The Information-Seeking Funnel

Daniel E. Rose
A9.com
130 Lytton Ave., Suite 300
Palo Alto, CA 94301 USA
+1 650-331-2670
danrose@a9.com

ABSTRACT

We introduce a new conceptual framework for information seeking called the information-seeking funnel, inspired by a common model of how people are attracted to products they end up buying. By using this model, we are able to better understand why current search tools are inadequate for several stages of the information-seeking process. At the same time, we explore what characteristics tools that are designed for these neglected stages might have, and how we might evaluate them.

1. INTRODUCTION

Information-seeking is a complex activity that can have varying degrees of directedness. Some parts of this activity involve focused search; others are more exploratory. But what does it mean to be exploratory? Is exploratory search an attitude or a process? What is the relationship between different kinds of information-seeking behaviors? We investigate these questions by using an analogy drawn on another type of familiar activity: shopping. While the task domains are quite different, they share the property that successive stages have different characteristics, and that different tools are appropriate at each stage.

2. SHOPPING FOR IDEAS

In the commercial world, there is a model called the *buying funnel* or the *sales funnel*. The buying funnel depicts the changing sets of people at different stages of the buying process, from all those who might possibly be interested in a product or service to those who actually purchase it. Sometimes the sets are described from the perspective of a sales representative, as "suspects," "prospects," "qualified leads," and "customers." Other times, the funnel is characterized by the mindset of the potential customer, where the four stages are "Awareness," "Desire," "Interest," and "Action," or (similarly) "Awareness," "Research," "Decision," and "Purchase." The "funnel" is so called because the graphical depiction illustrates the successively smaller sets of people that occupy each stage (see Figure 1). At each stage, the goal of the seller changes to correspond to the mindset of the buyer. For example, at the first stage, the seller's goal is to make the potential buyer aware of the product.

Prepared as a position paper for National Science Foundation workshop on Information-Seeking Support Systems (ISSS), June 26-27, 2008, Chapel Hill, NC. Workshop organizers: Gary Marchionini and Ryen White.

The buying funnel model seems to be widely used in business. It's mentioned everywhere from books on how to be a successful salesperson to blogs for search engine optimizers discussing how to get more clicks on paid search advertising. But what's striking is that the funnel embraces all phases of the user's mindset starting *before* she is even aware of the product or service.

2.1 The Information-Seeking Funnel

Now imagine that instead of potential customers buying products, we have users seeking information. By analogy, we can model the situation using an *information-seeking funnel*. And as with the buying funnel, we'll start before users are aware of an information need, and end when they know it has been satisfied. Just as we can think of a product's manufacturer hoping to draw a person deeper into the funnel, we can think of the supplier of some information (an author, perhaps) wanting that information to be consumed. Alternatively, we can think of the information nuggets themselves as memes [4] that "want" to be disseminated.

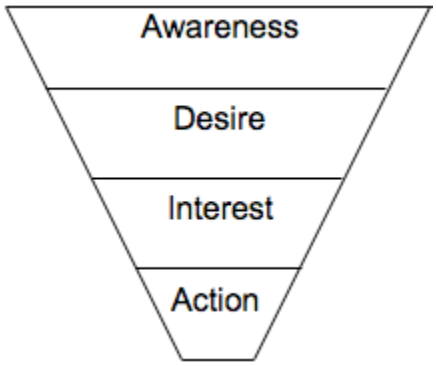

Figure 1: The buying funnel.

Of course, there have been many excellent models of information-seeking (e.g. ASK [3]) developed over the years as a result of studying how users search. But these models tend to focus on the situation after the user knows what information she is seeking. Some, like the Berrypicking model [1], acknowledge that the user's information need starts out poorly formed and evolves. But in general, there is an assumption of a certain amount of directedness toward an information goal.

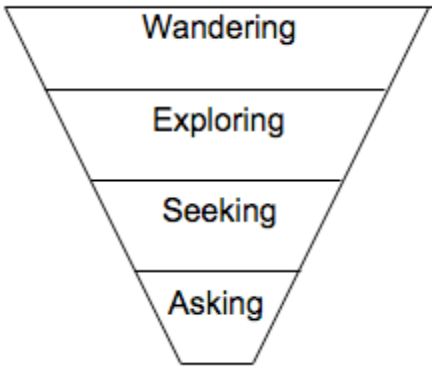

Figure 2: The information-seeking funnel.

The information-seeking funnel, shown in Figure 2, consists of four stages:

- *Wandering*. At this stage, the user does not have an information seeking-goal in mind. However, she may have a meta-goal (e.g. "find a topic for my final paper.")
- *Exploring*. In the Exploring stage, the user has a general goal (e.g. "learn about the history of communication technology") but not a plan for how to achieve it.
- *Seeking*. Here, the user has started to identify information needs that must be satisfied (e.g. "find out about the role of the telegraph in communication."), but the needs are open-ended.
- *Asking*. At the final stage, the user has a very specific information need that corresponds to a closed-class question ("when was the telegraph invented?").

