Clinic on Library Applications of Data Processing: 1997

Visualizing Subject Access for 21st Century Information Resources

Edited by
Pauline Atherton Cochrane
Eric H. Johnson

with the editorial assistance
of Sandra Roe

Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
1998

ISBN 0-87845-103-X ISSN 0069-4789

James S. Dowling, Managing Editor
Susan Lafferty, Production Assistant
PUBLICATIONS COMMITTEE
Leigh Estabrook, Janice Del Negro, Monica M. Walk, David Dubin

CONTENTS

The Role of Subject Access in Information Filtering

Raya Fidel and Michael Crandall

ABSTRACT

The sheer amount of electronic information makes filtering a vital component of contemporary information work. Field observations of managers and engineers at the Boeing Company who received filtered information about computer-related topics revealed criteria they used to select, and those they used to reject, documents within their subject interest. Responses to a questionnaire indicated that some criteria are used more frequently and are more important than others. The few criteria that related to the subject matter of the documents were not limited to a subject domain. Other criteria addressed the form of the documents, their content, and writing style. In addition, some criteria were stable and somewhat objective and others were situational and subjective. An examination of these criteria shows that many of them could be used in filtering, in addition to subject-based mechanisms, and that they might be particularly useful for systems with multiple sources because they can provide a useful filter that is not based on the subject domain.

INTRODUCTION

Electronic delivery of information from multiple sources, often covering a range of subjects, is becoming the norm. The sheer amount of information available makes filtering a vital component of contemporary information work. The essential role of filtering mechanisms has already been recognized by systems designers and developers who have proposed various algorithms and interface agents for information filtering (e.g., Anick et al., 1991; Maes, 1994; Shuldberg et al., 1993; Yan & Garcia-Molina, 1994).

As Belkin and Croft (1992) explain, filtering information differs from information retrieval (IR) for a search request in several ways. Filtering is

designed to deal with an incoming stream of unstructured, or semistructured, data and implies removing data, while IR deals with a search of a remote database with highly structured data and implies finding data. Further, filtering is based on individual or group profiles that may change in time but that typically represent continuing interests, whereas IR is based on a momentary information need.

To filter information requires building a model of user's interests, or a filtering profile, which is based on the topics of interest to the user. Such models are difficult to build because of semantic and contextual complexities and because users' interests are constantly changing (Stadnyk & Kass, 1992). It is important, therefore, to study the issues involved in filtering systems from the users' point of view. At the same time, it might be useful to look for more stable filtering criteria, possibly those that do not relate directly to the subject matter. Such criteria may prove particularly useful for systems with multiple sources that cover various subject domains.

To date, very few user studies have been carried out (Gant, 1995). The study reported here examined users' perceptions of a filtering system in a real-life situation at a particular setting and with a single source of information. The study explored various aspects of filtering. Here we report on one aspect—i.e., those quality criteria users employed that did not relate directly to the topics in their filtering profile.

The study was carried out at the Boeing Company in the Puget Sound area of Washington. During 1991-93, the Technical Library worked with the publisher of the *Gartner Group Reports* and various Boeing user groups to establish a company-wide contract for the reports, which included electronic delivery of the text to any Boeing employee.

During the study period, the reports were delivered to the company via the Internet each month and distributed throughout the Puget Sound area through two mechanisms. The first was an unfiltered bulletin-board type system using a Boeing internal variation of the Unix Newsgroup model, which enabled users to subscribe to the *Gartner Group Reports* and to other publications. The second mechanism allowed users to establish a subject profile (a model of their interests) through an intermediary librarian in the Technical Library. They then received only those reports which met their profile criteria via their e-mail system as ordinary mail messages. Profiles were developed and maintained using Verity's TOPIC information filtering software.

RESEARCH METHOD

The study employed a variety of methods and instruments to investigate the same process. It used a combination of qualitative and quantitative methods, including data collection through observation, verbal protocols, questionnaires, and interviews.

The study had three phases. The first phase used observation of, and interviews with, selected users to determine the patterns of searching behavior, the factors perceived to be important for the selection of a filtering method, and the elements of perceived satisfaction. During the second phase, these data were analyzed and used to design a questionnaire that was administered to all users. The third phase included statistical analyses of the data that had been collected from the questionnaires and interviews that supported the interpretation of the statistical results.

In the first phase, we observed a total of fifteen users, both engineers and managers, as they examined the reports they received on the Newsgroup or via e-mail. We observed a total of thirty-four sessions. Four of the participants received unfiltered information, and we observed each one only once. The other participants received filtered reports, and we observed most of them during three consecutive deliveries of the reports.

