

EVALUATING SEARCH FEATURES IN PUBLIC ADMINISTRATION WEBSITES

**Carpineto Claudio (*), D'Aloisi Daniela (*), Delogu Cristina (*),
Ragazzini Susanna (*), Bianchi Carlo (**), Virdis Giovanni (**),
Bernardini Andrea (***)**

()Fondazione Ugo Bordonj, Rome, ITALY*

e-mail: [carpinet](mailto:carpinet@fub.it), [daloisi](mailto:daloisi@fub.it), [cdelogu](mailto:cdelogu@fub.it), sragazzini@fub.it

*(**)Istituto Superiore delle Comunicazioni e delle Tecnologie dell'Informazione, Rome, ITALY*

e-mail: carlo.bianchi@istsupcti.it, giovanni.virdis@istsupcti.it

*(***)Università di Roma 3, Rome, ITALY*

e-mail: aberna@fub.it

Abstract

Search is one of the most important user interface elements in any complex Website, including the Public Administration ones. However, there is no standard methodology to evaluate a site's search quality and there is a dramatic shortage of evaluation studies. This paper represents a first step in this direction. We present an evaluation methodology that covers the three main dimensions of the search process, namely, efficiency, effectiveness, and user interface. The proposed methodology is applied to a sample of the Italian Public Administration Websites consisting of 155 sites. The most striking results of our evaluation study are that more than half of the sites did not have an internal search engine and that the quality of search features in the remaining sites was, in general, low, with some notable variations. In particular, search efficiency was fairly good, especially to the extent that response times were concerned, while search effectiveness and search interface were unsatisfying. One serious shortcoming of almost all sites examined was the lack of features for supporting interaction with users.

Keywords: web site search, evaluation of web site search, public administration web sites

INTRODUCTION

Search is the key to accessing the information published on the Web. It consists of providing some description of the information being sought on the part of the user and having the system retrieve database items that match the description. According to the biannual WWW surveys conducted at the Graphics, Visualization and Usability Center (GVUC) of the Georgia Institute

of Technology - one of the more reliable sources on the user data - about 85% of the Web users surveyed claimed to be using search engines or some kind of search tools to find specific information of interest.

Search is the predominant paradigm for accessing not only the information published in the whole Web but also in any complex Website. Website usability studies suggests that more than half of all users are search-dominant, about a fifth of the users are link-dominant, and the rest switch between search and link-following, without an inherent preference [Nielsen1997]. One of the main reasons for preferring search is that it leaves full control to the user; search is also seen as a way out when one gets stuck in navigation [Nielsen 2000].

Over the last few years, the research on more effective and usable search techniques has produced a number of advances, some of which have been incorporated into commercially available systems. Multilingual and cross-lingual information retrieval [Oard1999], link-based ranking [Henzinger2000], automatic query expansion [Carpineto2001], graphical user interfaces [Hearst1999, Berenci2000], XML search engines [Luk2000], and multimedia information retrieval [Hauptmann1997] are just a few examples of the renewed interest for information search and retrieval.

Search is especially important for institutional Websites. Institutional websites contain a large number of pages, they may offer multiple services in a single site, and their users present a large variation in terms of skills, Web experience, and knowledge of the domain being searched. Thus, the possession of search tools of good quality would be highly desirable for mastering the inherent complexity of such sites.

However, the problem of evaluating the quality of institutional Websites' internal search engines has so far received very little attention. Most of the studies that have recently attempted to evaluate the quality of the Italian Public Administration Websites focused on contents and administrative services (http://www.atenea.it/html/risorse/form_ARPA-C.html) or on different elements of the site's user interface ([http://www.aipa.it/attivita\[2/gruppi\[18/accessibilita\[3/bozza\[1/index.asp](http://www.aipa.it/attivita[2/gruppi[18/accessibilita[3/bozza[1/index.asp)), without addressing the search issue. This paper wants to help fill this gap by offering a two-fold contribution.

The first contribution is a methodology for evaluating the features of internal search engines. The methodology deals with all main aspects involved in the search process, including efficiency, effectiveness, and user interface. It is comprehensive, yet simple to use. The second main contribution is an evaluation study of the search features of Italian Public Administration Websites performed using the introduced methodology. The results of our study suggest that site's search is somewhat overlooked - internal search is, in general, either lacking or of low quality - and provide some indications as to how to improve its quality.

The rest of the paper has the following structure. We first discuss the main dimensions that affect the quality of search. Then we describe the evaluation methodology. At this point the evaluation study is presented. After describing the choice of the Website sample, the results are illustrated and their implications are discussed. Finally, some conclusions are drawn.

SEARCH QUALITY DIMENSIONS

Although statistics and demographics on search engine usage may vary between different sources, they seem to consistently rate among the most commonly experienced problems: (a) slow access, (b) inability to find relevant information, and (c) inadequacy of user interface.

The majority of web users surveyed at GVUC over a period of a year-and-a-half left a Website while searching for information because of slow access, and the problem appeared to be growing worse. Other studies - see [Huberman1998] and [Kobayashi2000] for a more extensive review - cite the number one and number two reason for dissatisfaction as "slow access" and "the inability to find relevant information".

The latter problem - mainly due to the limited ranking capabilities of available systems - may be exacerbated by a scarce amount of interaction between the user and the system, as observed in many Web search sessions. Users typically do not peruse long lists of documents (in this paper we will often use "document" and "page" interchangeably) obtained in response to a query and they are only moderately able or willing to refine their queries. Furthermore, users may make very quick judgements about a Website's value based on the quality of one or two sets of search results, after which they may decide to abandon the site altogether or to forgo the site's search in favor of external search engines like Google [Nielsen2000].

