A Dynamic Methodology for Improving the Search Experience

Marcia D. Kerchner

In the early years of modern information retrieval, the fundamental way in which we understood and evaluated search performance was by measuring precision and recall. In recent decades, however, models of evaluation have expanded to incorporate the information-seeking task and the quality of its outcome, as well as the value of the information to the user. We have developed a systems engineering-based methodology for improving the whole search experience. The approach focuses on understanding users' information-seeking problems, understanding who has the problems, and applying solutions that address these problems. This information is gathered through ongoing analysis of site-usage reports, satisfaction surveys, Help Desk reports, and a working relationship with the business owners.

Evaluation models

In the early years of modern information retrieval, the fundamental way in which we understood and evaluated search performance was by measuring precision and recall.[1] In recent decades, however, models of evaluation have expanded to incorporate the information-seeking task and the quality of its outcome, cognitive models of information behavior, as well as the value of the information to the user.[2] The conceptual framework for holistic evaluation of libraries described by Nicholson defines multiple perspectives (internal and external views of the library system as well as internal and external views of its use) from which to measure and evaluate a library system.[3] The work described in this paper is consistent with these frameworks as it emphasizes that, while efforts to improve search may focus on optimizing precision or recall, it is equally important to recognize that the search experience involves more than a perfect set of high-precision, high-recall search results. The total search experience and how well the system actually helps the user solve the search task must be evaluated.

A search experience begins when users enter words in a search box. It continues when the users view some representation (such as a list or a table) of candidate answers to their queries. It includes the users' reactions to the usefulness of those answers and their representa-

tion in satisfying information needs, and continues with the users clicking on a link (or links) to view content. Optimizing search results without considering the rest of the search experience and without considering user behavior is missing an opportunity to further improve user success. For example, the experience is a failure if typical users cannot recognize the answers to their information need because the items lack a recognizable title or an informative description, or they involve extensive scrolling or hard-to-use content.

Proposed solutions

Problems with search, such as low precision or low recall, are often addressed by either metadata solutions (adding topical tags to content objects based on controlled vocabularies) or replacement of the search engine. The problems with the metadata approach include the time and effort required to establish, evolve, and maintain taxonomies, and the need for trained intermediaries to apply the tags.[4] A community of stakeholders may be convened to define the controlled vocabulary, but often the lowest common denominator prevails, the champions and stakeholders leave, and no one is happy with the resulting standard. Even with trained intermediaries, inter-indexer inconsistency compromises this approach, and inconsistent term application can cause degradation of search results.[5]

Another shortcoming of the metadata approach is that a specific metadata classification is just a snapshot in time and assumes that there is only one particular hierarchy of the information in the corpus. In reality, however, there is almost always more than one way to describe a concept, and the taxonomy is the view of only one individual or group of individuals. In addition, topical metadata is often implemented with little understanding of the types of queries that are submitted or the probable user search behavior.

The other approach to improving search results—replacing a search engine—is not a guarantee to fixing the problem because it focuses only on improving precision (and perhaps recall as well) without understanding the true barriers to a successful search experience.

IRS.gov

IRS.gov, one of the most widely used government Web sites, is routinely accessed by millions of people each month (more than 27 million visits in April 2005). As an informational site, the key goal of IRS.gov is to direct visitors quickly to useful information, either through

Marcia D. Kerchner (mkerchner@mitre.org) is a Principal Information Systems Engineer at the MITRE Corporation, McLean, Va.

navigation or a search function. Given that there were almost 16 million queries submitted to IRS.gov in April 2005, search is clearly a popular way for its users to look for information. This paper offers an alternative to conventional search-improvement approaches by presenting a systems engineering-based methodology for improving the whole search experience. This methodology was developed, honed, and modified in conjunction with work performed on the IRS.gov Web site over a three-year period. A similar strategy of "sense-and-respond" for information technology (IT) departments of public organizations that involves systematic intelligence gathering on potential customer demand, a rapid response to fulfill that demand, and metrics to determine how well the demand was satisfied, has recently been described.[6]

The methodology described in this paper focuses on analyzing the information-seeking behaviors and needs of users and determining the requirements of the business owners (the IRS business operating divisions that provide content to IRS.gov, such as Small Business and Self-Employed, Wage and Investment) for directing users to relevant content. It is based on the assumption that a Web site must evolve based on its user needs, rather than expecting users to adapt to its singularities. To support this evolution, this approach leverages techniques for query expansion and document-space modification.[7] Dramatic improvements in quality of service to the user have resulted, enhancing the user experience at the site and reducing the need to contact the Help Desk. The approach is particularly applicable for those government, corporate, and commercial Web sites where there is some control over the content, and usage can be categorized into regular patterns. The rest of this paper provides a case study in the application of the methodology and the application of metrics, in addition to precision and recall, to measure search experience improvement.