Participants were asked first to explain why they looked for information and then to think aloud as they viewed the information on the screen. At the end of each session, users evaluated the session and its results. At the end of the observation period, we interviewed each participant to confirm our interpretation of the individual's searching behavior and to further investigate the reasons that led him or her to decide whether or not to filter the reports. All verbal protocols, think-aloud as well as interviews, were audio taped and transcribed.

Participants were extremely cooperative. They were very generous with their time and answered questions in great detail. Most of them found it comfortable to think aloud while browsing the reports. Generally, they liked to explain what they did and the reasons behind their decisions and actions.

In the second phase, we analyzed the data from the transcribed verbal protocols to identify the various factors that participants considered in the selection of filtering methods and in assessing their satisfaction. The analysis uncovered a host of quality criteria beyond the profiles' topics that the participants employed to determine whether or not a document was relevant.

This analysis guided the development of a questionnaire that was distributed to all users to validate the observation arrived at in the previous phase and to collect additional data. After a pilot test, the questionnaire was administered online. We attached it to the beginning of the next delivery of the *Gartner Group Reports* for all users who received them through the library filter and posted it twice on the newsgroup. The response rate from those users who received filtered information was 30 percent. However, because no list of the subscribers to the newsgroup was available, it could not be determined what the accurate response rate for this group was which comprised 15 percent of all the respondents. We received a total of eighty-three responses, and all were usable. The third phase of the study included statistical analyses of the data collected through the questionnaires.

QUALITY CRITERIA

Like most filtering mechanisms, the one used by the study participants was based on profiles that expressed topics or subjects. Respondents' profiles included words or phrases such as "computer security" or "client/server" that might occur in the text of a report. The average precision ratio at the observation phase was 23 percent for filtered information. This low figure suggested the need to look for additional criteria that could be used for filtering beyond topics or subjects. Supporting this approach were previous studies that were successful in identifying such criteria (e.g., Barry, 1994; Schamber, 1994), attempts at automatic recognition of whether or not an article is empirical (Haas et al., 1996), and the fact that the participants in the observation phase were highly articulate in expressing their reasons for accepting or rejecting reports. Thus, through the verbal protocols collected during the observation, we identified the criteria participants used to express the relevance of reports, and those they used to express the reasons a report was not relevant.

RELEVANCE CRITERIA

When participants in the observation phase examined a report on a topic of their interest, they still had to make a decision whether or not the report was relevant. They used various criteria as reasons for relevance. A report was considered relevant if one or more conditions were satisfied (see Figure 1).

- It was relevant to the Boeing Company
- It was about a product or a service that related directly to a project the participant was working on
- It was about new concepts, products, or services
- It was a case study
- It had hard data
- It displayed issues in a classified order and clearly (e.g., in the form of a list)
- It was written on a nontechnical level
- It described industry trends or gave predictions
- It was about a specific vendor, product, or service
- It confirmed or validated what the participant already knew
- It dealt with something the participant and his or her group had done
- It included background information or general information
- It had information that helped the participant keep up to date about a product with which he or she was familiar

Figure 1. Criteria for Judging a Report Relevant

NONRELEVANCE CRITERIA

Similarly, participants deemed reports on their topic as of nonrelevant interest using various criteria. A report was nonrelevant if one or more conditions were satisfied (see Figure 2).

- It was not relevant or applicable to the Boeing Company
- It was about something Boeing was already doing
- The participant had no influence over the issues the report raised
- The participant's group had already made a decision about the product or service that was addressed in the report
- It was about a technology that was not here yet
- It was completely nontechnical (e.g., about lawsuits or company analysis)
- It was about specific vendors
- It raised questions but gave no answers
- It expressed opinions rather than presenting facts
- The participant was not familiar with the product or the service the report was about
- It did not tell the participant anything he or she did not already know
- It took too long to understand what the report was about
- It was too basic or too general
- It was too detailed or too technical

Figure 2. Criteria for Judging a Report Nonrelevant

CATEGORIES OF QUALITY CRITERIA

These relevance and nonrelevance criteria could be viewed as factors to be used in addition to topics and subjects in filtering profiles. To explore the nature of these criteria, we examined the categories to which they belonged. We found that many criteria were attributes of the subject matter, and that these attributes were not limited to a subject domain.