The quality of retrieved information is cited as one of the main reasons for dissatisfaction also by [Lawrence1998]. In addition, they note that users are not necessarily satisfied with several user interface elements such as formats for inputting queries and presentation formats of the retrieval results. [Nielsen2001] analytically study the impact of the search interface details on systems's retrieval effectiveness and user satisfaction, pointing out a number of useful guidelines.

These observations and findings suggest that the quality of search features could be conveniently analyzed along three main dimensions, which we term, for the sake of simplicity, search efficiency, search effectiveness, and search interface. Although these dimensions may present some overlaps, they can be evaluated independently.

Search efficiency is mainly concerned with search speed and maintenance. Search speed is given by search engine search and retrieval time plus communication delays. While the latter may be

influenced by a number of factors such as the speed of connectivity to the network by users, the performance of the host Web server, and the uneven concentration of information packets on the network, the former is an isolated technical problem depending on the efficiency of the access and retrieval algorithms used by the search engines. Speed is not the only factor relevant to search efficiency. As many sites continue to grow and change fast, their internal search engines must be maintained and updated. The newly-created pages must be incorporated into the search engine's database, and possible dead links and duplicates should be eliminated.

Retrieval effectiveness is affected by the techniques used by the system for indexing the pages to be retrieved and for processing the user queries. It is also related to the possibility for the user to perform more sophisticated searches or to receive support from the system for entering queries and analyzing results. A number of specific details of search capabilities can be used to judge retrieval effectiveness, similar to feature charts used for rating external search engines (see, for instance, Web Search Engine Watch, <http://searchenginewatch.com/reports/index.html>).

The retrieval effectiveness of search systems can also be evaluated more quantitatively in terms of precision and recall, which measure, respectively, the ability of the system to retrieve only relevant documents and the ability of the system to retrieve all relevant documents. Most "quantitative" evaluation methods are based on a combination of these two criteria and require the possession of a set of documents, a set of topics, and relevance judgements for each topic (i.e., which documents are relevant to each topic). Several test collections of this kind are available, the best known of which are, perhaps, those developed by the TREC (Text Retrieval Conference) initiative [Voorhees2000]. Although this evaluation method too has shortcomings, it represents the most used and reliable way to measure the retrieval effectiveness of ranked output search systems for single-query searches. One practical difficulty of this method is that the system being evaluated must be run against the test collection, which may be unfeasible in many circumstances.

The user interface is the third main quality dimension that we consider. To build a Website that enables user to search successfully, it is necessary to look at the process from the user's point of view. Following [Nielsen2001], four main aspects should be dealt with: (a) the user notices the Website supports search, (b) the user expresses a question as a search query string, (c) the user enters the search string and submits the query, (d) the user receives and interprets the search results.

In the next section we describe a search quality evaluation methodology - based on a number of specific observable features - that accounts for all three main dimensions described above.

EVALUATION METHODOLOGY

We were interested in defining a methodology for evaluating the features of existing internal search engines, without assuming that their code is made accessible or that they may be used for indexing user-specified collections. Thus, we concentrated on those features whose quality measurement did not require experimenting with test collections.

For each dimension - search efficiency, search effectiveness, and search interface - we defined a number of features to be evaluated. For each feature, we defined an evaluation method and an evaluation measure. The search interface features are borrowed from [Nielsen2001], with some small changes due to the different nature of the domain of interest (Nielsen et al.'s study focused on commercial Websites).

In the following, the three quality dimensions are examined in turn. For each quality dimension, it is shown a table illustrating its associated features, including name, evaluation methodology, and evaluation measure. A complete list of the features that contains a more detailed discussion of their rationale as well as of their evaluation method is also provided.

Search efficiency

Fast response time. Speed of response time is assessed by performing a number of sample searches and averaging their response times. Communication delay is estimated by performing different operations that require no additional processing time (e.g., link navigation) on the same site at the same time. If the time strictly necessary for search engine search and retrieval exceeds two seconds, the response time is not deemed to be fast.

High coverage. The search engine should be able to retrieve any page contained in the site. The coverage of the search engine is considered to be high if all pages of a randomly-chosen page sample may be retrieved by the search engine.

| Name | Evaluation Methodology | Evaluation Measure |
|--------------------|------------------------------------|--------------------|
| Fast response time | Objective estimate by query sample | Yes/No |
| High coverage | Objective estimate by page sample | Yes/No |
| No broken links | Objective estimate by query sample | Yes/No |
| Freshness | Objective estimate by query sample | Yes/No |

Table 1 – Evaluation features for search efficiency

Freshness. The search engine should be incrementally updated as new pages are added to the site. Freshness is assessed by checking whether all pages of a sample of recently-created pages

may be retrieved by the search engine.

No broken links. The search engine should not return results containing broken links. The absence of broken links is assessed by checking if all queries of a query sample do not yield broken links.