Conceptual framework

While analysis of search results often focuses on search syntax and search-engine performance, there are actually several steps in the retrieval process, from the user identifying an information need to the user receiving and reviewing query results. As shown in figure 1, finding information is a holistic process. There are several opportunities to improve the whole user experience by fine-tuning this process with a variety of tools—from document engineering to results categorization. Once the user and business-owner needs are understood, the appropriate tools to address specific issues can be identified.

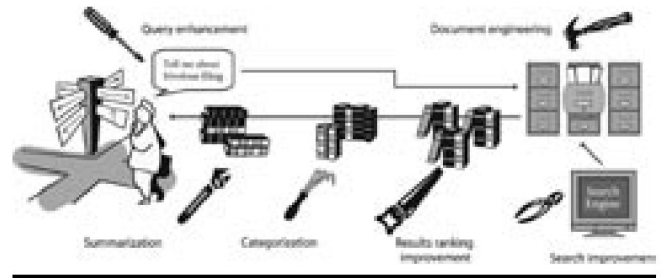The tools in our toolkit are described in the following sections.

Figure 1. The Information Retrieval Process

Document engineering

Document engineering includes:

- Document-space modification: Modifying the document space by adding terms to content (especially to titles) that are good discriminators and reflect terms commonly entered by users. This approach has the added benefit of making the content more understandable to users.
- Establishment of content-quality standards: Defining business processes that improve content quality and organization.

Document-space modification
There is significant syntactic and semantic impreciseness in the English language. In addition, because of the inadequacies of human or automatic keyword assignment, standard means of representing documents in indexes by statistical term associations and frequency counts or by adding metadata tags are not definitive enough to produce a space that is an exact image of the original documents. Document-space modification moves documents in the document space closer to future similar queries by adding new terms or modifying the weight of existing terms in the content (figure 2).[8] The document space is thus modified to improve retrieval. For IRS.gov, rather than adjusting content weights, titles and content are modified to adjust to changing terminology and user needs.

Establishment of content-quality standards
The quality of the search correlates with the quality of the content. Improved search results can be achieved by applying good content-creation practices. Retrieval can be significantly improved by addressing problems observed in the content. These problems include inconsistencies in term use—for example, Earned Income Credit (EIC) versus Earned Income Tax Credit (EITC)—duplicate content, insufficiently descriptive page titles, missing document summaries, misspellings, and inconsistent spellings.

Processes to improve content quality should establish standards for consistent term usage in content, as well as standards for consistent and descriptive naming of content types (for example, IRS types include forms, instructions, and publications). These processes will not only improve search precision, but will also help users identify appropriate content in the search results. For example, content entitled "Publication 503" in response to the query "child care" may be the perfect answer (with excellent precision and recall), but the user will not recognize it as the right answer. A title such as "Publication 503: Child and Dependent Care Expenses" will clearly point the user to the relevant information.

Usability tests conducted in March 2005 for IRS.gov confirmed that content organization plays an important role in the perceived success of a user's search experience. Long pages of links or scrolling pages of content left some users confused and overwhelmed, unable to find the needed information. For these queries, although the search results were perfect, with a precision of 100 percent after one document, the search experiences were still failures.

Query enhancement

The technique of relevance feedback for query expansion improves retrieval in an iterative fashion.[9] According to this approach, the user submits a query, reviews the search results, and then reports query-document relevance assessments to the system. These assessments are used to modify the initial query, that is, new terms are added to the initial query (hopefully) to improve it, and the query is resubmitted. If one visualizes the content in a collection as a space (figure 3), this approach attempts to move the query closer to the most relevant content.

