A Model of Consumer Search Behaviour

Tony Russell-Rose UXLabs London UK +44 (0)7779 936191 tgr@uxlabs.co.uk Stephann Makri University College London Interaction Centre, University College London, Gower St. London, WC1E 6BT, UK +44 (0)20 7679 0696 s.makri@ucl.ac.uk

ABSTRACT

In order to design better search experiences, we need to understand the complexities of human information-seeking behaviour. In previous work [13], we proposed a model of information behavior based on an analysis of the information needs of knowledge workers within an *enterprise search* context. In this paper, we extend this work to the *site search* context, examining the needs and behaviours of users of consumeroriented websites and search applications.

We found that site search users presented significantly different information needs to those of enterprise search, implying some key differences in the information behaviours required to satisfy those needs. In particular, the site search users focused more on simple "lookup" activities, contrasting with the more complex, problem-solving behaviours associated with enterprise search. We also found repeating patterns or 'chains' of search behaviour in the site search context, but in contrast to the previous study these were shorter and less complex. These patterns can be used as a framework for understanding information seeking behaviour that can be adopted by other researchers who want to take a 'needs first' approach to understanding information behaviour.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Search process; H.3.5 [Online Information Services]: Web-based services

General Terms

Human Factors.

Keywords

Site search, enterprise search, information seeking, user behaviour, search modes, information discovery, user experience design.

1. INTRODUCTION

Classic IR (information retrieval) is predicated on the notion of users searching for information in order to satisfy a particular 'information need'. However, it is now accepted that much of what we recognize as search behaviour is often not informational per se. For example, Broder [2] has shown that the need underlying a given web search could in fact be navigational (e.g. to find a particular site) or transactional (e.g. through online shopping, social media, etc.). Similarly, Rose & Levinson [12] have identified the consumption of online resources as a further common category of search behaviour.

Presented at EuroHCIR2012. Copyright © 2012 for the individual papers by the papers' authors. Copying permitted only for private and academic purposes. This volume is published and copyrighted by its editors.

In this paper, we examine the needs and behaviours of individuals across a range of site search scenarios. These are based on an analysis of user needs derived from a series of customer engagements involving the development of customised site search applications. In so doing, we extend and validate a model of information behaviours derived from a previous study of enterprise search users [13].

The model is based on a set of 'search modes' that users employ to satisfy their information search and discovery goals. It extends the IR concept of information-seeking to embrace a broader notion of discovery-oriented problem solving, addressing a wider range of information interaction and information use behaviours. The overall structure of the model reflects Marchionini's [9] framework, and consists of three lower-level 'lookup' modes (*locate, verify* and *monitor*), three "learn" modes (*compare, comprehend* and *explore*) and three higher-level "investigate" modes (*analyze, evaluate* and *synthesize*).

We investigate the degree to which the model extends to accommodate the domain of *site search* (i.e. consumer-oriented websites and search applications) and discuss some of the differences between the needs and goals of enterprise search users versus those of site search. We conclude by exploring the ways in which these modes combine to form distinct chains or patterns, and reflect on the value this offers as a framework for expressing complex patterns of behaviour.

2. MODELS OF INFORMATION SEEKING

The framework investigated in this study is influenced by a number of existing models. For example, Bates [1] identified a set of 29 search 'tactics' which she organised into four broad categories, including *monitoring* ("to keep a search on track"). Likewise, O'Day & Jeffries [11] examined the use of information search results by clients of professional information intermediaries and identified three categories of behaviour, including *monitoring a known topic or set of variables over time* and exploring a topic in an undirected fashion. They also observed that a given search scenario would often evolve into a series of interconnected searches, delimited by triggers and stop conditions that signalled transitions between modes within an overall scenario.

Cool & Belkin [3] proposed a classification of interaction with information which included *evaluate* and *comprehend*. They also proposed *create* and *modify*, which together reflect aspects of our *synthesize* mode.

