenges the Cartesian view of thinking happening within the head based on mental models of the world, and instead positions cognition in the interactions between humans and things in the world itself [[H90b]]. One of the classic studies here was of Polynesian navigation between islands: although in some sense none of the sailors 'know' the route, somewhere in the interworking of the crew, the boat, simple materials and the environment, they are able to successfully navigate many hundreds of miles across the open seas.

Perhaps most pertinent for the web has been the use of distributed cognition to understand basic mathematics. Although we may solve problems 'in our heads', we also often use materials such as calculators or paper, and may actually use the things being calculated themselves, perhaps holding a piece of word against another to compare lengths. In the use of paper particularly we have a case where the external medium is used to store temporary information that is not committed to memory. That is an off-loading of internal memory to an external medium.

As well as these more radical approaches there have been various efforts to make more traditional cognitive models capture more closely the interactions with the real world, for example Larkin's display-based problem solving [[L89]], and DTAG, which extends TAG to include the feedback from computer displays into a planning cycle [[HP90]. Indeed, Payne, Howes and Reader [[PHR01]] argue that the foundational assumptions of cognitive science are sufficient to explain the signature phenomena of distributed cognition. Interestingly one of the foci of display-based cognitive modelling has been calculator use.

New challenges of externalised knowledge strategies

Although we have seen there is promising work leading towards an understanding of the knowledge acquisition cycle for the web, this is still limited, and in particular the feedback nature of the cycle is hardly dealt with at a theoretical level. We can identify a number of issues that need to be addressed in order to complete this picture, each of which involves one potential consequence of dramatically lower-cost access to information on knowledge strategies: just-in-time knowledge; knowledge reconstruction; changes to encoding strategies; changes to the value of information; positive feedback.

Just-in-time knowledge

It is conceivable now to operate a just-in-time strategy for knowledge acquisition: choosing just when to access information and how much effort to put into learning. While the ability to process and use knowledge is currently more valued by society, ironically it may now be less important to commit facts to memory. Deploying a just-in-time strategy may have beneficial effects on the quality of information deployed but would also clearly carry risks. For example, a just-in-time strategy depends on individuals being able to identify and satisfy knowledge needs at the appropriate time.

The dynamics of the information environment

The information available to an individual on any particular topic changes with time. New empirical findings are added to scientific databases; new economic figures are calculated and released; news is generated. It is possible that the value of ‘staying in touch’ will change as information is accessed faster.

Knowledge reconstruction

Information retrieval from human memory can take as little as 50ms but may take much longer. In the extreme, hours may be required for information to be accessed by non-deliberative associative mechanisms, but more often a few seconds or minutes are required in order to elaborate potential cues, i.e. by reconstructing a retrieval context for the information that is required. While information can rarely be accessed in milliseconds using current search engines, it can be accessed in seconds. The overlap in the distribution of time costs for these external and internal retrieval mechanisms means that people now have a subtle choice between retrieving using a reconstructive mental process or retrieving with the aid of information technology. How is this choice made? It is as certain as it can be that the balance in favour of the technology will shift as bandwidth is increased, search engines compute better measures of relevance, and wireless devices become more available.

Encoding strategies

As the costs of web-based information access lowers, people will not only have a choice about how to retrieve information at the time of need but will also have a choice about how much effort to put into encoding information when they are exposed to it, and how much effort to put into abstracting or storing local copies of information (e.g. making notes). These decisions will be moderated by factors such as the likely projected future need for the information and the expected cost of retrieval from external and internal resources. People will also have a choice about what information to put effort into encoding. The cost/benefit equation for encoding the location of information will be different to the equation for the information itself, and in addition there will be choices about how to encode ‘location’, which, in a partially content addressable medium such as the web, may involve deliberate reshaping of conceptual structures to reflect those found in the information environment.

The value of information

People seek information in response to needs. These needs derive in the first instance from the fundamental needs of the human condition but more locally they derive from gaps and conflicts in available knowledge. People will seek information when its expected value exceeds the expected cost.

