

A Taxonomic Analysis of What World Wide Web Activities Significantly Impact People's Decisions and Actions

Julie B. Morrison, Peter Pirolli, Stuart K. Card

Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, California 94304
+1 650 812-4000
{morrison,pirolli,card}@parc.xerox.com

ABSTRACT

In this paper, we present three taxonomic classification schemes based on Web users' responses to what Web activities significantly impacted their decisions and actions. The taxonomic classifications focus on three variables: the *Purpose* of people's search on the Web, the *Method* people use to find information, and the *Content* of the information for which they are searching. These taxonomies are useful for understanding people's activity on the Web and for developing ecologically-valid tasks to be used when studying Web behavior.

Keywords

Information foraging, web usability, taxonomy

INTRODUCTION

The Graphics, Visualization, and Usability Center at Georgia Institute of Technology has conducted ten on-line surveys assessing Internet demographics and usage patterns. From October through December 1998, the following question was posted on the *Web and Internet Use* subsection of the survey [3]: "Please try to recall a recent instance in which you found important information on the World Wide Web, information that led to a significant action or decision. Please describe that incident in enough detail so that we can visualize the situation." This question will be referred to throughout as the *Significance* question.

The *Significance* question was a slight modification of the interview question used by Choo, Detlor, and Turnbull [1] when assessing how managers and IT specialists use the Web. Flanagan's [2] Critical Incident Technique, which emphasizes the importance of the evaluated incidents being recent, complete, and with known consequences, provided the basis for the original development of Choo and colleagues' question.

The purpose of posting the question was to determine the types of on-line activities in which people engage that they believe impact their lives. The responses were analyzed to provide two types of information. First, we wanted to

develop one or more taxonomies that appropriately identified the types of Web activities people think are significant. Second, we wanted to use the analysis of the responses to develop a laboratory experiment in which people complete the same types of tasks as the survey respondents (see *Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability*, Card et al., this volume).

TAXONOMIC CLASSIFICATIONS

Of the 3292 respondents to the *Web and Internet Use* survey, 2188 provided usable responses to the *Significance* question. Our interest in evaluating the responses was to understand the three components of Web searching: *why* people searched the Web, *how* they searched the Web, and for *what* information they searched.

Development of the specific taxonomic classifications addressing the *why*, *how*, and *what* questions was based on an earlier model of Web information seeking [1] as well as on a Web use taxonomy created by Stuart Card. A series of revisions of these classification schemes yielded the three taxonomies presented here. For example, Card's system was based on what people did on the World Wide Web as compared with what people did on the Web that they considered *significant*. When asked in a separate *Web and Internet Use* survey question how they spend their time on-line, 24% of the respondents said they read news, while only 2% mentioned reading news in response to the *Significance* question. After two iterations of revising the taxonomic classifications following the coding of random sets of 100 responses, we coded a final set of 100 responses. The data from this third coding is reported here.

The taxonomic classifications are a formalization of our original *why*, *how*, and *what* questions focusing on the *Purpose* of people's search on the Web, the *Method* people use to find information, and the *Content* of the information for which they are searching. The full taxonomies with descriptions of each component and accompanying graphs showing the distribution of responses (Figures 1-3) are on the following page. Those responses for which the coding disagreed, and those which were uncodable because the respondent failed to provide enough information, are not represented in the graphs. For the *Purpose*, *Method*, and *Content* taxonomies, 11%, 8%, and 6% of the responses were uncodable, and 33%, 38%, and 23% were disagreed upon, respectively. In general, inter-rater reliabilities were

© Copyright on this material is held by the Author(s).

acceptably high: Purpose Taxonomy $\alpha=.78$; Method Taxonomy $\alpha=.67$; and Content Taxonomy $\alpha=.82$.

Purpose Taxonomy

What was the primary reason for the respondent's search?