Search effectiveness

Case insensitivity. Upper- and lower-case characters should not be distinguished when indexing pages and queries. Case insensitivity is assessed by checking whether a randomly-chosen page is retrieved by the search engine using a same key word with different cases.

| Name | Evaluation Methodology | Evaluation Measure |
|--|-----------------------------------|----------------------|
| Case insensitivity | Objective estimate by page sample | Yes/No |
| Inexact matching | Objective estimate by page sample | Yes/No |
| Stemming | Objective estimate by page sample | Yes/No |
| Support for scoped search | Objective inspection | Yes/No |
| Support for Advanced Search | Objective inspection | Yes/No |
| Support for phrase searching | Objective inspection | Yes/No |
| Support for Boolean search | Objective inspection | Yes/No |
| Sort by relevance | Objective inspection | Yes/No |
| Multilingual support | Objective inspection | Yes/No |
| Selectable number of displayed results | Objective inspection | Yes/No |
| Support for refinement of search results | Objective inspection | Yes/No |
| Support for query reformulation | Objective inspection | Yes/No |
| Use of metadata | Objective estimate by page sample | Yes/No/Doesn't apply |

Table 2 – Evaluation features for search effectiveness

Inexact matching. Users may easily get confused by their search results when the search engine looks for exact matching only. The system should accommodate for multiple-word queries by allowing inexact matching between the words in the query and the words in each document. Inexact matching is assessed by checking whether a two-word query retrieves the documents that contain either word.

Stemming. It seems convenient to reduce each word to word-stem form, at least for the

standard inflections of nouns and adjectives. Stemming is assessed by checking whether a randomly-chosen page is retrieved by the search engine using a same word with different variant forms (i.e., singular/plural, male/female).

Support for scoped search. Scoped search lets user limit the search to results from specific areas of the site. This functionality may be especially useful when a single site offers multiple services.

Support for Advanced Search. Advanced search may be useful for experienced users in some circumstances, notably as an alternative option to default search after a poor search results page. Two specific forms of Advanced Search are considered below.

Support for phrase searching. On request, the system should be able to do an exact search for the whole string entered by the user.

Support for Boolean search. Boolean search should be supported, as this may be more effective than best matching search for experienced users.

Sort by relevance. Sorting the results by relevance is essential when the number of hits retrieved in response to a query is large, as it is often the case.

Multilingual support. Multilingual querying and document presentation could be useful for many potential users who cannot write or read Italian.

Selectable number of displayed results. The possibility for users to choose the number of results returned on the first page makes the interpretation or peruse of results more flexible.

Support for refinement of search results. The system should provide the user with mechanisms to refine or enlarge the search results, which are notably useful when the search engine returns too many results or no results at all.

Support for query reformulation. Search interfaces should show users easy ways to refine queries, mainly based on the number and characteristics of search results and on the knowledge of the content of the documents being searched.

Use of metadata. Metadata and annotation can facilitate fast and accurate search and retrieval. This aspect may become even more important as the semantic Web makes progress. Use of metadata is assessed by checking whether a randomly-chosen page is retrieved by the search engine using the metadata encoded in the page itself, provided that that page does use metatags.

Search interface

Accessibility for disabled persons. The search engine should be made accessible to people with handicaps. Platform- and browser-independent accessibility for disabled persons requires compliance with the W3C accessibility guidelines (<http://www.w3.org/TR/2000/NOTE-WCAG10-HTML-TECHS-20000920/>). This is ascertained by inspecting the source code of the Web page containing the search engine.

Search box on the home page. The search function should be clearly visible when entering a Website. For many users, tabs and links to a separate search page just don't work. It is preferable to use a text box.

| Name | Evaluation Methodology | Evaluation Measure |
|---|-----------------------------------|----------------------|
| Accessibility for disabled persons | Objective inspection | Yes/No |
| Search box on home page | Objective inspection | Yes/No |
| Large search box | Objective inspection | Yes/No |
| Search button | Objective inspection | Yes/No |
| Search function on every page | Objective estimate by page sample | Yes/No |
| Search box on every page | Objective estimate by page sample | Yes/No |
| Scoped search explained | Subjective inspection | Yes/No/Doesn't apply |
| Discreet Advanced Search | Subjective inspection | Yes/No/Doesn't apply |
| Context of search results | Objective inspection | Yes/No |
| Relevance scores hidden | Objective inspection | Yes/No |
| Matching keywords highlighted | Objective inspection | Yes/No |
| Visible "No Results" message | Objective inspection | Yes/No |
| Search function on a search results page | Objective inspection | Yes/No |
| Search box on a search results page | Objective inspection | Yes/No |
| Constructive advice for "No results" searches | Subjective inspection | Yes/No |

Table 3 - Evaluation features for search interface

Large search box. The search box should be able to handle long queries. We set the lower bound of the maximum number of characters accepted by a good search box to 30.

Search button. After entering a query, users should start a search by clicking on a button. The button label itself is less important.

Search function on every page. Easy access to search should be guaranteed from any page in the site, except for the cases when users would risk losing their work (e.g., form filling, checkout process).

Search box on every page. Even on pages other than the home page, it is preferable to use a text box.

Scoped search explained. The scope of the search - when applicable - should be explicitly stated at the top of the results page, and suggestions for limiting or enlarging the current scope should be provided.

Discreet link to Advanced Search - and back. Advanced Search - when applicable - should be offered with caution, as it may easily lead users into trouble. It is better not to make Advanced Search accessible from the home page and to make its link less prominent than the Search button. When the users get a No Results message in Advanced Search, a visible link back to simple Search should be provided.

Context of Search Results. After a user has entered one or more search keywords and activated the search, the Search Results page should always display the search criteria - possibly including the scope of search - and the total number of search results.