A drawback of relevance feedback is that it is not generally collected over multiple user sessions and over time, so the next user submitting the same query has to go through the same process of providing results evaluations for query expansion. Borlund has noted that, given that an individual user's information need is personal and may change over session time, relevance assessments can only be made by a user at a particular time.[10] However, on IRS.gov, where there are many common queries for which there is a clear best-guess response, there is valuable relevance information that, if captured once, could benefit tens of thousands of users for specific queries. In fact, in April 2005, the top four hundred queries represented almost half of all the queries.

Another drawback of the relevance-feedback approach is that it forces the user, novice or expert, to become engaged in the search process. As noted previously, users are generally not interested in becoming search experts or in becoming intimately involved in the process of search. The relevance-feedback approach tries

to change users' behavior and forces them to find the specific word or words that will best retrieve the relevant information. In fact, some research has shown that the potential benefits of relevance feedback may be hard to achieve primarily because searchers have difficulty finding useful terms for effective query expansion.[11]

To avoid requiring users to submit relevance-feedback judgments, the methodology uses alternative approaches for gathering feedback: (1) mining sources of input that do not require any additional involvement on the part of the users; and (2) soliciting relevance judgments from subject matter experts.

As noted above, while best results may be different per task and per user, particularly given the shortness of the queries, our goal is to maximize the good results for the maximum number of people. Best-guess results are derived from a variety of sources, including usability testing, satisfaction survey questionnaires, and business-content owners. For example, users entering the common query "1040ez" can be looking for information on the form or the form itself. Given that—as shown in table 1 (based on the responses of 11,715 users to satisfaction surveys in 2005)—the goal of 39 percent of IRS.gov searchers is to download a form as opposed to 28 percent seeking to obtain general tax information, the retrieval of the 1040ez form and its instructions is prioritized, while also retrieving any general related information.
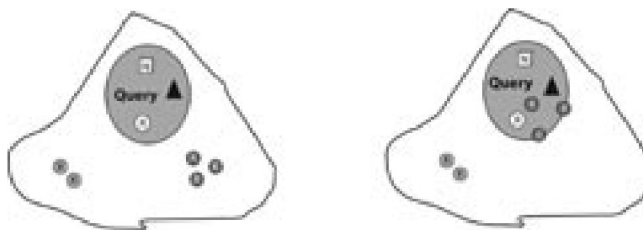

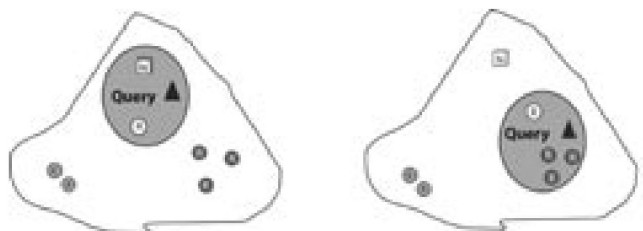**Figure 2.** Document-space modification

Figure 3. Query modification

We can determine the best-guess results as follows:

- Review the search results for terms that are on the frequently entered search-terms list
- Review Help Desk contacts, satisfaction-survey comments, and zero-results reports to identify information users who are having trouble finding or understanding
- Identify best results by working with the business owners as necessary
- Analyze why best results are not being retrieved for a particular query
    - Add appropriate synonyms for this and related queries
    - Engineer relevant documents (as described above)

In this way, the thesaurus, as the source for query enhancement, is an evolving structure that adapts to the needs of the users rather than being a fixed entity of elements based on someone's idea of a standardized vocabulary.

Search improvement

We can intercept very popular queries and return a set of preconfigured results or a quick link at the top of the search-results listing. For example, the user entering "1040" sees a list of the most popular 1040-related forms and instructions in addition to a list of other search results.

There were more than 31,000 users in April 2005 who requested the I-9 form. Since the form is not an IRS form, users are presented with a link to the Bureau of Citizen and Immigration Services Web site. The tens of thousands of users who look for state tax forms on IRS.gov are directed either to the specific state-tax-form Web-site page or to a page with links to state tax sites. This unique and user-friendly approach provides a significant improvement over a page that tells the user that there is no matching result, leaving him to fend for himself.

Another technique for improving search preci-

sion (not currently used for IRS.gov) is to tune and adjust parameters in the search engine, such as the relative weighting of basic metadata tags such as title (if they are used in the relevance calculation).

Results-ranking improvement

The search results can be programmatically re-ranked before being presented to the user. This approach (not used as yet on IRS.gov) is a variation on the quick links described above for re-ranking more than one result.