1. Find: Use of the web to
 - a. Download information
 - b. Get a fact
 - c. Get a document
 - d. Find out about a product
2. Compare/Choose: Use of the web to evaluate multiple products or pieces of information in order to help the respondent make a decision.
3. Understand: Use of the Web to help the respondent to understand some topic; generally includes locating facts or documents.

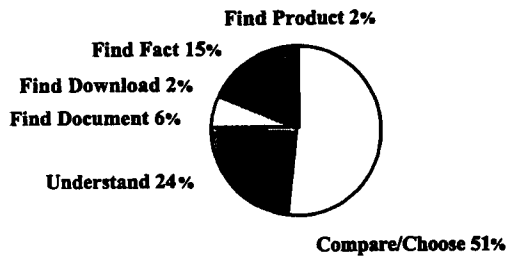


Figure 1. Response Distribution for the Purpose Taxonomy

Method Taxonomy

How did the respondent find the information? What was the respondent's goal?

1. Explore: General searching for information. The search is not triggered by a particular goal.
2. Monitor: Repeated visits to specific websites to update information. The search is not triggered by a particular goal; it is a routine behavior.
3. Find: Searching for a particular fact/document/piece of information. Search is triggered by a goal.
4. Collect: Searching for multiple pieces of information. Searcher is open to any answer, not looking for a particular one. A goal drives the searcher's behavior.

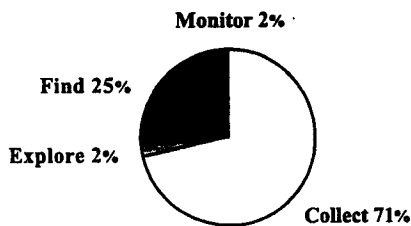


Figure 2. Response Distribution for the Method Taxonomy

Content Taxonomy

For what type of information is the responded searching?

- | | |
|------------------|----------------------------------|
| 1. Business | 8. People |
| 2. Education | 9. Product Info & Purchase (I/P) |
| 3. Finance | a. Computer |
| 4. Job Search | b. Vehicles |
| 5. Medical | c. Download |
| 6. Miscellaneous | d. Other |
| 7. News | 10. Travel |

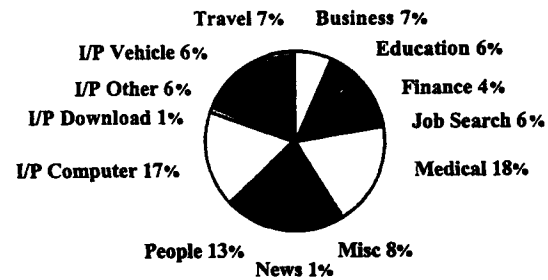


Figure 3. Response Distribution for the Content Taxonomy

CONCLUSIONS

In order to study search behavior on the Web, it is essential to define a set of tasks that replicate real user behavior. Prior to this study, no task bank existed, nor was there a system for identifying and developing such tasks. The three taxonomic classification systems we developed identify the activities Web users engage in when using the Web to make decisions or to inform actions. By separating search behavior in terms of the *Purpose* of the search, the *Method* used to find the information, and the *Content* of the searched-for information, we have been able to create ecologically-valid tasks used to guide our understanding and modeling of search behavior on the Web.

ACKNOWLEDGMENTS

We would like to thank the Georgia Tech Gvu for collecting data on the Significance question as part of their WWW User Survey. This work is supported by an Office of Naval Research grant No. N00014-96-C-0097.

REFERENCES

1. Choo, C.W., Detlor, B., & Turnbull, D. (1998). A behavioral model of information seeking on the web: Preliminary results of a study of how managers and IT specialists use the web. *Proceedings of the 61st Annual Meeting of the American Society for Information Science* (pp. 290-302), Pittsburgh, PA.
2. Flanagan, J.C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327-358.
3. Georgia Tech Graphics, Visualization, and Usability Center. (1998). *Gvu's 10th WWW User Survey*, http://www.gvu.gatech.edu/gvu/user_surveys/survey-1998-10/.