Ellis and his colleagues [4, 5, 6] developed a model consisting of a number of broad information seeking behaviours, including *monitoring* and *verifying* ("checking the information and sources found for accuracy and errors"). In addition, his *browsing* mode ("semi-directed searching in an area of potential interest") aligns with our definition of *explore*. He also noted that it is possible to display more than one behaviour at any given time. In revisiting Ellis's findings among social scientists, Meho and Tibbo [10] identified *analysing* (although they did not elaborate on it in detail). More recently, Makri et al [8] proposed *searching* ("formulating a query in order to locate information"), which reflects to our own definition of *locate*.

In addition to the research-oriented models outlined above, we should also consider practitioner-oriented views. Spencer [14] suggests four modes of information seeking, including *known*-*item* (a subset of our *locate* mode) and *exploratory* (which mirrors our definition of explore). Lamantia [7] also identifies four modes, including *monitoring*.

In this paper, we use the characteristics of the models above as a lens to interpret the behaviours found in a new source of empirical site search data. We also explore the combinatorial nature of the modes, extending Ellis's [5] concept of mode co-occurrence to identify and define a set of repeating patterns and sequences.

3. CONSUMER SEARCH BEHAVIOUR

3.1 Data Acquisition

The primary source of data in this study is a set of 277 information needs captured during client engagements involving the development of a number of custom site search applications. These information needs take the form of 'micro-scenarios', i.e. a brief narrative that illustrates the end user's goal and the primary task or action they take to achieve it, for example:

- Find best offers before the others do so I can have a high margin.
- Get help and guidance on how to sell my car safely so that I can achieve a good price.
- Understand what is selling by area/region so I can source the correct stock.
- See year-on-year ad spend trends for TV and online to supply to the Head of Global Media.

The scenarios were collected as part of a series of requirements workshops involving stakeholders and customer-facing staff from the respective client organisations. They were generated by participants in individual breakout sessions, and then moderated by the workshop facilitator in a group session to maximise consistency and minimise redundancy or ambiguity. They were also prioritised by the group to identify those that represented the highest value both to the end user and to the client organisation.

This data possesses a number of unique properties. In previous studies of information seeking behaviour (e.g. [5], [10]), the primary source of data has traditionally been interview transcripts that provide an indirect, verbal account of end user *information behaviours*. By contrast, the current data source represents a self-reported account of *information needs*, generated directly by end users (although a proportion were captured via proxy, e.g. through customer facing staff speaking on behalf of the end users). This change of perspective means that instead of using information behaviours to infer information needs and design insights, we can adopt the converse approach and use the stated needs to infer information behaviours and the interactions required to support them.

Moreover, the scope and focus of these scenarios represents a further point of differentiation. In previous studies, (e.g. [8]), measures have been taken to address the limitations of using interview data by combining it with direct observation of information seeking behaviour in naturalistic settings. However, the behaviours that this approach reveals are bounded by the functionality currently supported by existing systems and working practices, and as such do not reflect the full range of *aspirational* or *unmet* user needs encompassed by the scenarios in this study.

Finally, the data is unique in that is constitutes a genuine practitioner-oriented deliverable, generated expressly for the purpose of designing and delivering professional site search systems. As such, it reflects a degree of realism that interview data or other research-based interventions might struggle to replicate.

3.2 Data Analysis

These scenarios were analyzed using the model derived previously for the domain of enterprise search [13]. In this respect, the process was partially deductive, applying the model in a top-down fashion to classify the data. But it was also partially inductive, applying a bottom-up, grounded analysis to identify new types of behaviour not present in the original model or to suggest revised definitions of the existing categories.

Although the original study involved three separate analysts, the behaviours this time were identified by the first author alone. The current analysis approach is therefore much more subjective. However, the first author was also the facilitator at each of the requirements workshops at which the scenarios were generated, and was able to again a deep insight into the needs, goals and motivations of the participants. This allowed him to be as confident as possible in his understanding of the users' information needs and consistent in his interpretation of the information behaviours required to satisfy a particular need.