Categorization
A large set of search results can be automatically categorized into subsets to help the user find the information he needs. In addition, a "search within a search" function is available to help the user narrow down results. Research to be conducted on commercial products to support automatic categorization is planned for the future.

Summarization
As noted earlier, a barrier to a successful user experience can be the lack of informative descriptions in the search results. Therefore, an important tool for search-experience improvement is to make sure that content titles and summaries are informative, or as a second choice, that the search engine dynamically generates informative summaries. Passage-based summaries and highlighted search terms in the summary and the content have become a feature of many commercial search engines as another way to improve the usability of the returned results. In

Table 1. Reasons for using IRS.gov

Reason for coming to IRS.gov	% of total site visitors	% of total search users
Download a tax form, publication, or instructions	39	39
Obtain general tax information	27	28
Obtain information on e-file	10	10
Other	6	6
Obtain info on tax regulations or written determinations	4	4
Order forms from the IRS	3	4
Sign up or login to e-services	3	3
Link and learn (VITA/VCE) training	3	3
Obtain info on the status of your tax return	2	2
Use online tax calculators	1	1
Obtain info on revenue rulings or court cases	1	1
Obtain an Employer Identification Number (EIN)	1	—

Note: Due to rounding, totals may not equal 100%.

addition, for those PDF publications that lacked informative titles in the title tag, descriptive information from a different metadata field was added to the search display programmatically, which improved the usability of such results significantly.

Methodology

The methodology for evolving the search functionality is based on a logical, systems-engineering approach to the issue of getting users the information they seek: understanding the problems, understanding who has the problems, and applying solutions that address the problems. Usability studies, weblogs, focus groups, Help Desk contacts, and user surveys provide different perspectives of the information system.

The steps of the methodology are:

1. Understand the user population.
2. Identify the barriers to a successful search experience.
3. Analyze the information-seeking behaviors of the users.
4. Understand the needs of the business owners.
5. Identify and use the appropriate tools to improve the user's search experience.
6. Repeat as needed.
7. Monitor new developments in search and analytic technologies and replace the search engine as appropriate.

Step 1: Understand the user population

The first step is to profile and understand the user population. As mentioned above, an online satisfaction survey was conducted during a six-week period in January–February 2005, to which 11,715 users responded. The users were asked the frequency of their usage of the site, their primary reason for coming to IRS.gov, their category (individual, tax professional, business representative), and how they generally find information on IRS.gov.

As shown in tables 1–4, 76 percent of the IRS. gov visitors use it once a month or less (the largest group being those who use it every six months or less), or were using it for the first time; 64 percent are individual taxpayers; 10 percent are tax professionals; 39 percent visit

the site to download a form or publication; and 27 percent come for general tax or e-file information. Forty-nine percent use the search engine. Not surprisingly, 44 percent of the frequent visitors (those who visit once a week or more) are tax professionals, while 72 percent of the infrequent visitors are individuals or those who represent a business. The most common task of both the most frequent and infrequent visitors is to download a form, publication, or instructions, followed by obtaining general tax information. Most frequent and infrequent visitors use the search function to locate their information.

Thus, the largest group of IRS.gov users consists of average citizens, unfamiliar with the site, who have a specific question or a need for a specific form or publication. These users require high-precision, highly relevant results, and a highly intuitive search interface. They do not want or need to read all the material generated by their search, but they want their question answered quickly.

These users are generally not experienced with sophisticated query language syntax, and because they come to the site no more than once a month, they are not likely to be familiar with its navigational organization. As studies demonstrate, users in general do not want to learn a search engine interface or tailor their queries to the design of a particular search engine.[12] They want to find their information *now* before "search rage" sets in. One study observed that, on average, searchers get frustrated in twelve minutes.[13]

Tax professionals form a small but important group of IRS.gov users that includes lawyers, accountants, and tax preparers. They generally use the site on a regular basis, which could be daily, weekly, or monthly. Some of these users, particularly lawyers and accountants, require high relevance in their search results; it is critical that they retrieve every relevant piece of information (e.g., all the tax regulations) related to a tax topic. They may be willing to sift through large results sets to make sure they have seen all the relevant items. In contrast, many tax preparers use the site primarily to download forms and instructions.