But how do people assign value to information? Work in micro-economics has analysed various types of commodity, but there is little work on the immediate cognitive assessment of the value or utility of information itself. The recent approach to subjective utility by Kahneman and colleagues [[K00]] in the field of behavioural decision making ("evaluation by moments") might fruitfully be extended to knowledge workers' evaluation of information by exploring the patterns in readers' local judgements of interest or pleasure, and testing how these relate to patterns of access and use.

As the cost of information access decreases it might be expected that people will seek more information, but a complicating factor, as revealed by economic theory, is that the value of information is sometimes subject to a variety of "network effects" [[WWW00,LM98]]. For example, the value of information is sometimes dependent on how many people have access to it – in many business contexts information is more valuable if fewer people have access to it. Such network effects suggest that the relationship between cost of access and amount of information accessed is non-linear. And as judgements of human intelligence are often relative to other people and to the available resources, there are concomitant consequences for understanding the nature of human intelligence in an information rich society.

Positive feedback

Another consequence of faster access to information is an increased prevalence of positive feedback in information seeking behaviour. In contrast to most other commodities, satisfying an information need can itself beget more information needs. This is a major challenge for the economic models of information seeking that we have alluded to above and that are the basis of some important existing approaches to information seeking, such as the work on information foraging by Pirolli and Card [[PC99]].

Meta-knowledge

A recurrent theme through the above issues is the importance of meta-knowledge – not just knowing things or finding knowledge, but knowing what can be found, how to find it, and if you have found it once, whether it is worth remembering it rather than simply finding it again. The knowledge gained may also be about the process of finding out, not the results of that finding.

The importance of this higher level knowledge has been noted in related fields. Kidd emphasises that the crucial role of information for knowledge workers (and by extension those acting temporarily in such roles) is the transformation of the subject's own thought patterns [[K94]]. In "The marks are on the knowledge worker", she found that it was rare that previous facts were required – the filing cabinet is essentially write-only! Instead the way in which the subject approached future problems was subtly influenced by past information. The significant thing is not the marks left on the paper, but those on the knowledge worker's own mind.

Talking informally with expert web users it is evident that they both employ a variety of strategies and are able to *articulate* these strategies: "I don't bookmark X, because I know I can always find it using Google". One of the authors recently talked to someone who wanted to point to his home page, but instead of giving his URL said "it's easiest just to type 'John+Smith' into Google" (anonymised). In observations of students' browsing behaviour McManus also found that they employed sophisticated strategies to deal with download delays [[M97]]. These all portray not just models of information, but also models of their own and others' information seeking – meta-knowledge. The models of less expert users may be less well informed about mechanisms, but, given general human model-making behaviour, it is fairly certain that novices will also have process models of some sort.

In figure 1 we drew a general knowledge acquisition process; however, it seems more likely that the models used by users will be both more domain specific (finding telephone numbers different from finding email addresses) and more access method specific (using a library different from searching the web). However, several web access designs use metaphors of bookshelves and books, for example the web book [[CRY96]], implicitly assuming that users will have similarities between their models of different access methods, or at least are able to use one as a metaphor for another.

Whereas these book metaphors are exploiting largely unconscious models of search, work by Walkerdine and Rodden has looked at the explicit sharing of search terms [[ER01]]. This is exploiting two things: first that the discovery of collections of search terms is often itself significant and worth passing on to others, and second that these search terms are often usable between different search engines. The search terms embody a set of words that are sufficient to elicit a large enough class of relevant entries whilst excluding the irrelevant. The latter is perhaps particularly important in sifting through the vast volume of uninterpreted data that is often the result of what appear to be quite reasonable search requests. Indeed it is not uncommon to see links on web pages that are not links to particular web researches, but instead a 'live' link to a search engine with particular keywords. It is interesting that whereas this appears to be a new phenomenon it was one of the key skills that library staff learnt when dealing with some of the earliest online bibliographic databases. In those days connection was via telephone modem with expensive telephone and database connection charges. Letting 'ordinary' users loose on the system would inevitably lead to overlong lists of results and long sessions as they tried to refine the searches from 1000s of results without over-confining them and getting empty results. The librarians learnt the knack of choosing appropriate terms and rapid refinement although this was a higher order domain independent skill and often tuned to particular databases.

