



Alta Vista vs. Lycos



The Great Search Engine Debate - Tracey Stanley compares Alta Vista and Lycos (comparison done at the end of February/start of March 1996).

Introduction

Searching for information on the World Wide Web (WWW) can often be a long and tedious task. The WWW has grown phenomenally since its origin as a small-scale resource for sharing information and finding the information you need amongst this huge collection of resources can be very difficult without effective tools.

As the WWW has grown it has become necessary to provide a quick and easy method of rapidly searching webspace; and search tools - often known as search engines - have been developed which can perform this activity. Search engines provide a front end to a database of indexed WWW resources, into which search keywords can be typed.

However, the number of search engines available on the WWW has also grown quickly over a very recent period, and this has posed new problems for WWW users. There are now a bewildering variety of search engines available - each offering different features and interfaces. Many are linked to sizeable catalogues of WWW resources, and some claim to offer a comprehensive index of the entire WWW.

Given the problem with choosing a search engine from the range of tools on offer; this article concentrates on two particular tools - Lycos (<http://www.lycos.com>) and Alta Vista (<http://www.altavista.digital.com>). These two search engines have been chosen because they offer two of the largest databases of WWW documents - at the last count Lycos claimed to have indexed 10 million WWW documents, and Alta Vista claims to have a database of over 21 million documents. They also offer a number of fairly sophisticated searching features. This article will attempt to compare the performance of these two tools in order to provide some recommendations for their use.

Which has the best interface?

	Alta Vista	Lycos
Ability to restrict a search by date	Yes	No
Relevancy Ranking	No	Yes
Adjacency indicators	No	Yes
Ability to refine a search	Yes	Yes
"Help" documents	Yes	Yes
Simple keyword searching	Yes	Yes
Abstracts provided of "hits"	Yes	Yes

WWW interfaces to search engines generally consist of a form which appears on a WWW page. Keywords can be typed into the form and a button is provided which can be clicked on with a mouse in

order to activate the search. Other features such as small menus for selecting Boolean operators may also be present.

Both of these search engines effectively have two interfaces - one for simple keyword searching and another for more advanced queries using Boolean operators. Simple keyword search interfaces are located on the home page of each search engine, so these are the first interfaces that the user sees, and many users tend to use these by default rather than exploring the other options available. These interfaces generally provide a fast and easy to use tool for very simple WWW searching, but their use may be problematic given the size of the WWW and the generally diverse nature of material available.

Alta Vista offers its Advanced Query mode under its clickable logo at the top of the home page. Lycos offers its advanced search form under a hypertext link from the home page called Enhance your search. For the purpose of this article I've concentrated largely on features offered by the advanced searching options of these tools, as I feel that the extended features they offer are of considerably more use than simple keyword tools in searching a large and disparate database like the WWW.

Figure 1, below, displays the Alta Vista Advanced Query Interface.

Alta Vista

Simple Query Advanced Query Purchase Help

Search the Web and Display the Results in Standard Form

Selection Criteria: Use only Advanced Query Syntax with AND, OR, NOT and NEAR. Simple Query Syntax will not work!

Results Ranking Criteria: documents containing these words will be listed first. If left blank, the matching documents will not be sorted.

Start date: End date: e.g. 21/Mar/96

Alta Vista provide a facility to search both the WWW and Usenet News from their interface, and these options can be selected by clicking on the 'down' arrow to select from the Search menu. Another menu gives you the option of choosing how you would like your results displayed. Standard Form gives a short abstract of each document retrieved. Detailed Form can be used to display information on the publication dates of documents. On-line help in using the search interface can be accessed by clicking on the hypertext links - these point to a standard help file. A query can be typed into the Selection Criteria box, and relevancy ranking can be activated by typing keywords which have a high priority in the query into the Results Ranking Criteria box. Boxes are also provided which enable you to set a date threshold for your search. Once a query has been formulated the search can be initiated by clicking on the Submit Advanced Query button.

Figure 2, below, displays the Lycos enhanced search interface.

Lycos Search

Query:

Search Options:

Display Options:

- [Search language help](#)
- [Formless Interface](#)

Lycos provides a small query box into which search keywords can be typed. The Search Options menu enables the user to choose and and or Boolean operators (the not operator can be implemented by placing a dash (-) in front of a keyword). Various levels of matching can also be chosen - a loose match on the keyword computer would also retrieve the plural computers. The user can also determine the format of the results they retrieve, so that Standard Results will give a title and brief abstract of a document whereas Detailed Results will give information on indexing keywords and date of publication. Help is provided through a hypertext link to Search Language Help. Once a search has been constructed it can be initiated by clicking on the Search button.