A number of the scenarios focused on needs that did not involve any explicit information seeking or use behaviour, e.g. "Achieve a good price for my current car". These were excluded from the analysis. A further number were incomplete or ambiguous, or were essentially feature requests (e.g. "Have flexible navigation within the page"), and were also excluded. This process resulted in further confirmation and validation of the nine search modes identified in the original study, but with revised definitions to reflect a broader scope:

1. **Locate**: *To find a specific (possibly known) item*, e.g. "Find my reading list items quickly". This mode encapsulates the stereotypical 'findability' task that is so commonly associated with site search, consistent with (but a superset of) Spencer's [14] *known item* search mode. This was the most frequent mode in the site search scenarios (120 instances).

2. Verify: To confirm that an item meets some specific, objective criterion, e.g. "See the correct price for singles and deals". Often found in combination with locating, this mode is concerned with validating the accuracy of some data item, comparable to that proposed by Ellis et al. [5] (39 instances).

3. Monitor: Maintain awareness of the status of an item for purposes of management or control, e.g. "Alert me to new resources in my area". This activity focuses on the state of asynchronous responsiveness and is consistent with that of Bates [1], O'Day and Jeffries [11], Ellis [4], and Lamantia [7] (13 instances).

4. **Compare**: To identify similarities & differences within a set of *items*, e.g. "Compare cars that are my possible candidates in detail". This mode has not featured prominently in previous models (with the possible exception of Marchionini's), but was found to be a significant component of enterprise search behaviour [13]. Moreover, it is a common feature of product search and navigation on many ecommerce sites. However, it occurred relatively infrequently in the site search scenarios (2 instances).

5. **Comprehend**: *To generate independent insight by interpreting patterns within a data set*, e.g. "Understand what my competitors are selling". Like *compare*, this mode was found to be a key element of the enterprise search scenarios, and also features in the models of Cool & Belkin [3] and Marchionini [9]. It occurred relatively frequently in site search (50 instances).

6. **Explore**: To investigate an item or data set for the purpose of knowledge discovery, e.g. "Find useful stuff on my subject topic". In some ways the boundaries of this mode are somewhat less prescribed than the others, but what the instances share is the characteristic of open ended, opportunistic search and browsing in the spirit of O'Day and Jeffries [11] exploring a topic in an undirected fashion and Spencer's [14] exploratory. This mode was the second most common in site search (110 instances).

7. Analyze: To examine an item or data set to identify patterns & relationships, e.g. Analyze the market so I know where my strengths and weaknesses are". This mode features less prominently in previous models, appearing as a sub-component of the processing stage in Meho & Tibbo's [10] model, and overlapping somewhat with Cool & Belkin's [3] organize. This definition is also consistent with that of Makri et al. [8], who identified analysing as an important aspect of lawyers' interactive information behaviour and defined it as "examining in detail the elements or structure of the content found during information-seeking." (p. 630). Although the most common element of the enterprise search scenarios, it was less prevalent in site search (59 instances).

8. Evaluate: To use judgement to determine the value of an item with respect to a specific goal, e.g. "I want to know whether my agency is delivering best value". This mode is similar in spirit to verify, in that it is concerned with validation of the data. However, while verify focuses on simple, objective fact checking, our conception of evaluate involves more subjective, knowledge-based judgement, similar to that proposed by Cool & Belkin [3] (61 instances).

9. Synthesize: To create a novel or composite artefact from diverse inputs, e.g. "I need to create a reading list on celebrity sponsorship". This mode also appears as a sub-component of the *processing* stage in Meho & Tibbo's [10] model, and involves elements of Cool & Belkin's [3] *create* and *use*. Of all the modes, this one is the most commonly associated with information *use* in its broadest sense (as opposed to information *seeking*). It was relatively rare within site search (5 instances).