Relevance scores hidden. Relevance scores are just noise for most users, who aren't aware of or interested in the factors that influence ranking. Scores should be used only for ranking the list from highest to lowest.

Matching keywords highlighted. Highlighting the document keywords which matched the query keywords and showing the sentences surrounding the matching keywords is a simple and effective way to show why results were included and avoid displaying seemingly irrelevant hits.

Visible "No Results" message. If the search fails, one must make it sure that the user notices the "No Results" message, without mistaking the No Results page for a Search Results page.

Search function on a search results page. Users should not be forced to navigate backward from the Search Results or No Results page to try another query. Search functionality should be placed directly in the content area of such pages.

Search box on a search results page. Search box is recommended also for Search Results or No Results pages.

Constructive advice for "No results" searches. After making clear why the search failed, the system should suggest possible ways to avoid the No Results search (e.g., checking the spelling, broadening the search) or alternative ways of locating information (e.g., navigation, table of contents, site map, site's FAQ).

A CASE STUDY: THE ITALIAN PUBLIC ADMINISTRATION WEBSITES

We applied the evaluation methodology described above to the Websites of the Italian Public Administration. In the next two sections, we describe the Website sample chosen for evaluation and the results of our study.

Website sample

In the study, conducted in June 2001, the following sites were considered.

- The Chief Towns of Italy. Of the 103 administrations, 4 did not have a Website, 8 had an unavailable Website, 65 did not have a search engine or had a search engine which did not work, and only 26 had a running system.
- The Regions of Italy. All 20 regions had a Website, 12 with and 8 without an internal search engine.
- The Presidency of the Republic and the two branches of Parliament. The Presidency of the Republic and the Chamber of Deputies had a search engine, the Senate had a Website without search engine.
- The Presidency of the Council of Ministers and the Ministries. The former and 12 ministries had a search engine, while 7 ministries had a Website without search engine.
- Miscellaneous of main institutional bodies. We selected 9 administrations (see Table 5 for a detailed list), all of which had a site's search engine.

In all, of the 155 administrations considered for evaluation, 4 did not have a Website, 8 had a Website which, however, did not happen to work, 81 (more than half of the sample) had a Website but they did not have a search engine or had a search engine which did not work. Thus, we actually evaluated the search quality of the remaining 62 sites.

The evaluation was performed using Internet Explorer on a Personal Computer with a fast connection to the Internet. In the next section we present the results.

Results

For the sake of illustration, the evaluation features are numbered. The numerical labels assigned to the features are shown in Table 4. In Table 5 we report, for each site (denoted by its URL without the prefix "www."), the complete list of search quality features possessed by that site (i.e., "x" means that the evaluation measure for the corresponding site-feature pair is "yes"). These raw results are the starting point of our analysis.

| | |
|--|----|
| High speed | 1 |
| High coverage | 2 |
| No broken links | 3 |
| Freshness | 4 |
| Case insensitivity | 5 |
| Inexact matching | 6 |
| Stemming | 7 |
| Support for scoped search | 8 |
| Support for advanced search | 9 |
| Support for phrase searching | 10 |
| Support for boolean search | 11 |
| Sort by relevance | 12 |
| Multilingual support | 13 |
| Selectable number of displayed results | 14 |
| Support for refinement of search results | 15 |
| Support for query reformulation | 16 |
| Use of metadata | 17 |
| Accessibility for disabled persons | 18 |
| Search box on home page | 19 |
| Large search box | 20 |
| Search button | 21 |
| Search function on every page | 22 |
| Search box on every page | 23 |
| Scoped search explained | 24 |
| Discreet Advanced Search | 25 |
| Context of search results | 26 |
| Relevance scores hidden | 27 |
| Matching keywords highlighted | 28 |
| Visible "no results" message | 29 |
| Search function on a search results page | 30 |
| Search box on a search results page | 31 |