Recent empirical work adds further support for users' ability to articulate these meta-knowledge issues [[M02]]. In exploratory interviews aimed primarily at visualisation requirements for web history, an issue which repeatedly arose was the interviewees' desire to be able to classify bookmarks/favourites at the moment they were 'remembered' rather than as a secondary exercise. Currently all common browsers force this bookmark-now, sort-later mode of working. This led directly to experiments looking at sorting during browsing compared with sorting afterwards and its effect on later (online) recall. It was postulated that although the interviewees had expressed a desire for 'during' sorting this would lead to less clear classifications as the participants would not know what was still coming. This was borne out by the results which did show significantly better recall for the 'after' condition. However, in a small number of retests a week later this advantage appeared to disappear almost completely and may have simply been due to recency effects. In a post-test questionnaire the participants preferred the 'after' sorting, in direct contrast to the interviewees' imagined preference. However, again this result needs to be read with care as the 'during' condition used an interface initially designed for 'after' sorting!

There are two main lessons from this last experiment. First the interviewees were able to articulate desires relating to meta-knowledge issues – the timing of bookmark classification. Second, the actual running of the experiment also showed how complex the issues are, especially because we are looking at relatively long-term effects that are hard to capture fully within a laboratory setting.

Summary – the way forward

We have argued that the web and related forms of rapid information access are changing the nature of the relationship between people and knowledge. This is not the first such change. Writing made it possible to keep tallies, lists, etc. – the management of human and physical resources that were previously intractable now became subject to reason and control . Print in contrast changed mainly the inter-relationship between people – the knowing and the unknowing blurred by the use of the book. Whereas the web is seen largely as an extension of the latter, the issues we have raised suggest that its effect is more like the former – like writing substantially changing the modes of cognition itself.

Although we can recruit existing work from HCI and cognitive science, the more specific issues that we have highlighted are to date understudied. There are some notable exceptions, most significantly information foraging theory, but the majority of the existing work, including this, is focused on the search phase of the knowledge acquisition cycle. However, the new issues we have highlighted concern aspects of the cycle as a whole including the important feedback paths.

A full analysis of these knowledge strategy issues will require a combination of techniques: more ecologically valid studies of actual use, knowledge elicitation from real users, controlled semi-real experiments like this and more 'traditional' psychological studies of fine issues.

This paper is mainly laying out this area and we look forward to further work leading towards a more complete framework to address the peculiar properties of information and information technologies which are the matrix of our global life.

Acknowledgements

This work has been supported by the EPSRC under the EQUATOR Interdisciplinary Research Collaboration. and grant GR/M43302 ("Cognitive strategies and design heuristics for dealing with information overload on the web"). Thanks also to Jason Marshall for preliminary acxces to has research results.

References

[[AAAI98]] AAAI-98 Workshop on Recommender Systems, 1998. http://rsv.ricoh.com/~marko/rec98/

[[BP98]] S. Brin and L. Page (1998). "The Anatomy of a Large-Scale Hypertextual Web Search Engine," Proceedings of the 7th World-Wide Web Conference, 1998. http://www7.scu.edu.au/programme/fullpapers/1921/com1921.htm

[[CRY96]] S. K. Card, G. G. Robertson, and W. York. The WebBook and the Web Forager: an information workspace for the world-wide web. In Proceedings of CHI'96, pages 111-117. ACM Press, New York, 1996.

[[C02]] S.K. Card (2002) Cognitive science as the engine of innovation: beyond human-computer interaction. In W. Gray and C. Schunn (Eds.) Proceedings of the 24th Annual Conference of the Cognitive Science Society. August 7-10, George Mason University, Fairfax, VA.