To demonstrate this, it is easiest to first identify the categories that clearly do not relate to subject matter. Two categories that did not relate to the subject matter presented themselves immediately—i.e., style of writing and form or nature of a report. Two of the thirteen relevance criteria and four of the fourteen nonrelevant ones refer to issues of writing style. Some participants said they would read a report if it displayed issues clearly in a classified order (e.g., in the form of a list), or if it was written on a nontechnical level. To express writing style issues for deleting a report, participants said they would reject a report if it was completely nontechnical, it took too long to understand what the report was about, it was too basic or too general, or it was too detailed or too technical.

Similarly, three relevance criteria and two nonrelevance ones addressed the form or nature of a report. A report would be relevant to some participants if it was a case study, it had hard data, or it included background information. Alternately, it might be rejected if it raised questions but gave no answers, or it expressed opinions rather then presenting facts.

More than half of the quality criteria (eight of the thirteen relevant ones and eight of the fourteen nonrelevant ones) did not fit into these two categories. A close examination showed that they were all related, directly or indirectly, to the subject matter of a report. Consider, for example, the statement that one would keep a report because it was relevant to Boeing. What is it about the report that was relevant to Boeing? Definitely not the writing style or the form. Clearly, the subject matter was relevant to the company. For instance, a report might be on a client/server system that Boeing had already decided not to purchase. Another example is the case where a participant considered a report nonrelevant because she had no influence over the issues the report raised. Here again, the subject matter of the report played a central role in her decision to reject it.

Generally speaking, all the criteria that did not address the writing style or the form of a report could be converted to statements of the form: "It was about. . . ." As such, they all related to the subject matter of a report. Unlike the filter's topic, however, these relevance and nonrelevance criteria expressed *attributes* of the subject matter, whether subjective or objective, rather than the subject matter itself. Moreover, they were independent of the subject matter itself and thus were not limited to a subject domain. This analysis showed, therefore, that in addition to topics and subjects, their attributes should be used for filtering information.

THE VALUE OF THE QUALITY CRITERIA TO FILTERING

Both sets of quality criteria include those that are situational, depending on the individual's knowledge and situation at a certain time, as well as general or more objective ones. Criteria that are situational, such as "it confirmed what I already know" or "I had no influence over the issue," could be applied when filtering is done by a program that is constantly and directly negotiating the screening profile with each user. To integrate effectively such situational factors into filter construction would require a program with a learning mechanism and the use of artificial intelligence techniques.

On the other hand, the general and objective criteria, such as "it has hard data" or "it is about technology that is not here yet," could be integrated into filtering systems already available to improve their filtering performance. To demonstrate how such criteria could be employed, we analyzed three groups of criteria: those that were opposites in each list, those that were common to both, and those that were unique to each list.

Criteria that were Opposites

Criteria that appeared in both lists but in opposite directions might be considered strong indicators of relevance. These were attributes that caused respondents to conclude that a report was relevant, and the opposite of the same attributes served as a reason for nonrelevance. These are presented in Figure 3.

- It was relevant/nonrelevant to Boeing
- It did/did not present facts
- The participant was familiar/not familiar with the product
- It did/did not have new information
- It was written on a nontechnical/too technical level

Figure 3. Criteria that were Opposites

These criteria are important indicators for relevance because their presence implies relevance and their absence nonrelevance. As such, they can be used in filtering and are adequate to use for individuals as well as for group filtering. The situational criteria informed us that participants were very focused in their selection. They rejected reports about unfamiliar products and decided to examine reports about familiar products. Similarly, they rejected reports that included no new information and considered those with new information relevant. It is not advisable, however, to use the concept of "new information" as a relevance criteria because it is situational and relative as well. Two colleagues working on the same project, and with the same interests, experience, and knowledge, may disagree about whether or not a report includes new information. The criteria outlined here, however, suggest that a report that includes information never published before is likely to be of great relevance.

Moving away from the situational criteria, the list provides three factors that can be used for screening reports. Filtering for the participants in this study should eliminate reports (or move them to the bottom of a ranked list) that are not relevant to Boeing, those that present no facts, and those that are very highly technical. Similarly, reports with high or immediate relevance to Boeing, those rich with data and facts, those that present absolutely new information, and those that are written on a nontechnical level should be at the top of a ranked display.

CRITERIA COMMON TO BOTH

Criteria common to both were criteria that some participants used to explain why a report was relevant but others used for exactly the opposite reason—to explain why a report was nonrelevant. These are presented in Figure 4.