Lycos provides a feature known as relevancy and adjacency ranking; and this is implemented when it sorts your search hits. Documents which contain a high incidence of your particular choice of keywords, and a high incidence of your keywords appearing close together within a document will score higher than those which don't and so appear higher up your list of hits. Alta Vista only performs relevancy ranking on your hits if you specifically ask it to by specifying keywords in the Results Ranking Criteria box.

WWW search engines generally seem to go through a process of more or less constant upgrade to their interfaces, and this can be very annoying for users who have spent time becoming familiar with an interface only to find that it has changed when they next want to use it. The Lycos interface has gone through numerous incarnations over the last year or so, and it seems likely that Alta Vista will also follow this pattern. The Alta Vista interface is currently a test version, and upgrades will be made once the developers have evaluated user responses.

Both Alta Vista and Lycos provide the ability to refine your search once the results of an initial search have been posted on screen. Alta Vista provides the full search form at the top of a list of hits so that a user can immediately refine a search if necessary. This form also retains the query structure of the previous search so that it can be quickly modified. Lycos, however, defaults automatically to its simple search form which means that a number of the enhanced features are lost.

Which has the best query language?

	Alta Vista	Lycos
Phrase searching	Yes	No
Full Boolean searching	Yes	No
Wildcard	Yes	No
Ability to restrict search to a particular field (eg: title)	Yes	No

Alta Vista supports a number of options which Lycos doesn't have, and this increases its functionality greatly. Phrase searching can be implemented by placing quotation marks around a phrase to ensure that Alta Vista only retrieves documents which contain the keywords in direct proximity to each other.

Its also possible to use a wildcard when searching Alta Vista to ensure that a number of possible variations on a word are included in search results. The wildcard symbol is *.

Alta Vista also enables a user to restrict their search to particular fields of a document. The fields available are the title of a document, the URL, the host (a page's server) and the links contained within a document. These features can be implemented by typing -for example - title: directly in front of a keyword.

Alta Vista also supports full Boolean searching, and it is possible to use and, or, not and near to expand or constrict a search. Parentheses can also be used to control nesting.

Which gives the best results?

The best way to illustrate the features of these two tool, and their differences, is to conduct a few searches and measure how they perform against each other.

To test the capabilities of the two search engines I performed a number of searches, aimed at testing how features such as phrase searching and use of wildcards can actually influence the results of a search. The aim was, wherever possible, to compare like with like, so I performed parallel searches on the same days in both Lycos and Alta Vista using the different features they offer, and compared my results.

The first search I performed was to search for information on the *Communications Decency Act*. As I anticipated that there would be a lot of material on this particular topic I decided to narrow my search by restricting it to sources which contained the phrase 'freedom of information'.

I first performed the search using Alta Vista's Advanced Query tool. This enabled me to perform a phrase search by typing my query as "*Communications Decency Act*" and "*freedom of information*".

This search returned 173 hits. I then decided to further refine my query by introducing a date restriction, and searched again, limiting the date to material published between 1 January 96 and 23 February 96 (the date on which I performed the search). This gave me 6 documents. All of this material looked relevant to my subject; there were documents from the Electronic Freedom Foundation on 'censorship and free expression' and a number of resources devoted directly to the discussion of the *Communications Decency Act*. It was interesting to note, however, that all of the documents listed had been published in January 1996, and there was no material available for February. I felt this was incongruous given that one of the key features of the WWW is that material is often available which is extremely current and up to date. At the time of performing my search the *Communications Decency Act* was a hot topic on many American newsgroups and mailing lists, and I felt that there should be some very recent material available on the WWW.

This led me to question how up to date the Alta Vista database actually is. Alta Vista don't provide any information in their 'help' files about how often the database is updated with new material.

I then performed a search on Lycos on the subject of the *Communications Decency Act*. I used the 'enhanced search' option in Lycos and typed the keywords *communications decency act freedom information*, setting the search options menu to 5 words to enable Boolean and. Lycos doesn't enable

phrase searching so I was unable to specify phrases in my query. Lycos also ignores word of less than three characters in length, such as of, so I had to leave this word out of my query.