[[CP95]] L. Catledge and J. Pitkow, "Characterizing browsing strategies in the World-Wide Web",. Proceedings of the Third International World Wide Web Conference, Darmstadt, Germany, 1995. http://www.igd.fhg.de/www/www95/papers/

[[CC98]] C. Chen and M. Czerwinski, "From Latent Semantics to Spatial Hypertext – An Integrated Approach", Hypertext'98, ACM Press, pp. 77–86, 1998.

[[CKPT92]] Cutting, D.R., Karger, D.R., Pedersen, J.O. and Tukey, J.W. A cluster-based approach to browsing large document collections. Proceedings SIGIR '92, Copenhagen, 1992, ACM Press, 318–329

[[DM97]] A. Dix and R. Mancini (1997). Specifying history and backtracking mechanisms. In Formal Methods in Human-Computer Interaction, Eds. P. Palanque and F. Paterno. London, Springer-Verlag. pp. 1-24. http://www.comp.lancs.ac.uk/computing/users/dixa/papers/histchap97/

[[DFAB98]] A. Dix, J. Finlay, G. Abowd and R. Beale (1998). Human-Computer Interaction, second edition. Prentice Hall. http://www.hcibook.com/

[[D98]] A. Dix, J. Wilkinson and D. Ramduny (1998). Redefining Organisational Memory – artefacts, and the distribution and coordination of work. In Understanding work and designing artefacts (York, 21st Sept., 1998). http://www.hiraeth.com/alan/papers/artefacts98/

[[D99]] A. Dix (1999). Design of User Interfaces for the Web (invited paper), User Interfaces to Data Intensive Systems - UIDIS 1999, eds. N. Paton, T. Griffiths, Edinburgh 5th - 6th Sept. 1999. IEEE Computer Society, 1999. pp. 2-11 http://www.hiraeth.com/alan/papers/UIDIS99/

[[DRROSM02]] A. Dix with D. Ramduny, P. Rayson, V. Ochieng, I. Sommerville and A. Mackenzie (2002). artefacts speak and artefacts to speak Position paper for "Analyzing Collaborative Activity" - CSCW 2002. http://www.hcibook.com/alan/papers/cscw-artefact-2002

[[JD96]] J. Finlay and A. Dix (1996). An Introduction to Artificial Intelligence. UCL Press / Taylor and Francis. http://www.hiraeth.com/books/ai96/

[[H90]] A. Howes and S. Payne (1990). Display-based competence: towards user models for menu-driven interfaces. Int. J. of Man-Machine Studies, 33:637-655.

[[H95]] J. Hughes, J. O'Brien, M. Rouncefield, I. Sommerville and T. Rodden (1995). Presenting ethnography in the requirements process. In Proc. IEEE Conf. on Requirements Engineering, RE'95. IEEE Press, pp. 27–34.

[[H90b]] E. Hutchins (1990), The Technology of team navigation. In Intellectual teamwork: social and technical bases of collaborative work. Gallagher, J., Kraut, R. and Egido, C., (eds.), Lawrence Erlbaum.

[[IX02]] Inxight Software (2002). Hyperbolic Tree. http://www.inxight.com/

[[KG01]] Kaasten, S. and Greenberg, S. Integrating Back, History and Bookmarks in Web browsers. Extended Abstracts of ACM CHI'01, 379-380. 2001.

[[KGE02]] S. Kaasten, S. Greenberg. and C. Edwards (2002) How People Recognize Previously Seen WWW Pages from Titles, URLs and Thumbnails. In X. Faulkner, J. Finlay, F. Detienne (Eds) People and Computers XVI (Proceedings of Human Computer Interaction 2002), BCS Conference Series, 247-265, Springer Verlag.

[[K00]] Kahneman, D. (2000) Evaluation by moments. In D. Kahneman and A. Tversky (Eds) Choices, values and frames. Cambridge University Press.