4. MODE SEQUENCES AND PATTERNS

Applying the modes described above provides a framework for understanding the needs of site search users, and an insight into their likely behaviours. But as with the previous study [13], their real value lies not so much in the individual instance data but in the patterns of co-occurrence they reveals. In most scenarios, modes combine to form distinct chains and patterns, echoing the transitions observed by O'Day and Jeffries [11] and the combinatorial behaviour alluded to by Ellis [5], who suggested that information behaviours can often be nested or displayed in parallel.

Just as new definitions were needed to accommodate the new domain, new patterns of occurrence were identified in the data. Typically these consisted of chains of length two or three, of which the following were most frequent:

- 1. **Insight-driven search**: (Explore->Analyze-> Comprehend): This patterns represents an exploratory search for insight to resolve an explicit information need, e.g. "Assess the proper market value for my car" (45 instances)
- 2. **Opportunity-driven search**: (Explore-Locate-Evaluate): In contrast to the explicit focus of the pattern above, this sequence represents a less directed exploration in the prospect of serendipitous discovery e.g. *"Find useful stuff on my subject topic"*(31 instances)
- 3. **Qualified search** (Locate-Verify) This pattern represents a variant of the stereotypical findability task in which some element of immediate verification is required, e.g. "*Find trucks that I am eligible to drive*" (29 instances)

A deeper insight into these patterns can be obtained by presenting them in diagrammatic form, as a network (Figure 1). This diagram illustrates the three sequences outlined above plus other commonly found patterns. It also reflects an outcome of the pervious study, in that certain modes tend to function as "terminal" nodes, i.e. entry points or exit points to a scenario. For example, *Explore* typically functions as an opening, while *Comprehend* and *Evaluate* function in closing a scenario. *Analyze* typically appears as a bridge between an opening and closing mode.



Figure 1. Mode network for site search

4.1 Site search vs. Enterprise Search

The sequences described above also allow us to reflect on some of the differences between the needs of site search users and those of enterprise search. One of the most fundamental differences is an emphasis on simpler "lookup" modes such as *Locate* and *Verify*: these were relatively rare in the enterprise search data, but prominent in site search (120 and 39 instances respectively). Enterprise search, by contrast, emphasised higher-level "investigate" behaviours such as *Analyze* and *Evaluate* (modes which also appeared frequently in site search, but not as prominently: 58 and 61 instances respectively). However, in neither case was the stereotype of 'search as findability' borne out: even in site search (where it was the most common mode), *Locate* was accountable for no more than a quarter of all instances.

But perhaps the biggest difference was in the composition of the chains: while enterprise search was characterised by a wide variety of heterogeneous chains, site searched focused on a small number of common trigrams and bigrams. Moreover, these chains displayed little evidence of the composite nature observed in enterprise search, in which certain chains were seen to be embedded within others to create larger, more complex sequences of behaviour.

5. DISCUSSION

A key feature of the current model is its emphasis on the combinatorial nature of search modes, and the value this offers as a framework for expressing complex patterns of behaviour. Such an approach is not unique: the second author, for example, has also previously explored the concept of mode chains to describe information seeking behaviours observed in naturalistic settings. However, his approach was based on the analysis of complex tasks observed in real time, and as such was less effective in revealing consistent patterns of atomic behaviour such as those found in the current study.

Conversely, this virtue can also be a shortcoming: the fact that simple repeating patterns can be extracted from the data may be as much an artefact of the medium as it is of the information needs it contains. These scenarios were expressly designed to be a concise, self-contained deliverable in their own right, and applied as a simple but effective tool in the planning and prioritisation of software development activities. This places a limit on the length and sophistication of the information needs they encapsulate, and hence a natural boundary on the scope and extent of the patterns they represent. Their format also allows the analyst to apply perhaps an unrealistic degree of top-down judgement and iteration in aligning the relative granularity of the information needs to existing modes; a benefit that is less readily available to those whose approach involves real-time, observational data.