Table 4 - Labeled search quality features

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | | | | | |
|---------------------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|--|--|--|--|--|
| quirinale.it | x | x | x | | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| palazzochigi.it | x | x | x | x | x | x | | x | x | x | | x | | | | | | | x | | x | x | | x | x | | | | x | x | x | | | | | | |
| camera.it | x | | x | | x | x | | x | x | | | x | | | | | | | x | x | x | x | | x | x | | | | x | x | x | | | | | | |
| mincomes.it | | x | x | | x | x | | | | | x | x | | | | | | | | | x | x | | | | | | | | | | | | | | | |
| comunicazioni.it | x | x | x | x | x | | x | | | | | x | x | | | | | | | | x | x | x | | | | x | x | | | | | | | | | |
| esteri.it | x | x | x | | x | x | | | | | | | x | | | | | | | x | x | x | | | | | | | | | | | | | | | |
| finanze.it | x | x | x | x | | x | | x | x | x | x | x | | | | | | | x | x | x | | | | | x | x | | | | | | | | | | |
| funzionepubblica.it | | | x | | x | | | | | | x | x | x | | | | | | | | x | x | x | | | | | | | | | | | | | | |
| giustizia.it | | | x | | x | x | | | | | x | | | | | | | | | | x | x | | | | | x | x | x | x | x | | | | | | |
| minlavoro.it | x | x | x | | x | x | | | | | x | x | | | | | | | | x | | | x | x | | | x | x | | | | | | | | | |
| llpp.it | x | x | x | | x | x | | | | | x | x | | | | | | | | | x | x | | | | | | | | | | | | | | | |
| istruzione.it | | | | | x | | | x | x | x | | | | | | | | | | x | | x | x | x | | | | | | | | | | | | | |
| murst.it | x | x | x | x | x | | | | | | | x | x | | | | | | | x | | | | | | | | | | | | | | | | | |
| sanita.it | x | | x | x | x | | | | | | | x | | | | | | | | x | x | x | | | | | x | x | | | | | | | | | |
| trasportinavigazione.it | x | | x | x | x | x | | | | | | x | | | | | | | | | | | | | | | x | | | | | | | | | | |
| regione.abruzzo.it | x | | x | x | x | | | | | | | | | | | | | x | | | x | x | x | x | | | | | | | | | | | | | |
| regione.emilia-romagna.it | x | x | | x | x | x | x | | | | x | x | | | | | | | | x | | x | x | x | | x | x | x | | | | | | | | | |
| regione.fvg.it | | | x | | x | | | x | x | | | | | | | | | | | x | x | | x | | | | | | | | | | | | | | |
| regione.liguria.it | x | x | x | x | x | | | | | | | x | x | | | | | | | | | | | | | | | | | | | | | | | | |
| regione.lombardia.it | x | x | x | x | | x | | | | x | | | | | | | | | | x | | x | x | x | | | | | | | | | | | | | |
| regione.piemonte.it | x | x | x | x | x | | | | | | | | x | | | | | | | | | | | | | | | | | | | | | | | | |
| regione.taa.it | x | | | | | | | x | | | x | x | | x | | | | | | | x | x | x | x | | | | | | | | | | | | | |
| regione.umbria.it | x | x | x | x | x | | | | | | x | x | | | | | | | | | x | x | x | x | | | | | | | | | | | | | |
| regione.sardegna.it | | x | x | | x | | | | | x | | x | x | | | | | | | x | | x | x | | | | | | | | | | | | | | |
| regione.sicilia.it | | | x | | x | | | | | | | | x | | | | | | | | x | | x | | | | | | | | | | | | | | |
| regione.vda.it | | | x | x | | | | x | x | x | | | | x | | | | | | | x | x | x | x | | | | | | | | | | | | | |
| regione.veneto.it | x | | x | | x | x | | | | | | | x | | | | | | | | x | x | x | | | | | | | | | | | | | | |
| comune.ancona.it | x | | x | x | x | | | | | | | x | x | | | | | | | x | | | | | | | | | | | | | | | | | |
| comune.asti.it | | x | x | x | x | | | | | | | | x | x | | | | | | | x | x | | | | | | | | | | | | | | | |
| comune.bergamo.it | x | | x | | x | | | | | | | | | | | | | | | | x | x | | | | | | | | | | | | | | | |
| comune.bologna.it | x | | x | x | x | x | | x | x | | | x | x | | | | | | | | x | | | | | | | | | | | | | | | | |
| comune.bolzano.it | | | x | x | x | x | | | | | | | | | | | | | | | x | x | x | x | | | | | | | | | | | | | |
| comune.cosenza.it | x | | x | | x | | | x | | | | | | | | | | | | | x | | x | x | | | | | | | | | | | | | |
| comune.ferrara.it | | | x | | x | x | x | | | | x | | | | | | | | | | x | x | x | x | | | | | | | | | | | | | |
| comune.firenze.it | x | x | x | x | x | x | | | | | x | x | | x | x | x | | | | | x | x | x | x | | | | | | | | | | | | | |
| comune.genova.it | x | | x | x | x | | | | | | | | | | | | | | | | x | | x | | | | | | | | | | | | | | |
| comune.mantova.it | x | | x | x | x | | | | | | | | | | | | | | | | | | x | x | | | | | | | | | | | | | |
| comune.milano.it | x | x | x | | x | x | | x | | | | | | x | | | | | | | | | x | x | | | | | | | | | | | | | |
| wcomune.modena.it | x | | x | x | x | | | | | | | x | x | | | | | | | | x | | | | | | | | | | | | | | | | |
| comune.palermo.it | x | x | x | | x | x | | | | | | | x | | | | | | | | | | x | x | | | | | | | | | | | | | |
| comune.piacenza.it | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.pisa.it | x | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.pistoia.it | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.prato.it | x | x | x | | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.reggio-calabria.it | x | | | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| municipio.re.it | | | | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.roma.it | | x | x | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.sassari.it | x | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.torino.it | x | x | | | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.trento.it | x | | x | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.vicenza.it | x | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.venezia.it | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| comune.udine.it | x | | x | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| agcom.it | x | | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| aipa.it | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| autorita.energia.it | x | | x | | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| consob.it | | | | | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| eurispes.com | x | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| garanteprivacy.it | x | | | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| inps.it | x | x | x | | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| istat.it | x | | x | | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| poste.it | x | | x | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Table 5 - Site-feature relation

It should be noted that, using this methodology, for some features we were truly evaluating their quality (e.g., is the search engine fast?, does it have a good coverage?, etc.), whereas for other features we only ascertained that they were made available to users, without investigating their quality (e.g., how good was the document ranking produced in response to a user query?). Moreover, it may be the case that the possession of high-quality features will not automatically result in performance improvement, as, for instance, with some advanced search capabilities which are in fact neglected by most users.