The funnel shape now represents the narrowed scope of relevant information. At the first stage, all information in the world is potentially relevant; at the last stage, only a small amount of information is relevant.

3. INTERACTION AT EVERY STAGE

It is fair to say that most existing work on technology that supports information-seeking has focused on the last two stages, where the information need is more narrowly focused and the information-seeking task more directed.

If that's true, then opportunities exist to make a difference in neglected stages of the process. What might tools that are aimed at these other parts of the funnel look like? To investigate this, it is instructive to return to the shopping experience and think about the many types of interaction that are supported throughout the buying funnel.

3.1 Variety of Buying Funnel Interactions

In the early stages, users are interacting with the entire world. In order to gain their awareness, marketers rely on brand advertising in a variety of media. When a new class of products is introduced, the seller often has to convey some information about why someone would find such a product useful. Essentially, part of the task is education – initially, people may not even know what a phonograph / clothes dryer / cell phone is or why they would want one.

In the middle stages, users – now potential customers – are starting to seek out information about products. They may go to a store and look at the product or try it out; they may read reviews in online or offline media; they may rely on recommendations, either from friends or strangers, as in collaborative filtering; they may look at ratings from others who have bought the product.

Importantly, users at these stages employ a variety of modalities. For example, online purchases often involve complex combinations of searching, navigating, filtering through a faceted metadata space, and simply following suggestions that attract their attention.

As users narrow in on specific products, merchants provide specialized tools to facilitate that process. For example, articles of clothing sold online may be viewed from several different angles or with the option to zoom in on details. The task changes from learning about these kinds of products to learning about this specific product. What's important is not only that there are such a wide variety of tools, but also that the tools are appropriate to the funnel stage.

3.2 Stateful Search

Returning to the information-seeking funnel, what might some tools to support the upper stages look like? It is hard to predict which of the many new approaches being explored in the research community are likely to be successful, but it seems clear that at least one aspect of these systems will be what we call *stateful search*.

In recent years, most general-purpose search engines tend to behave in a stateless way. That is, each interaction with a user is performed in isolation, without taking into account any previous interactions with users.[1]

In contrast, a stateful search system treats the process, not just the result, as a first-class object of interest. A stateful search system will not only keep a history of each step in the information-seeking process (searching, browsing, viewing results, etc.), it will also allow users to interact with that state. Furthermore, such a system will allow users to gather, analyze, annotate, and organize information, creating new explicit representations of state alongside those kept internally by the system.

But doing all this is going to require different kinds of user interactions than search engines expect today – interactions that in many cases might be characterized as "more work" than they expend when entering a query today. Is it reasonable to expect that users can change their behavior?

3.3 The Myth of the Two-Word Query

In 1996, the ATG Information Access Group at Apple packaged the V-Twin search software we had developed[2] into a free web site search engine called "Apple e.g.," which powered an early

[1] In a sense, "personalized" search systems, which analyze the past actions of individual users to learn their general preferences, may be considered stateful, but not at the level of granularity we mean here.

[2] Doug Cutting was the architect and lead developer of V-Twin.

version of Apple's own web site as well as several other university sites. Observing that user queries on these sites were quite short, we came up with a method for improving performance of the engine on short queries. We presented our results at TREC [8], where we simulated the behavior of short queries on the TREC ad hoc task by truncating the topic titles to the two words with the highest term weight. At the time, we had trouble convincing people that worrying about two-word queries was worthwhile. I overheard another participant at the conference questioning the value of our work during one of the breaks. "If we're only given two words to work with," our critic said, "then there's pretty much nothing we [IR researchers] can do."

A few years later, the pendulum had swung absurdly far in the opposite direction, where it remains today. A series of web use studies starting in 1998 [10] confirmed that short queries are the norm on the web. Not only are all major search engines optimized for two-word queries, there is now conventional wisdom that says that users will never type more than this. This is sometimes characterized as "users are lazy," and other times as "users are efficient." In some ways, the latter sentiment is worse, implying that search engines are so good that no more than two words are necessary for describing a user's information need.

As we have discussed above and elsewhere [7, 9], the existing search engine interaction paradigm of typing two words in a text box and getting a linear list of ten results is ill-suited to many information-seeking tasks. In particular, supporting all stages of the information-seeking funnel will require a variety of interaction modalities, just as is the case with the buying funnel. But are users so strongly trained in the existing paradigm, and so reluctant to enter more than two words, that tools with alternate interaction styles are doomed to failure?

Several pieces of evidence suggest that they are not. We know from multiple studies [2, 5] that the user interface – even the shape of the search box – can influence the length of user queries. Perhaps more importantly, there are thriving communities where users routinely expend much more effort and enter far more information. At Yahoo! Answers [11], a service where users ask questions and others in the community answer them, the "short" version of the question is typically 10 or 12 words – and this is usually followed by a more detailed version that can be as long as a paragraph. A few recent short questions are:

- *Does Barack Obama play video games?*
- *Where can I find a good mariachi in LA county or OC for my wedding?*
- *What documents are required to bring the child of my fiancée to the USA on a K1 visa?*

Just part of one typical detailed question starts:

> *I am 12 years old and i need a dress for my friends bar mitzvah. I want it to be either red, gray, or black and not too expensive. I got some really cute black wedges but i dont have a dress but i could exchange the shoes. I need it by the 21st so it has to be a store i can go buy it at…*

What these examples suggest is that users are not only adaptable but also quite willing to switch to a different interaction style if they believe that style will provide more value in that system. If the application is compelling enough, users will happily take the time to provide the information that will make them successful.