Lycos wasn't able to find any documents which matched my query so I broadened my search and tried again with just the keywords *communications decency act*. Lycos found 94 documents, and displayed the first ten of these.

All of these 10 documents contained my three keywords, and one document - entitled 'Stop the Communications Decency Act' - was dated 10th February 96.

So in this case, Alta Vista was able to retrieve many more documents than Lycos and it was extremely useful to be able to qualify my search using phrases and date restrictions. However, Lycos did find a document which was more up to date than anything which appeared to be in the Alta Vista database at the time of writing.

My second search was on the subject of *dyslexia or special needs education*. I first performed this search in Alta Vista, using the Advanced Query option, and I decided to test the Boolean searching capabilities by including parentheses in my search. My search was structured as follows: (*dyslexia or "special needs"*) and *education*. This search retrieved over 40000 documents - obviously far too much information to wade through. The first hit which Alta Vista posted on the screen was an advertisement for an American gift shop which contained the line 'we provide gifts selection for clients who have special needs'! I concluded from this that I would need to further refine my search in order to retrieve more useful information. I decided to use the title tag to qualify my search, and I structured it as follows: (*title:dyslexia or title:"special needs"*) and *education*. This restricted the information retrieved to only those documents which contained the words *dyslexia or special needs* in their titles, rather than anywhere in the entire document.

This second search retrieved 183 documents, and all of my first 10 hits were relevant to my query. However, I again had the problem that all of my first 10 hits were dated from before January 1996, so I would need to refine my search again to discover very up to date material.

I then performed a search on this topic using Lycos. The keywords *dyslexia special needs education* were used and I set the search option to match 2 terms in order to implement Boolean or; this gave me 294 hits. All of my first 10 results were relevant to my query, and five of these results were dated February 1996. The keywords *special needs, special education or dyslexia* appeared in the titles of all of the ten documents. Lycos doesn't support use of a title tag to restrict queries, although it has a much narrower indexing policy than Alta Vista so this may explain why the search results appear to be more efficient.

Conclusion

Criteria	The Winner
The most frequently updated database	Lycos
The best query language	Alta Vista
The best interface	A draw
The best search results	A draw

Overall, both of these tools can potentially be extremely useful in searching the WWW, and they both

provide a number of options which are invaluable in helping the user to retrieve relevant information. However, there are a number of issues which this article raises which need to be addressed before these tools can be wholeheartedly recommended.

Alta Vista seems to have a significant problem with updates to its database; it may be that their hardware and software is having difficulty with keeping up with the rapid growth of the WWW, and that database updates are lagging behind as a result. However, Lycos update their entire database every month, and new pages are added on a weekly basis so at present it seems to have the lead over Alta Vista as regards up to the minute information.

As Alta Vista has such a big database it is very important to qualify a search as much as possible in order to retrieve information which may be relevant to your query. As demonstrated in the searches I performed, it is often necessary to make use of features such as field searching in order to maximise the effectiveness and accuracy of your search. For the user who is unaware of these features it could be fairly difficult to get good results using Alta Vista.

In many cases single keyword searching isn't really effective when searching such a huge and diverse database as the WWW, and I find it worrying that both of these search engines seem to be promoting single keyword searches through the layout of their pages. Users are encouraged to use the basic search forms which appear on the home pages of Lycos and Alta Vista and it isn't really made clear that using the advanced or enhanced query options will result in more efficient searches. Lycos even hides away the route to its enhanced search form under a rather vague 'enhance your search' link.

It's also worth noting that the Lycos simple search form has the Boolean 'or' operator set as a default. This can be confusing for users who are trying to perform a search. For example, its quite common for a user to want to perform a search on a phrase such as *user education*, and it would seem logical for a user to type these keywords into the simple search form in order to retrieve information on this topic. However, the default setting results in a search on *user* or *education*. At the time of writing, this search would retrieve over 36000 documents from Lycos!

It's important to point out to users the drawbacks of using the simple search options, and try to encourage them to use the other options available in order to enhance their searches. However both of these tools provide help files which are very limited in their scope and users are often confused by the query syntax available. Neither of these tools can therefore offer an ideal solution to searching the WWW.

Acknowledgement

With thanks to Karen G. Schneider for information and discussion relating to this article.

Contents Page - Electronic Libraries Programme and Project Information
Search Ariadne - Editorial for this issue - News desk - Reviews - Front Cover

March 22nd 1996 - Comments can be emailed to Ariadne