Table 6 shows the number and percentage of sites which comply with each feature. The most unusual features (with less than 25% of complying sites) are underscored, the most usual features (with more than 75% of complying sites) are bold. The majority of the most unusual features belongs to the search effectiveness dimension (i.e., 9, 13, 14, 15, 16, 17) - in particular, to the set of features which should support the interaction with the user. One feature (i.e., stemming) refers to search effectiveness and four features (18, 24, 25, and 32) to search interface. By contrast, the most usual features are concentrated in the search efficiency group (i.e., speed and no broken links, with all search efficiency features having more than half of complying sites) and in the search interface group (21, 29, 30), with only one feature (i.e., case insensitivity) being related to search effectiveness.

The results shown in Table 6 are useful to consider the behavior of single features, but they do not show the distribution of features in the sites under examination.

| % Sites | #Sites | Features |
|-----------|-----------|-----------|
| 77 | 48 | 1 |
| 53 | 33 | 2 |
| 87 | 54 | 3 |
| 53 | 33 | 4 |
| 94 | 58 | 5 |
| 45 | 28 | 6 |
| <u>5</u> | <u>3</u> | <u>7</u> |
| 29 | 18 | 8 |
| 19 | 12 | 9 |
| 27 | 17 | 10 |
| 53 | 33 | 11 |
| 53 | 33 | 12 |
| <u>2</u> | <u>1</u> | <u>13</u> |
| 11 | 7 | 14 |
| <u>5</u> | <u>3</u> | <u>15</u> |
| <u>18</u> | <u>11</u> | <u>16</u> |
| <u>15</u> | <u>9</u> | <u>17</u> |
| 0 | 0 | 18 |
| 65 | 40 | 19 |
| 50 | 31 | 20 |
| 94 | 58 | 21 |
| 65 | 40 | 22 |
| 32 | 20 | 23 |
| <u>15</u> | <u>9</u> | <u>24</u> |
| 19 | 12 | 25 |
| 76 | 47 | 26 |
| 55 | 34 | 27 |
| 31 | 19 | 28 |
| 90 | 56 | 29 |
| 92 | 57 | 30 |
| 65 | 40 | 31 |
| <u>11</u> | <u>7</u> | <u>32</u> |

Table 6 - Number and percentage of sites per feature.

Figure 1 depicts how the number of sites which possess at least a given number of search quality features decreases as the number of features is increased. The figure clearly shows that the overall search quality of the sites examined was modest. In particular, while almost all sites had at least 10 features, the number of complying sites decreased sharply when passing from 11 to 20 features. For 17 features, which is just more than 50% of the features considered for

evaluation, the number of sites reduced to 11 (16 % of the sites). Only two sites had twenty or more features and no site had at least 25 features.