4. UPPER-FUNNEL EVALUATION

Once we have these new tools for supporting the upper states of the information-seeking funnel, how will we evaluate them? For specific information-seeking tasks at the lower half of the funnel, we can evaluate a system based on how well it helps users satisfy their information need. Metrics such as recall, precision, and discounted cumulated gain (DCG) can answer questions like "did the system find all the material on the topic" and "were the results presented in the optimal order." But those metrics are not useful when the information need is not yet clearly defined.

Some systems are also judged by the *speed* with which users can "get answers to their questions." I have heard many product managers in the search industry, and even some HCI professionals, claim that a good search engine is one in which users "get in and get out as quickly as possible." Not only does this perspective treat the *quality* of the information obtained as irrelevant, it is predicated on the assumption that there is no value to the user in exploration [6].

In short, metrics like these can't tell us whether a system did a good job helping users explore different ideas or prompting them to make connections.

4.1 Collateral Knowledge

At the top half of the funnel, then, what we need are ways to measure the ability of the system to convey unsolicited information – essentially, to put the user in an information-seeking frame of mind.

We began exploring this concept during research at Apple in the 1990s on the roles of browsing and searching in information-seeking systems.[3] We hypothesize that while a user is conducting an information-seeking task, he acquires insights and information along the way that he was not explicitly looking for, but that may prove to be valuable. This serendipitously-acquired information we call *collateral knowledge*.

To measure the ability to convey collateral knowledge, we designed the following methodology: First, we gave users a pre-test to measure their general knowledge of a certain subject area. Next, we gave them a set of directed information-seeking tasks using a particular system. Finally, we gave them a post-test on the subject, including topics they had not been asked to learn about, but which they might have been exposed to during the process. By comparing the post-test answers to the pre-test answers, we were able to measure the collateral knowledge obtained during the information seeking process.[4]

Since users who are at the early stages of the information funnel don't exactly know what they're looking for yet,

[3] This work was done with Paul V. Biron.

[4] The specific study we conducted at Apple was intended to compare the abilities of search-oriented vs. browse-oriented systems to convey collateral knowledge, but for a variety of logistical reasons, the study was never completed.

collateral knowledge may be an interesting measure of how well the system is supporting exploratory search.

5. CONCLUSIONS

The information-seeking funnel is a model of the information-seeking process encompassing the period from before the user knows that an information need exists, to the point at which an information need is satisfied. By means of its analogy to the buying funnel, it serves as a way to frame the problem of information-seeking tools and their evaluation. It helps explain why mainstream search engines aren't doing a good job of supporting exploratory search. It makes clear that – just as with shopping – different tools with different interaction styles will be needed at each stage. It illustrates why metrics for the latter stages make little sense for the earlier ones. In summary, the information-seeking funnel provides a conceptual framework for thinking about the problem space.

6. REFERENCES

[1] Bates, M. J. 1989. The Design of Browsing and Berrypicking Techniques for the Online Search Interface. *Online Review*, 13(5):407-24.

[2] Belkin, N. J., et al. 2001. Rutgers' TREC 2001 Interactive Track Experience. *Proceedings of the Tenth Text Retrieval Conference* (TREC-10).

[3] Belkin, N. J., Oddy, R. and Brooks, H. 1982. ASK for Information Retrieval. *Journal of Documentation*, 38, 61-71 (part 1) & 145-164 (part 2).

[4] Dawkins, R. 1976. Memes: The New Replicators, in *The Selfish Gene*. Oxford University Press.

[5] Karlgren, J. and Franzen, K. 2000. Verbosity and Interface Design. *SICS Technical Report T00:04*, Swedish Institute for Computer Science, Stockholm.

[6] Rose, D. E. 2004. Why Is Web Search So Hard... to Evaluate? *J. Web Eng.* 3(3-4): 171-181.

[7] Rose, D. E. 2006. Reconciling Information-Seeking Behavior with Search User Interfaces for the Web. *JASIST* 57(6):797-799.

[8] Rose, D. E. and Stevens, C. 1997. V-Twin: A Lightweight Engine for Interactive Use, *Proceedings of the Fifth Text Retrieval Conference* (TREC-5).

[9] Rose, D. E. and Raju, S. 2007. Encouraging Exploration with Elroy: A New User Experience for Web Search. *Proceedings of the ACM SIGCHI 2007 Workshop on Exploratory Search and HCI* (ESI 2007), 52-55.

[10] Spink, A., Bateman, J., and Jansen, B. J. 1998. Users' Searching Behavior On The Excite Web Search Engine. *WebNet 1998.*

[11] Yahoo! Answers, http://answers.yahoo.com, accessed 12 June 2008.