While these different sets of users have different levels of expertise using the site and somewhat different precision and recall requirements, they do have one characteristic in common—they are not interested in search

Table 2. Frequency of visits to IRS.gov

	First time	Every six months or less	About once a month	About once a week	Daily	More than once a day
Site visitor	29%	34%	13%	13%	7%	4%
Search user	26%	34%	14%	14%	7%	5%

for its own sake. Approaches to improving retrieval results that focus on forcing users to use tools to refine their query to get presumably better search results (e.g., leveraging the power of Boolean or other search syntax) are not desirable in a public Web site environment. The complexity of the search must be hidden behind the search box and users must be helped to find information rather than be expected to master a search function.

Step 2: Identify the barriers to a successful search experience

There are several categories of reasons why finding information on a public Web site can be frustrating for the user.

- Mismatch between user terminology and content terminology
    - The user search terms may not match the ter-

Table 3. IRS.gov user types

Type of user	% of total site visitors	% of total search users
Individual taxpayer	64%	64%
Representing a business	11%	11%
Tax professional	10%	11%
Representing a charity or nonprofit	3%	3%
VITA/VCE volunteers	3%	3%
Representing a government entity	2%	2%
Student	2%	1%
IRS employee	1%	2%
Other	4%	3%

Table 4. How users find information on IRS.gov

How do you usually find information on IRS.gov?	% of total site visitors
Search engine	49%
IRS keyword	18%
Navigation to the Web page	11%
Internet search engine (e.g., Google, Yahoo)	7%
Site map	5%
Other	4%
Bookmarks	3%
Links to IRS.gov from other Web sites	3%

minology or jargon used in the content (e.g., users ask for "Tax Tables" or "Tax Brackets"; the IRS names them "Tax Rate Schedules").
- Multiple synonymous terms or acronyms are found because different authors are providing content on similar topics (e.g., "EIN," "employer identification number," "federal id number"; "EIC" versus "EITC").
- Users request the same information in a variety of ways (e.g., "1040ez," "1040-ez," "ez," "form1040EZ," "1040ez form," "2005 1040ez," "ez1040").
- Related content may be inconsistently named, complicating the user's search process (e.g., "1040X" form versus "1040-X" instructions).
- The user may use a familiar acronym that is spelled out in the content (e.g., "poa" for "power of attorney").
- Mismatch between user requests and actual content
    - Many users ask for information that they expect to find on the site but is actually hosted at another site (e.g., "ds156," a Department of State form; "IT-201," a New York State tax form).
- Issues with results listing and content display
    - Content may lack informative titles.
    - Automatically generated summaries may not be sufficiently descriptive for users to recognize the relevant material in the results listing.
    - Content may consist of long, scrolling pages, which users find hard to manage.
- Incomplete user queries
    - Very short search phrases (average length of less than two words) can make it difficult for a search algorithm to deduce the specific content the user is seeking.

Step 3: Analyze the information-seeking behaviors of the users

Site-usage reports, satisfaction surveys, Help Desk contact reports, zero-results reports, focus groups, and usability studies are valuable sources of information. They should be mined for information-seeking behaviors of the site's users and other barriers to a successful search experience, as follows:

- Review site-usage reports for the most frequently entered search terms and popular pages (both may change over time) and the zero-results search terms. Look for:
    - New terms
    - Variations on popular terms
    - Common misspellings or typos
    - Common searches, including searches for items

not on the site, that could be candidates for pre-programmed "quick links"

- Frequently entered terms—review search results to identify candidates for improvement
- Review satisfaction surveys over time
    - Look for new problems that caused satisfaction to decrease
    - Analyze answers to questions asking what people *could not* find, potentially identifying new barriers to success
- Conduct usability studies
    - Identify issues with the user interface as well as with content findability and usability
- Review Help Desk contact reports
    - Identify which topics users are having trouble finding or understanding

Step 4: Understand the needs of the business owners

The business owners are the IRS business operating divisions that provide content to IRS.gov, such as Small Business and Self-Employed, Wage and Investment. It is important to involve them in the process of enhancing the user experience, because they may have specific goals for prioritizing information on a particular topic or may be managing campaigns for highlighting new information. Thus it is desirable to:

- Meet with business owners regularly to understand their goals for providing information to users
- Work with them to increase the findability of their content

For example, when an issue in finding a particular content topic is identified (e.g., through an increase in Help Desk contacts), one approach is to show the business owner the actual results that common queries (based on the site-usage reports) on the topic retrieve and then present suggested alternative results that could be retrieved with a variety of enhancement techniques, such as thesaurus expansion or title improvement. The business owner can then evaluate which set of results presents the content in the most informative manner to the user.