A further caveat is that in order to progress from understanding an information need to identifying the information behaviors required to satisfy those needs, it is necessary to *speculate* on the behaviours that a user *might* perform when undertaking a task to satisfy the need. It may transpire that users actually perform different behaviours which achieve the same end, or perform the expected behavior but through a combination of other nested behaviours, or may simply satisfy the need in a way that had not been envisaged at all.

Finally, the process of inferring information behaviour from selfreported needs can never be wholly deterministic, regardless of the consistency measures discussed earlier. In this respect, further steps should be taken to operationalize the application of the framework and apply some independent measure of stability or objectivity in its usage.

6. CONCLUSIONS

In this study we have investigated a model of information seeking behaviour derived from the domain of enterprise search, and validated its extensibility to users of consumer-oriented websites and search applications. In so doing, we explored a novel, goaldriven approach to eliciting user needs, and identified some key differences in user behaviour between the two domains.

In addition, we have demonstrated the value of the model as a framework for expressing complex patterns of behaviour, extending the IR concept of information-seeking to embrace a broader range of composite information interaction and use behaviours. Moreover, we propose that our method can be adopted by other researchers who want to take a 'needs first' approach to understanding information behaviour.

7. REFERENCES

- [1] Bates, Marcia J. 1979. Information Search Tactics. *Journal* of the American Society for Information Science 30: 205-214
- [2] Broder, A. 2002. A taxonomy of web search, *ACM SIGIR Forum*, v.36 n.2, Fall 2002
- [3] Cool, C. & Belkin, N. 2002. A classification of interactions with information. In H. Bruce (Ed.), *Emerging Frameworks* and Methods: CoLIS4: proceedings of the 4th International Conference on Conceptions of Library and Information Science, Seattle, WA, USA, July 21-25, 2002, (pp. 1-15).
- [4] Ellis, D. 1989. A Behavioural Approach to Information Retrieval System Design. *Journal of Documentation*, 45(3).
- [5] Ellis, D., Cox, D. & Hall, K. 1993. A Comparison of the Information-seeking Patterns of Researchers in the Physical and Social Sciences. *Journal of Documentation* 49(4).
- [6] Ellis, D. & Haugan, M. 1997. Modelling the Informationseeking Patterns of Engineers and Research Scientists in an Industrial Environment. *Journal of Documentation* 53(4), pp. 384-403.
- [7] Lamantia, J. 2006. 10 Information Retrieval Patterns JoeLamantia.com, http://www.joelamantia.com/informationarchitecture/10-information-retrieval-patterns
- [8] Makri, S., Blandford, A. & Cox, A.L. 2008. Investigating the Information-Seeking Behaviour of Academic Lawyers: From Ellis's Model to Design. *Information Processing and Management* 44(2), pp. 613-634.
- [9] Marchionini, G. 2006. Exploratory search: from finding to understanding. *Commun. ACM* 49(4): 41-46
- [10] Meho, L. & Tibbo, H. 2003. Modeling the Informationseeking Behavior of Social Scientists: Ellis's Study Revisited. *Journal of the American Society for Information Science and Technology* 54(6), pp. 570-587.
- [11] O'Day, V. and Jeffries, R. 1993. Orienteering in an information landscape: how information seekers get from here to there. *INTERCHI* 1993: 438-445
- [12] Rose, D. and Levinson, D. 2004. Understanding user goals in web search, *Proceedings of the 13th international* conference on World Wide Web, New York, NY, USA
- [13] Russell-Rose, T., Lamantia, J. and Burrell, M. 2011. A Taxonomy of Enterprise Search and Discovery. *Proceedings* of HCIR 2011, California, USA.
- [14] Spencer, D. 2006. Four Modes of Seeking Information and How to Design for Them. Boxes & Arrows: http://www.boxesandarrows.com/view/four_modes_of_seeking_information_and_how_to_design_for_them.