

# The Impact of Automated Assistance on the Information Retrieval Process

**Bernard J. Jansen**

School of Information Sciences and Technology  
The Pennsylvania State University  
University Park, PA, 16801, USA  
[jjansen@acm.org](mailto:jjansen@acm.org)

**George K. Kroner**

School of Information Sciences and Technology  
The Pennsylvania State University  
University Park, PA, 16801, USA  
[gik103@psu.edu](mailto:gik103@psu.edu)

## ABSTRACT

Advanced information retrieval systems providing automated assistance offer the opportunity to greatly enhance the effectiveness of the information retrieval process. One issue in designing such systems is determining the effect that the automated assistance has on the tasks and sequence of tasks within this process. Using verbal protocol data and transaction log analysis, we present a taxonomy of tasks when utilizing information retrieval systems with automated assistance, along with a temporal analysis of when interaction with the automated assistance occurs. Results indicate that there is a predictable pattern of user interaction with automated assistance with implications for the design of information retrieval systems.

## Keywords

Information retrieval, user interaction, explanations

## INTRODUCTION

There has been considerable research into automated information retrieval systems (e.g., [1]), in order to resolve some of the issues users have during the information retrieval (IR) process and when interacting with IR systems [2]. These automated IR systems attempt to assist the user during the search process by either executing search tactics for or offering searching advice or strategy to the user. Although the success of these systems depends in part on their ability to generate appropriate advice or action at appropriate times, there has been little empirical research examining to the impact of automated assistance on the information retrieval process. The research results presented in this article identify the tasks user initiate when using advanced IR systems. This knowledge can be used to improve the design of future IR systems. We discuss the experiment conducted, the results of our experiment and the implications for IR system design. Directions for future research are also presented.

## EMPRICAL TESTS

Our research goal was to identify the individual tasks when using an automated system during the information retrieval process. The IR system utilized for the empirical study was MG [3] enhanced with an agent that provided searching assistance [4]. The interface was a web browser coded in TCL/TK. The agent was programmed to offer assistance whenever it has assistance to offer.

The document collection utilized was the Text REtrieval Conference (TREC) volumes number 4 and 5, approximately 2GB in size and containing approximately 550,000 documents. The TREC topics utilized for this evaluation were Number 301: International Organized Crime and Number 340: Land Mine Ban.

The subjects for the evaluation were 30 freshman college students (26 males and 4 females) in their second semester at a four-year undergraduate university. The subjects were provided 5 minutes to find as many relevant documents as possible and instructed to verbalize their actions during the search process. We determined the length of the search session by measuring the length of time it would take to implement a typical Web search session, as reported in [2].

The users were video taped during the searching process and a transaction log recorded user – system interactions. Using the video, we analyzed the user's utterances using the thinking-aloud protocol where the verbalization occurs in conjunction with a task. The combination of the protocol analysis, transactions log, and subject evaluations provided a robust data source to conduct our analysis. All 30 subjects utilized the full five minutes for a total of 150 minutes of video for analysis.

The coded protocols from the video analysis identified the specific user actions on the system. These interactions represent tasks occurring during the search process. From this task analysis, we developed a taxonomy of user – system interactions, as shown in Table 1. The 7 categories account for 99.66% (501) of all user interactions with the system. There were 2 user actions occurring a single time that this taxonomy does not describe, which were one action to resize the browser and one action to access the normal help features of the system.

The search process unfolds sequentially over time as a searcher enters a query, evaluates the results list or individual results. If the searcher does not locate documents or enough documents to satisfy their