Steps 1–4 facilitate work behind the scenes to gather the data needed to improve precision and recall and to make information more findable. The remaining steps use these data to adapt proven, widely used techniques for improving search experiences to a Web site's specific environment.

Step 5: Identify appropriate tools to improve the information-retrieval process

As described in the previous section, the tools in our toolkit are document engineering, query enhancement, search improvement, results-ranking improvement, categorization, and summarization.

Step 6: Repeat as needed

The process of improving the user search experience is ongoing as the site evolves. At IRS.gov, different search terms appear on the site-usage reports over time, depending on whether or not it is filing season, or as new content and applications are published. Human intervention (with the help of applicable tracking software) is essential for incorporating business requirements, evaluating human behavior, and identifying changing terms.

Step 7: Monitor new developments in search and analytic technologies and replace the search engine as appropriate

Although a new search engine will not address all the issues that have been described, new features such as passage-based summaries and term-highlighting can improve the search experience. Of course, one should consider replacing a search engine if new technology can demonstrate significantly improved precision and recall.

The application of the methodology and the use of the toolkit for IRS.gov will be described in the next section.

■ Findings

Site-usage reports

In 2003, an example of a serious mismatch in user and content terminology was discovered when site-usage reports were analyzed. Users entering the equivalent terms EIN, employer number, employer id number, and employer identification number retrieved significantly different sets of results. We met with the business owner, who identified a key-starting page that should be retrieved along with other highly relevant pages for all of these query terms. We recommended that "EIN" be added to the title of the key page because, although EIN is a very popular query, the acronym was not used in the content, but was instead spelled out. As a result, the key page was not being retrieved. Synonyms were added to the query enhancement thesaurus to accommodate the variants on the EIN concept.

After these steps were implemented, the results were as follows:

- For the query ein, the target page moved from #16 to #1
- For the query *ein number,* it moved from #17 to #5

- For the query *employer identification number,* it moved up to #2 (it was not in the top 20 previously)
- All search results now retrieved on the first page for these terms were highly relevant

In January 2004, there were approximately twenty thousand queries using these terms, so the search experience has been improved for tens of thousands of users in one month and hundreds of thousands of users throughout the year.

Review of Help Desk contacts

Help Desk reports summarize, for each call or e-mail, the general topic of the user's contact (filing information, employer ID number, forms, and publications issues) and the specific question. For example, the report might indicate that a user needed help in finding or downloading the W-4 form or did not understand the instructions for amending a tax return. As Help Desk contact reports were reviewed, clusters of questions emerged indicating information that many users could not find or understand. By analyzing approximately 9,800 contacts (e-mail, telephone, chat) during a peak five-day period in April 2003, four particular areas were identified that were ripe for improvement: 480 users could not find previous years' forms, which, although they can be found on the site, are not indexed and thus not findable through search; 250 users had questions about where to send their tax returns; 170 users had questions about getting a copy of their tax return or W-2 form; and 77 users had problems finding the 1040X or 1040EZ forms.

Utilizing the information retrieval toolkit, the following improvements were implemented:

a) Search for previous years' forms
   Tool used: Results-ranking improvement

A user requesting a previous year's forms (for example, 2002 1099misc) is now presented with a link directly to the page of forms for that specific year, as follows:

Recommendation(s) for: **2002 1099misc**

- **2002 Forms and Publications**
  2002 Forms, instructions, and publications available in PDF format

b) Request for filing address
   Tools used: Document engineering and query enhancement

A new "where to file" page was created. Synonyms were added to the thesaurus to accommodate the variations on how people make this request (address, where to send, where to mail) and to prioritize retrieval of the "where to file" page.

c) Request for information about obtaining a copy of a tax return or W-2 form
   Tools used: Results-ranking improvement and query enhancement

A "quick link" was created to the target page for getting a copy of returns and W-2 forms and synonyms were added to the thesaurus to prioritize related content for any query containing the word "copy."

d) Requests for 1040X or 1040EZ forms or instructions
   Tool used: Query enhancement

Synonyms were added to the thesaurus to address both the variations on how users requested the 1040X and 1040EZ forms and instructions, and the inconsistencies in the titling of these documents (for example, the form and the instructions have different variations of the compound name).

