Investigating Sense Making and Information Triage in Web Search

Introduction

It seems obvious to state that web search is one of the most common uses of the internet but a look at Alexa’s top five most popular websites shows that three of these are search engines\(^1\). Combine this with comScore’s February 2010 study\(^2\) that tells us that 14.5 billion searches were performed by Americans in a single month and it begins to put the popularity of web search into perspective. Deborah Fallows performed a study in 2008 that found that 49% of internet users used web search daily, a number that had steadily increased since her previous study in 2002\(^1\). It can be hypothesised that the presence of “always-on” home internet connections as well as internet capabilities of mobile phones will make users more likely to turn to the internet to answer any questions they may have.

It is this existence of ubiquitous internet access that inspired my research into information seeking and the interface presented between the user and the service that is provided by the search tool. This is represented in my third year project where I looked at the use of tag clouds in the search interface and the effects this has on the learning abilities of the user. The project involved the design and creation of search interfaces and the study of twelve participants. The study itself had each participant write a brief summary of a subject they were given before performing a timed search on the subject, then summarising the subject once more. The summaries were compared for any learning effect that might have occurred during the search. While the study is still being analysed the preliminary outcome is looking like the inclusion of the tag cloud positively affects a user’s learning.

Performing the study created further questions that were linked deeper to user experience during web search, particularly related to sense making. I intend to follow up from this study and investigate how people make sense of information while searching.

Background

Sensemaking is currently a popular subject within the world of HCI with the CHI conference running sensemaking workshops in both 2008 and 2009. One paper from the 2008 workshop by Abraham, Petre and Sharp\(^2\) defined sensemaking as “…the strategies and behaviours evident when users collect, evaluate, understand, interpret, and integrate new information for their own specific problem/task needs”. They performed a study to investigate how experienced information processors interact with information sources that have been gathered following a web search. Eight participants were asked to gather information about a subject and outline a 15 minute slide presentation to ensure that they used external representations such as notes, diagrams and cut-and-paste operations. They found that the participants tended to use a number of resources at the same time, whether it is pen and paper, a word processing document or the slides themselves. Most commonly the participants used an intermediate external representation before working with the


slides directly, with the most popular being pen and paper. They conclude that tools should be designed that support these dual resources but also comment that they should support what are now everyday web based practices, providing benefits throughout the entire range of planning, searching and sensemaking.

As referenced in their conclusion, web search is not limited to just the sensemaking that is performed once suitable resources are found but also the search itself. Information triage is the process that allows a user to refine the information they gather, promoting resources that they consider most valuable to their current search as well as removing resources that have no relevance. For example, to write this proposal a search on Google for “sensemaking” returned 142,000 results while “information triage” returned 32,600. While obvious that not all those resources are needed the process of narrowing down the results and finding the most relevant is not as obvious. A folder was created to hold shortcuts to the pages that are most relevant and useful, but information triage does not stop at the search stage. Once the resources were gathered the sensemaking stage begins, but information triage still occurs. Outlining academic papers requires focused information triage on a single source, selecting information that is important to reference while leaving out details that are not pertinent. All the while making notes in a text document to keep track of what has been gathered already.

The company Cirilab deal with the information triage process to help users deal with “information overload” and outline on their blog the process in which their “speed reading” software can help. They break information triage into five steps that involve finding themes in a document, reading a synopsis, a detailed summary, finding relevant paragraphs then finally deciding whether or not to read the whole document. The software they provide runs on a local computer and analyses documents, but the inclusion of such a tool in a web search environment is certainly worth considering.

Proposal

As mentioned previously, the work undertaken in my third year project led to further questions and interest in how people make sense of information when performing web searches. This project would primarily involve further research into sensemaking as well as information triage and the relation between the two. Investigation into both previous and current tools available and employed by users to aid these processes will drive toward my own understanding of the software design involved and will hopefully lead to my own undertaking of developing and studying such a tool.

References