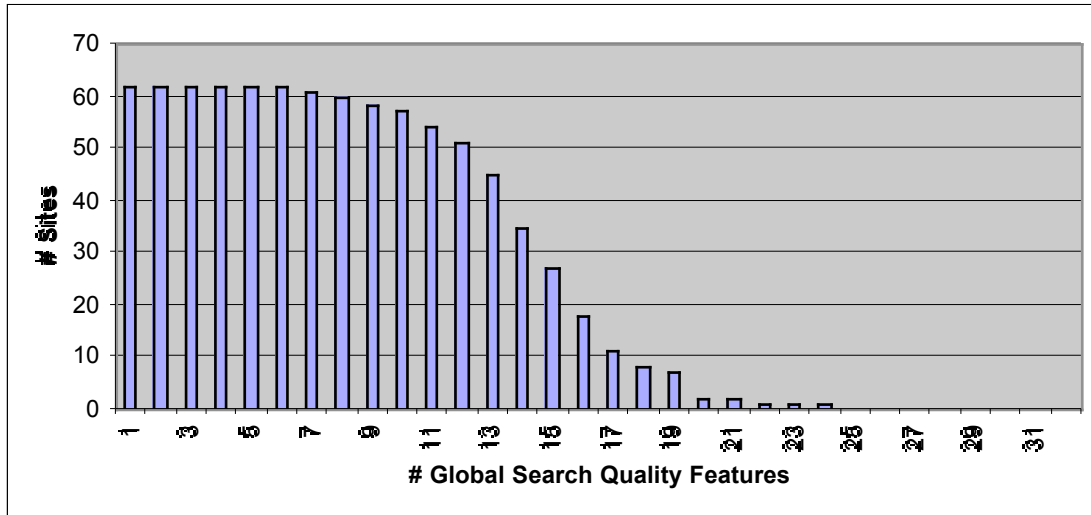


Figure 1 - Number of complying sites versus number of search quality features

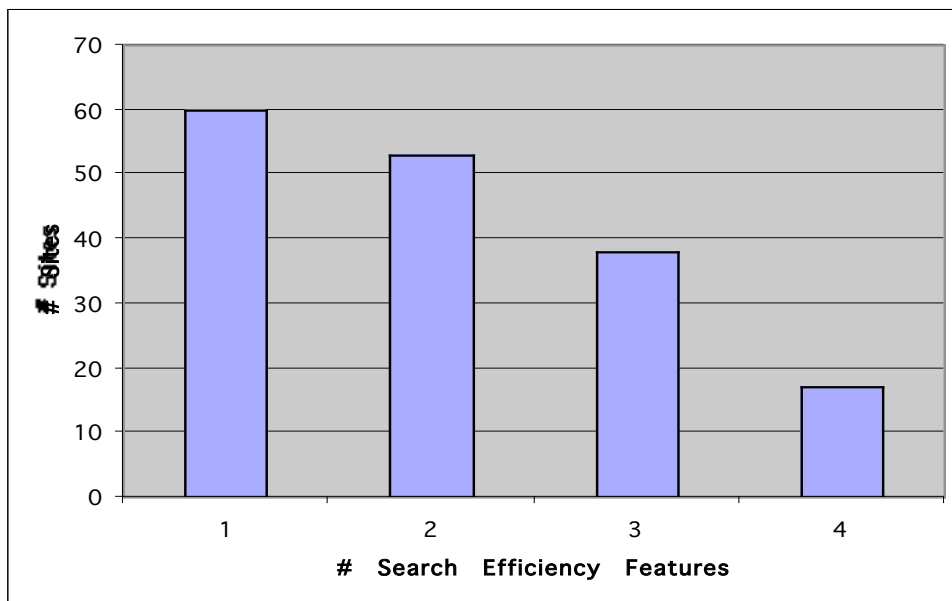


Figure 2 - Number of complying sites versus number of search efficiency features.

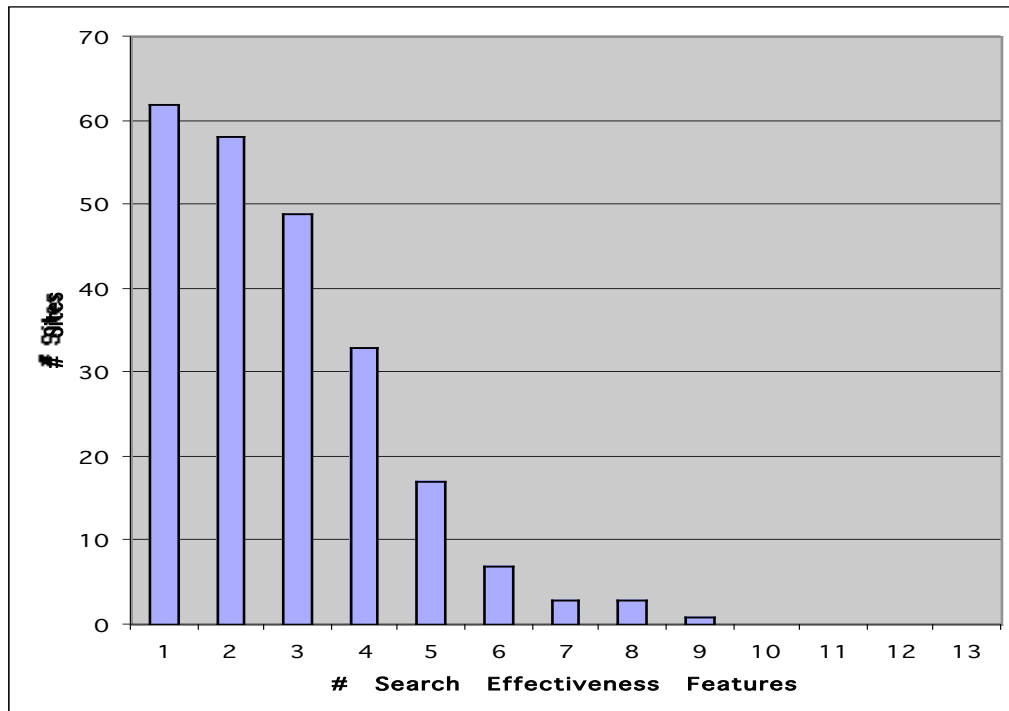


Figure 3 - Number of complying sites versus number of search effectiveness features

As we are interested in evaluating the behavior of the examined sites with respect to the main search quality dimensions, in the following we show the analogous of Figure 1 for each quality dimension.

The quality of search efficiency, as illustrated in Figure 2, is more favourable than the general picture. About 50 sites were found to have at least two features and nearly 40 sites possessed three of the four features considered for evaluation. Search speed, as already pointed out, was pretty good while search maintenance was slightly less satisfying with both high coverage and freshness scoring about 50% of sites.

Figure 3 shows the search effectiveness quality results. The number of complying sites falls sharply after a few features, reducing to only 10% of sites for less than 50% of features. Clearly, the low quality of search effectiveness was heavily influenced by the lack of features supporting the interaction with users. In fact, this is one of the most visible and consistent findings of our study.

Finally, in Figure 4 we report the search interface results. The quality was better than search effectiveness but it was still unsatisfying. Less than 50% of sites had more than 50% of

features, and some features, as shown in Table 6, were very poorly represented.