Results

In 2004, approximately 4,200 contacts were reviewed with the Help Desk during the same time period (the week before April 15) to see whether the changes actually did help users find the information. It should be noted that, during this period from April 2003 to April 2004, many other improvements to the user search experience based on the methodology were deployed.

Although the number of visits to IRS.gov increased by approximately 50 percent compared with the same period in 2003, the total number of contacts with the Help Desk decreased by 47 percent (there were approximately 9,800 contacts in this period in 2003). The results for the specific improvements are shown in table 5. The average decrease in contacts for those four topics was 68 percent, compared with the average decrease of 47 percent.

This approach has significantly improved the user experience by identifying and addressing subject areas users have trouble finding or understanding on IRS.gov, eliminating the need for them to contact the Help Desk. As a result, an increase of resources at the Help Desk was avoided and, hopefully, user satisfaction improved.

Conclusions

While the case presented in this article was specific to IRS.gov, the methodology itself has wide application across domains. Customer service for most government and commercial organizations depends on providing users with relevant information effectively and efficiently. There are many aspects to achieving this elusive goal of matching users with the specific information they need. In this paper, it has been demonstrated that, rather than focusing just on optimizing the search engine or developing a metadata-based solution, it is essential to view the user search experience from the time content is created to the moment when users have truly found the answer to their information needs. There is no one surefire solution, and one should not assume that enhanced metadata or a new search engine is the only solution to retrieval problems.

The methodology described in this paper assumes that users, especially infrequent users of public Web sites, do not wish to become search experts; that intuitive interfaces and meaningful results displays contribute to a successful user experience; and that keeping business owners involved is important.

The methodology is based on understanding the behavior of a site's users in order to identify barriers to a successful search experience, and on understanding the needs of business owners. The methodology focuses on adapting the site to its users (rather than vice versa) through document modification, improved content-development processes, query enhancement, and targeted search improvement. It includes improvements to the results phase of the search process, such as improved titles and summaries, as well as to the search-and-retrieval phase.

This toolkit-based approach is effective and low-cost. It has been used over the past four years to improve the user search experience significantly for the millions of IRS.gov users. Interesting follow-on research could focus on identifying to what degree this methodology can be automated and how to leverage new tools to provide automated support for usage log analysis (such as MondoSearch by Mondosoft).

It is clear from this case study that it is time to apply systems engineering rigor to search-experience improvement. This approach confirms the need to extend metrics for evaluating search beyond precision and recall to include the totality of the search experience.

Table 5. Comparison of 2004 and 2003 Help Desk contacts

Problem area	Number of contacts 2003	Number of contacts 2004	Change
1040X, 1040EZ	77	19	-75%
Prior year forms	480	103	-78%
Copy of return	170	91	-47%
Where to file	250	104	-58%
Total	977	317	-68%

Future work

Teleporting has been defined as an approach in which users try to jump directly to their information targets.[14] Trying to achieve perfect search results supports the information-seeking strategy of teleporting. But the search process may involve more than a single search. People often conduct "a series of interconnected but diverse searches on a single, problem-based theme, rather than one extended search session per task."[15] This approach is similar to the sport of orienteering with searchers using data from their present situation to determine where to go next—that is, looking for an overview first and then submitting more detailed searches. Given the general, nonspecific nature of the short queries submitted by IRS.gov users, the orienteering approach may well describe the information-seeking behaviors of many users. This paper is limited to the improvement of search results for individual searches, but the need to investigate improving the search experience to support orienteering behavior is acknowledged. Future research will investigate how to leverage the theoretical models of the information-search process, such as the anomalous states of knowledge (ASK) underlying information needs and the Information Search Process model.[16]

References and notes

1. "Common Evaluation Measures," *The Thirteenth Text Retrieval Conference,* NIST Special Publication SP 500-261 (Gaithersburg, Va.: National Institute of Standards and Technology, 2004), appendix A.