- It was nontechnical
- It was about a specific vendor
- The participant's group had made a decision about the product
- It was basic or general

Figure 4. Criteria Common to Both

That is, some participants wanted to read reports because they were nontechnical, about a specific vendor, or basic and general, while others decided to delete them for the very same reasons. These criteria are important for indexing because they were used to determine the relevance of reports, but they cannot be used automatically for filtering aimed at the participants as a group. Unlike other criteria, they have no absolute relevance-related value because, for some participants, they indicated relevance and for others nonrelevance. If they were used for indexing, participants could have known ahead of time, before they read a report, whether it was nontechnical, about a specific vendor, or whether it was basic or general. Participants could then make a selection according to their individual inclinations. These attributes can be easily determined and assigned during the intellectual processing of the reports (e.g., writing the abstracts), and thus improve the ease and efficiency of filtering and browsing.

UNIQUE CRITERIA

Some criteria represented no overlap between relevance and nonrelevance so that some were unique in determining relevance and others in determining nonrelevance, as shown in Figures 5 and 6.

- It displayed issues in a list form
- It described industry trends
- It confirmed or validated what the participant already knew
- It helped the participant keep up to date
- It related directly to the participant's project

Figure 5. Unique for Relevance

- It was about something Boeing was already doing
- The participant had no influence about the issues raised
- It was about a technology that was not here yet
- It raised questions but gave no answers
- It took too long to understand what the report was about

Figure 6. Unique for Nonrelevance

These unique criteria could be construed as an indication of preferences or a wish list. A user may read a report, for example, because it displays issues in a list form, but the same user would probably not reject all reports that had no lists. By itself, this would not be enough; reports would have to have other negative attributes to cause the user to reject them. Similarly, users, for example, may decide to delete a report because they think they have no influence over the issues raised. But no participant claimed to read a report just because he or she had influence in the matter. Again, other attractive attributes would have had to play a role when a user decided to read a report. Despite their somewhat weaker status, respondents considered some of these unique criteria as highly important. This is discussed in the next section.

While these unique attributes represent preferences rather than absolute criteria, they do play a role in relevance judgment and, therefore, should be considered when profiles are constructed. These unique criteria illuminate an important point—i.e., when constructing a filtering profile, users should be asked to indicate both relevance and nonrelevance criteria. This conclusion is important for all feedback processes to determine relevance. Criteria for both relevance and nonrelevance need to be ascertained because one is not always the opposite of the other, and because at times it is easier for a user to determine why a document is relevant and at other times it is easier to see why it is nonrelevant.

TOP CRITERIA ACCORDING TO USERS' CHOICE

During the observation, it became clear that some criteria were used more frequently than others, and that some were more important than others. In the questionnaire, we asked respondents to mark all the criteria from the nonrelevant list they would use to delete a report in the area of their interest. We then queried them to indicate which three they used most frequently and which three they considered most important. Similarly, we asked them to mark all the criteria from the relevance list that they would use to save or keep a report in the area of their interest and to indicate the top three in frequency and importance.

While each criteria proved useful to at least some respondents, data analysis showed that some criteria were much more significant than others. Further, when we compared the most popular criteria—i.e., those which the largest percentage of respondents reported employing, with those they ranked high in frequency of use and in importance—we found them to be the same. That is, the top criteria, measured according to popularity, were also rated by respondents as top in frequency of use and in importance. Figures 7 and 8 list these criteria in ranked order.

- It was about a product with which the participant was not familiar with
- It was about a technology not available
- It was about a specific vendor
- It was about something Boeing was already doing
- It was not relevant or applicable to the Boeing Company
- It included no information that was new to the participant

Figure 7. Top Nonrelevance Criteria

- It was related directly to the respondent's project
- It was about new concepts
- It was about industry trends
- It kept the respondent up to date
- It had hard data
- It was relevant to the Boeing Company

Figure 8. Top Relevance Criteria

An examination of the top criteria shows that most of these criteria, for nonrelevance as well as for relevance, are attributes of subject matter.

DISCUSSION

Two measurements assessed the value of the quality criteria to filtering. The first assumed that criteria in the relevance list that are the opposites of criteria in the nonrelevance list are the most promising criteria for filtering (see Figure 3). The second asked users to assess the weight of each relevance criteria as well as each of the nonrelevance ones. This assessment generated the list of top criteria (see Figures 7 and 8). Three criteria of those are prominent because they were ranked highly by both measurements. These criteria are:

- Whether or not a report was relevant to Boeing
- Whether or not a report had new information
- Whether or not the user was familiar with the product discussed in a report

The study findings showed that the top criteria had distinct characteristics. First, most of them were attributes of the subject matter, although participants in this study used these in addition to the topics of their interest. These criteria, however, were not limited to a specific subject domain because they could apply to any topic, from computer security to client/server systems.