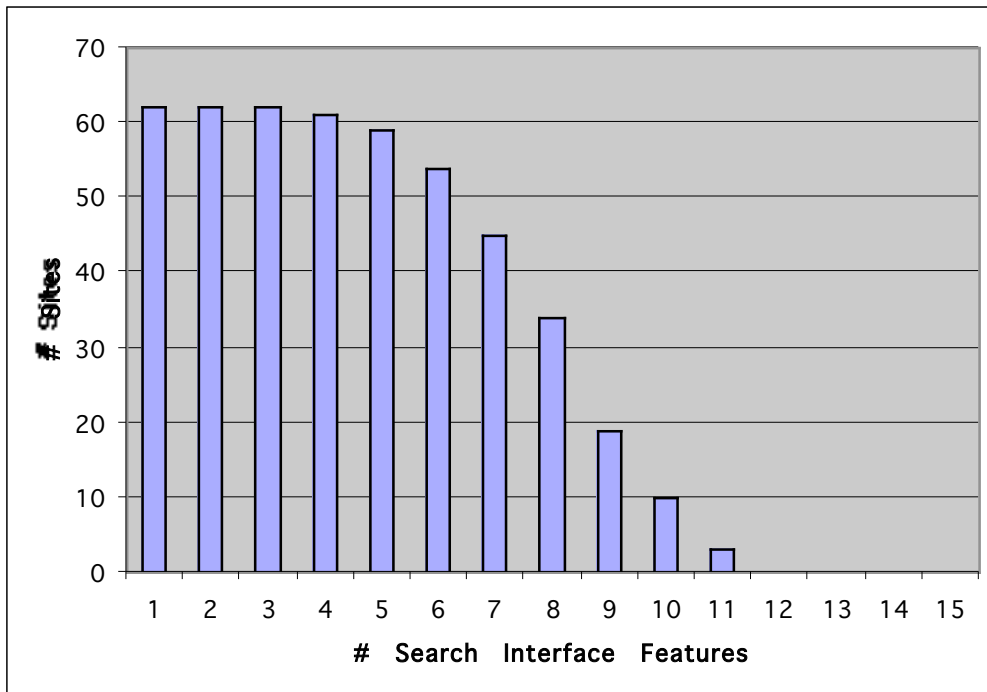


Figure 4 - Number of complying sites versus number of search interface features

CONCLUSIONS

We presented a comprehensive methodology for evaluating the quality of a site's search and applied it to a sample of the Italian Public Administration Websites.

The main conclusions of our evaluation study are the following.

- Half of the sites under consideration were not equipped with an internal search engine.
- The overall search quality of the sites with an internal search engine was low.
- The quality of search efficiency was fairly acceptable, with good response times and slightly less satisfying search maintenance.
- The quality of search effectiveness and search interface was unsatisfying, with a serious shortage of features supporting the interaction between users and system.

Finally, we would like to emphasize that due to the difficulty of accessing the source code of

the deployed systems, their ranking effectiveness for single-query searches was not taken into account in our evaluation. This aspect complements the others very well and can now be reliably assessed using electronically available test collections, provided that the search systems are made available for testing.

Acknowledgments

We would like to thank an anonymous reviewer for his useful comments. This research was done as a part of the project "InfoWeb", which aims at assessing and improving the quality of the Italian Public Administration Websites. Infoweb is conducted by Fondazione Ugo Bordoni and ISCTI.

REFERENCES

- [Berenci2000] Berenci, E., Carpineto, C., Giannini, V, and Mizzaro S.. Effectiveness of keyword-based display and selection of retrieval results for interactive searches. *International Journal on Digital Libraries*, 3, 3, 2000, pp. 249-260.
- [Carpineto2001] Carpineto, C., De Mori, R., Romano, G., Bigi, B. An information theoretic approach to automatic query expansion, *ACM Transactions on Information Systems*, 19, 1, 2001, pp. 1-27.
- [Hauptmann97] Hauptmann, A., and Witbrock, M. Informedia: News-on-demand multimedia information acquisition and retrieval. In *Intelligent Multimedia Information Retrieval*, M. Maybury (Ed.), AAAI Press, Menlo Park, CA, 1997.
- [Hearst1999] Hearst, M. User interfaces and visualization. In R. Baeza-Yates and B. Ribeiro-Neto (Eds.), *Modern information retrieval*, pp. 257-322, ACM Press, New York, 1999.
- [Henzinger2000] Henzinger, M. Link analysis in web information retrieval. *Bulletin of the Technical Committee on Data Engineering*, IEEE Computer Society, 2000, pp. 3-8.
- [Huberman1998] Huberman, B., Pirolli, P., Pitkow, J., and Lukose, R. Strong regularities in world wide web surfing, *Science*, 280, 1998, pp. 95-97.
- [Kobayashi2000] Kobayashi, M, and Takeda, K. Information retrieval on the Web. *ACM Computing Surveys*, 32, 2 , 2000, pp. 144-173.
- [Lawrence 1998] Lawrence, S. and Giles, C. Searching the world wide web. *Science*, 280, 1998, pp. 98-100.

- [Luk2000] Luk, R., Chan, A., Dillon, T., and Leong, H. A survey of search engines for XML documents. In Proceedings of the ACM SIGIR Workshop on XML and Information Retrieval, Athens, Greece, July 2000.
- [Nielsen1997] Nielsen, J. Search and you *may* find. Jakob Nielsen's Alertbox, July 15, 1997, <http://www.useit.com/alertbox/9707b.html>.
- [Nielsen 2001] Nielsen, J. Search: Visible and Simple. Jakob Nielsen's Alertbox, May 13, 2001, <http://www.useit.com/alertbox/20010513.html>.
- [Nielsen2000] Nielsen, J., Molich, R., Snyder, C., and Farrel, S. Search: 29 Design Guidelines for Usable Search, 2000, <http://www.nngroup.com/reports/ecommerce/search.html>
- [Oard1999] Oard, D., Peters, C., Ruiz, M., Frederking, R., Klavans, J., and Sheridan, P. Multilingual information discovery and access (MIDAS): A joint ACM DL '99 / ACM SIGIR '99 workshop. D-Lib Magazine, October 1999, <http://www.dlib.org>
- [Voorhees2000] Voorhees, E., and Harman, D. Overview of the Eighth Text Retrieval Conference (TRC-8). In Proceedings of the 8th Text REtrieval Conference (TREC-8), Gaithersburg, MD, 2000, NIST Special Publication 500-246, pp. 1-24.