2. Kalervo Jarvelin and Peter Ingwersen, "Information-Seeking Research Needs Extension towards Tasks and Technology," *Information Research* 10, no. 1 (2004), http://InformationR.net/ir/10-1/paper212.html (accessed Feb. 2, 2006); K. Fisher, S. Erdelez, and L. McKechnie, eds., *Theories of Information Behavior* (Medford, N.J.: Information Today, 2005); T. Saracevic and Paul B. Kantor, "Studying the Value of Library and Information Services, Part I: Establishing a Theoretical Framework," *Journal of the American Society for Information Science.* 48, no. 6 (1997): 527–42.

3. Scott Nicholson, "A Conceptual Framework for the Holistic Measurement and Cumulative Evaluation of Library Services," *Journal of Documentation* 60, no. 2 (2004): 164–82.

4. Avra Michelson and Michael Olson, "Dynamically Enabling Search and Discovery TEM," Internal MITRE presentation, McLean, Va., Mar. 30, 2005.

5. Lawrence E. Leonard, "Inter-indexer Consistency Studies, 1954–1975: A Review of the Literature and Summary Of Study Results," Occasional Paper Series, No. 131, Graduate School of Library Science, University of Illionois, Urbana-Champaign, 1977; Tefko Saracevic, "Individual Differences in Organizing, Searching and Retrieving Information," in *Proceedings of American Society for Information Science '91* (New York: John Wiley, 1991), 82–86; G. Furnas et al., "The Vocabulary Problem in Human-System Communication," *Communications of the ACM* 30, no. 11 (1987): 964–71.

6. Rajiv Ramnath and David Landsbergen, "IT-enabled Sense-and-Respond Strategies in Complex Public Organizations," *Communications of the ACM* 48, no. 5 (2005): 58–64.

7. T. L. Brauen et al., "Document Indexing Based on Relevance Feedback," Report No. ISR-14 to the National Science Foundation, Section XI, Department of Computer Science, Cornell University, Ithaca, N.Y., 1968; M. C. Davis, M. D. Linsky, and M. V. Zelkowitz, "A Relevance Feedback System Employing a Dynamically Evolving Document Space," Report No. ISR-14 to the National Science Foundation, Section X, Department of Computer Science, Cornell University, Ithaca, N.Y., 1968; Marcia D. Kerchner, *Dynamic Document Processing in Clustered Collections,* Report No. ISR-19 to the National Science Foundation, Ph.D. thesis, Department of Computer Science, Cornell University, Ithaca, N.Y., 1971.

8. Ibid.

9. Gerard S. Salton, *Dynamic Information and Library Processing* (Englewood Cliffs, N.J.: Prentice-Hall, 1975).

10. P. Borlund, "The IIR Evaluation Model: A Framework for Evaluation of Interactive Information Retrieval Systems," *Information Research* 8, no. 3 (2003), http://informationr.net/ir/8-3/paper152.html (accessed Feb. 15, 2006).

11. Ian Ruthven, "Re-examining the Effectiveness of Interactive Query Expansion," in *Proceedings of the 26th International ACM SIGIR Conference on Research and Development in Information Retrieval* (New York: ACM Press, 2003), 213–20.

12. Marc L. Resnick and Rebecca Lergier, "Things You Might Not Know about How Real People Search," 2002, www.searchtools.com/analysis/how-people-search.html (accessed Oct. 1, 2005).

13. Danny Sullivan, "WebTop Search Rage Study," The Search Engine Report, 2001, http://searchenginewatch.com/sereport/article.php/2163451 (accessed Sept. 10, 2005).

14. J. Teevan et al., "The Perfect Search Engine Is Not Enough: A Study of Orienteering Behavior in Directed Search," in *Proceedings of Computer-Human Interaction Conference '94* (New York: ACM Press, 2004), 415–22.

15. Vicki O'Day and Robin Jeffries, "Orienteering in an Information Landscape: How Information Seekers Get from Here to There," in *Proceedings Interchi '93* (New York; ACM Press, 1993), 438.

16. N. J. Belkin, R. N. Oddy, and H. M. Brooks, "ASK for Information Retrieval, Part I. Background and Theory," *The Journal of Documentation* 38, no. 2 (1982): 61–71; N. J. Belkin, R. N. Oddy, and H. M. Brooks, "ASK for Information Retrieval, Part II. Results of a Design Study," *The Journal of Documentation* 38, no. 3 (1982): 145–64; Carol C. Kuhlthau, *Seeking Meaning: A Process Approach* (Norwood, N.J.: Ablex, 1993).