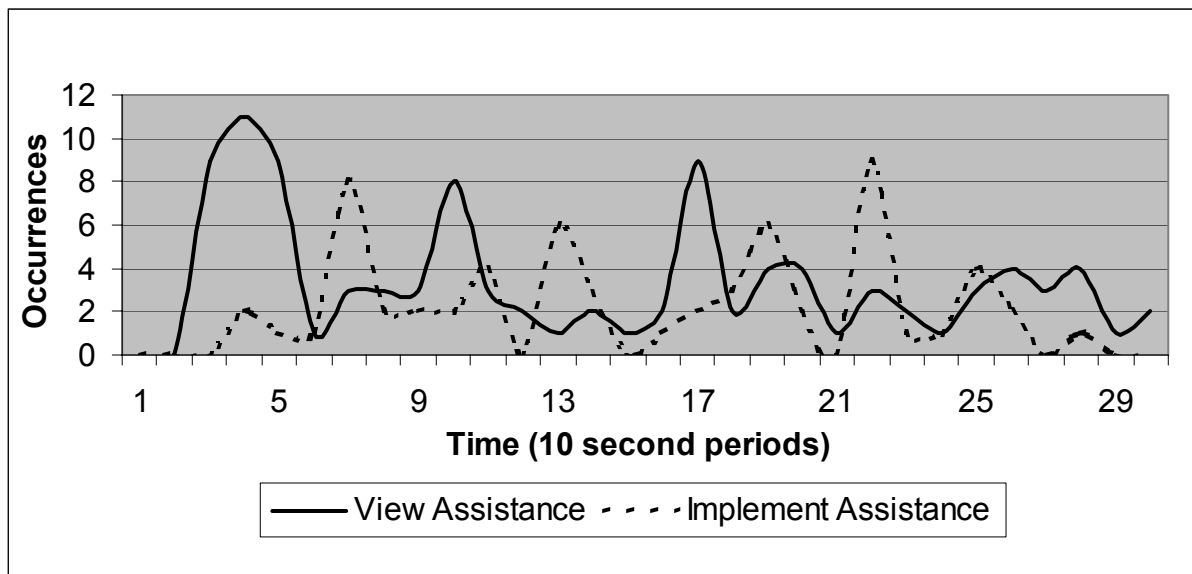
information need, they may mentally reformulate the query, repeating the process [5]. Therefore, we present a temporal view of when in the search process searchers

interacted with the automated assistance, as shown in Figure 1.

**Table 1: Taxonomy of User – System Interactions.**

	Taxonomy Category	Number of Occurrences	Percentage of All Occurrences		Taxonomy Category	Number of Occurrences	Percentage of All Occurrences
1	View Results Listing	122	24.3%	5	Implement Assistance	63	12.5%
2	View Particular Document	105	20.9%	6	Navigation	57	11.3%
3	View Offered Assistance	76	15.1%	7	Action Indicating Relevance	10	2.0%
4	Search	68	13.5%		<b>Total Actions</b>	501	99.60%

**Figure 1: Temporal View of Interaction with Automated Assistance.**



From Figure 1, there are peaks and valleys of viewing the assistance and implementing the assistance, with a noticeable lag between the two tasks. There was a reduction in the viewing of the automated assistance after the initial two periods representing a drop of approximately 4% to 15% from the first and second periods. The implementation held relatively steady until the last period.

**CONCLUSION AND SIGNIFICANCE OF RESEARCH**

From our analysis of user interactions with automated searching assistance, the searchers interacted with the automated searching assistance 54% of the time (76 of the 141 offers). The temporal examination of the data shows that the searchers utilized the automated assistance more in the beginning of the search process compared to later in the search process. Users most commonly implemented the assistance immediately after viewing the feedback. However, rather surprisingly, users also took some other action 22% of the time between viewing and implementing assistance. This would indicate that users, even after viewing assistance, prefer to attempt to search on their own first. By detecting the patterns of user – system interaction, designers can tailor IR systems to provide searching assistance at opportune states when the probability is greatest that the user is willing to view the assistance and

implement the assistance. This may improve the benefit these systems have on the searching process.

**REFERENCES**

- [1] S. Lawrence, C. L. Giles, and K. Bollacker, "Digital Libraries and Autonomous Citation Indexing," *IEEE Computer*, vol. 32, pp. 67-71, 1999.
- [2] B. J. Jansen, A. Spink, and T. Saracevic, "Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web," *Information Processing and Management*, vol. 36, pp. 207-227, 2000.
- [3] I. Witten, A. Moffat, and T. C. Bell, *Managing Gigabytes: Compressing and Indexing Documents and Images*. New York, NY: Van Nostrand Reinhold, 1994.
- [4] B. J. Jansen, "A Software Agent for Performance Improvement of Existing Information Retrieval Systems," presented at The International ACM Conference on Intelligent User Interfaces, Los Angeles, CA, 1999.
- [5] S. Robertson, "Theories and Models in Information Retrieval," *Journal of Documentation*, vol. 33, pp. 126-148, 1977.