Second, about half of the top criteria, and all of the prominent three, were subjective or situational and at times difficult to apply. Consider, for example, the prominent criteria. First is the statement "I kept it because it was relevant to Boeing" or its mirror image "I rejected it because it was not relevant to Boeing." This might seem a stable and objective statement. One might even suggest that information filters for all Boeing employees should consider the mission statement of the company. In reality, however, there are no general standards about what is relevant or not relevant to Boeing. Participants used their own perception of the company's interests when they used this criteria for relevance judgment.

Can such subjective criteria be considered when building a profile? In general, subjective criteria should not create any difficulties. After all, profiles are often built for an individual. The purpose is to integrate personal and subjective requirements and preferences. For this specific criteria, however, difficulties might arise when users are asked to provide comprehensive statements about their perception of what is relevant to Boeing. It is likely that they would be able to easily determine whether or not something is relevant to Boeing but would be unable to articulate the full range of considerations. Thus, to implement the strongest filtering criteria would require special investigative methods to elicit this information from the user.

The next prominent criterion is whether or not the information was new to the user. As mentioned earlier, this criterion is not only subjective and situational, it is relative as well. Important as it is, it is not stable and cannot be used as a filtering criterion, but only to recommend that reports with entirely new information should be ranked at the top of the list for all users. It is interesting to note that, while novelty was considered an important factor, some respondents reported that they would consider a report relevant if it confirmed or validated what they already know. That is, reports might have been considered relevant by some even if they did not carry new information as long as they confirmed or validated what these respondents already knew.

Third, users' choice of top quality criteria showed that the participants in the study were very focused on the tasks they had to perform. Not only did they reject reports that were not relevant to Boeing or that included no new information, they were interested in reports that related directly to their projects, kept them up to date, were not about products with which they were not familiar, or that were not about a technology that was not there yet. While this observation cannot be generalized to all users, it is plausible to assume that managers and engineers in other companies would exhibit similar filtering behavior. Thus, this study recommends that, in building filtering profiles in comparable environments, special attention be placed on considering subject attributes that are personal and subjective and that relate directly to the task at hand.

CONCLUSION

Our quest for stable filtering criteria for electronic delivery from multiple sources revealed that subject matter played a central role in filtering and on various levels:

- Relevance judgments based on subject matter, and thus filtering criteria, may present themselves in various ways. In addition to a straightforward statement of the topical relevance, other attributes of the subject matter might be important for filtering.
- Filtering criteria related to the subject matter can sometimes be expressed most effectively in a negative way, pointing to subject attributes that should not be present in retrieved documents. At times, users might consider such criteria important.
- In some contexts, users are likely to consider most important subject-related filtering criteria that are subjective and situational.

ACKNOWLEDGMENT

The study was supported by a grant from the Special Libraries Association.

REFERENCES

Anick, P. G.; Fly, R. A.; & Hansenn, D. R. (1991). Addressing the requirements of a dynamic corporate textual information base. In A. Bookstein, Y. Chiaramella, G. Salton, & V. V. Salton (Eds.), *Proceedings of the Fourteenth Annual International ACM/SIGIR Conference on Research and Development in Information Retrieval* (pp. 163-172). New York: Association for Computing Machinery.

Barry, C. (1994). User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science, 45*(3), 149-159.

Belkin, N. J., & Croft, W. B. (1992). Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM, 35*(12), 29-38.

Gant, S. P. (1995). A portrait of potential adopters of information filters. In T. Kinney, M. G. Lippert, & E. L. Steele (Eds.), *Proceedings of the 58th ASIS Annual Meeting* (pp. 167-171). Medford, NJ: Information Today.

Haas, S. W.; Sugarman, J.; & Tibbo, H. R. (1996). A text filter for the automatic identification of empirical articles. *Journal of the American Society for Information Science, 47*(2), 167-169.

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM, 37*(7), 31-40.

Schamber, L. (1994). Relevance and information behavior. In M. E. Williams (Ed.), *Annual review of information science and technology* (vol. 29, pp. 3-48). Medford, NJ: Learned Information.

Shuldberg, H. K.; Macpherson, M.; Humphrey, P.; & Corely, J. (1993). Distilling information from text: The EDS template filler system. *Journal of the American Society for Information Science, 44*(9), 493-507.

Stadnyk, I., & Kass, R. (1992). Modeling users' interests in information filters. *Communications of the ACM, 35*(12), 49-50.

Yan, T. W., & Garcia-Molina, H. (1994). Index structures for information filtering under the vector space model. In *Proceedings of the 10th International Conference on Data Engineering* (Houston, Texas) (pp. 337-347). Los Alamitos, CA: IEEE Computer Society Press.