[[K94]] A. Kidd (1994). "The Marks are on the Knowledge Worker." Human Factors in Computing Systems: CHI'94 Conference Proceedings (Boston, MA, April 24--28, 1994), pp. 186--191. ACM Press: New York, NY, 1994.

[[LR96]] Lamping, J. and Rao, R. Visualizing Large Trees Using the Hyperbolic Browser. Proceedings CHI '96, Vancouver, April 1996, ACM Press, 388-389

[[L89]] Larkin, J H. Display-based problem solving. In Complex Information Processing: The impact of Herbert A Simon, D Klahr and K Kotovsky, Eds. Erlbaum, Hillsdale NJ, 1989. 319-341.

[[LM98]] S. Liebowitz and S. Margolis (1998). "Network Externalities (Effects)" entry in The New Palgraves Dictionary of Economics and the Law, MacMillan, 1998.

[[L98]] C. Lohr, Alexa Internet Donates Archive of the World Wide Web to Library of Congress. Alexa Internet Press Release, 13th Oct. 1998. http://www.alexa.com/company/inthenews/loc.html

[[M97]] . McManus (1997). Compensatory Actions for Time Delays. Time and the Web, Staffordshire University, 19th June 1997. (also reported in A. Dix. Time and the Web (report). SIGCHI Bulletin, 30(1) 1998 pp. 30–33) http://www.hiraeth.com/conf/web97/papers/barbara.html

[[M02]] J. Marshall (2002). An Exploratory Study in World Wide Web Navigation History. MRes Dissertation. Lancaster University, UK.

[[ODP02]] Open Directory Project (2002). http://dmoz.org/

[[PHR01]] Payne, S.J., Howes, A., Reader, W.R. (2001). Adaptively distributing cognition: a decision-making perspective on human-computer interaction. Behaviour and Information Technology, 20, 5, 339-346.

[[PCMT01]] Pettifer,S., Cook,J., Mariani,J., Trevor, J.,(2001),Q-Space: A Virutal Environment for Interactive Visualisation of Abstract Data, Proceedings of IEEE Virtual Reality 2001, eds Takemura,H., Kiyokawa,K., Yokohoma, Japan, ISBN: 0-7695-0948-7, pp.293-294, March 2001.

[[PSHD96]] Pirolli, P., Schank, P., Hearst, M. and Diehl, C. Scatter/Gather browsing communicates the topic structure of a very large text collection. Proceedings CHI '96, Vancouver, May 1996, ACM Press, 213–220

[[PC99]] Pirolli, P., & Card, S. K. (1999). Information foraging. Psychological Review, 106, 643-675.

[[RV97]] Resnick, P. and H.R. Varian (guest editors), Special Issue on Recommender Systems. CACM. 40(3):56–89, 1997.

[[RC98]] Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D. & van Dantzich, M. (1998). Data Mountain: Using Spatial Memory for Document Management, Paper to appear in Proceedings of ACM UIST '98 Symposium on User Interface Software & Technology, November, San Francisco, CA.

[[S87]] L. Suchman (1987). Plans and Situated Actions: The problem of human–machine communication. Cambridge University Press,.

[[TG97]] L. Tauscher and S. Greenberg, "How people revisit web pages: empirical findings and implications for the design of history systems", International Journal of Human Computer Studies, 47(1), 1997. http://www.cpsc.ucalgary.ca/grouplab/papers/1997/

[[WR01]] James Walkerdine, Tom Rodden (2001). "Sharing Searches: Developing Open Support for Collaborative Searching", In the proceedings of Interact 2001, Japan, 9th-13th July 2001

[[WWW00]] T. Weitzel; O. Wendt, and F. von Westarp (2000). Reconsidering Network Effect Theory. In: 8th European Conference on Information Systems (ECIS 2000); Wien, Österreich

[[Y02]] Yahoo (2002). http://www.yahoo.com/