

# 8 ALTAVISTA

AltaVista is the largest, fastest Internet search engine but it can also be the most frustrating to use. It is sometimes difficult to narrow down your results or get a result which is not dominated by US information, even with advanced search techniques. In fact it seems that the more “advanced” your search strategy, the worse it gets. This is sometimes a consequence of the time-out “feature” that AltaVista applies to all searches. AltaVista allocates the same amount of time to each search that is submitted. When that time is up it presents the searcher with the results that it has found up to that point. This means that you may see completely different results with an identical search strategy, even if it is run only a few minutes later. The more complex and sophisticated your search and the busier AltaVista is, the more likely it is to return incomplete results. On some occasions, it has been known to return a results list which bears no resemblance to the search strategy whatsoever.

## 8.1 Simple search

The default search screen is Spartan when it comes to options but there is a lot you can do to make your search more specific once you know the commands.

The “Simple search”, as AltaVista calls it, is the default. The title is misleading as one can build very complicated searches on this screen. As with any search engine, you simply type in your terms and press Search.

The screenshot shows the AltaVista search page with several callout boxes pointing to specific elements:

- Option for selecting language in which the Web pages have been written:** Points to the "any language" dropdown menu.
- Click here for the Advanced search screen for Boolean searching, or limiting a simple search by date:** Points to the "Advanced Text Search" link.
- Select to search Web pages, Images, Video, Audio:** Points to the radio buttons for "Web Pages", "Images", "Video", and "Audio".
- Enter your search terms here:** Points to the search input field.
- Set your Preferences Set up default search and display options.:** Points to the "Set your Preferences" link in the footer.

The interface includes a search bar with a "Search" button, a "Help" link, and various navigation links like "Shopping.com" and "Zip2.com". It also features sections for "ALTAVISTA CHANNELS", "USEFUL TOOLS", "CATEGORIES", "ALTAVISTA HIGHLIGHTS", "OUR SPONSORS", and "ALTAVISTA INFORMATION".

The standard advanced search symbols can be used: double quotes for phrases, plus sign to make a term mandatory, minus sign to exclude pages. In addition, you can search areas of a Web page by using a range of prefixes. These are all given in the summary sheet later but the most useful ones are:

- url: to search for a term anywhere in the URL e.g. url:daimler
- domain: to search in the *last* part of the domain name only. This is useful if you want to restrict your search to a country e.g. +domain:uk or to a type of organisation but where there is not a country code at the end of the domain. For example, +domain:gov will find US government Web sites e.g. www.epa.gov but not the UK government Web site at www.open.gov.uk. If you want to restrict your search to a particular country's government sites use the url: option, for example +url:gov.uk for UK government sites.
- host: To search pages from a particular web site e.g. host:widget.co.uk
- title: to search for words in the title of the page e.g. title:chocolate

## 8.2 Results

AltaVista presents you with a list of sites ranked by relevance. Unless you specify otherwise in the search strategy, AltaVista will match as many of your terms as possible but not necessarily all of them.

The screenshot shows the AltaVista search interface. At the top, the search bar contains the query '+chocolate per capita consumption +UK'. Below the search bar is a banner for '50% off! Conversion to BT Highway'. The search results are displayed below the banner. The first result is 'History of Per Capita Consumption of Chocolate' with a URL of '207.197.202.83/conschoc.html'. The second result is 'Per Capita Meat Consumption' with a URL of 'meat.tamu.edu/consum.html'. There are also advertisements for Amazon.com and Shopping.com.

The Ask Jeeves service is checked to see if any of its questions and answers matches all or part of your strategy.

Number of pages that AltaVista has found matching one or more of your terms (Only the top 200 can be displayed)

Although AltaVista often tells you that it has found some enormous number of documents that match your search, the ones at the bottom of the list will probably only mention one of your terms once.

**ALTAVISTA**

There is no way that this can be checked with an extremely long results list because AltaVista allows you to display only the first 200.

The way in which AltaVista ranks the search results varies from time to time. When this update was written, AltaVista attempted to find phrase matches first rather than do a simple “or” search, which it claims is the default. Thus the order in which you type in your search terms does matter. For example, these are the results we obtained when we carried out variations on a search on low fat food:

<u>Search Strategy</u>	<u>Number of web pages found</u>
<u>Default search screen</u>	
low fat food	1,745
“low fat food”	1,745
food low fat	652,745
low food fat	3,282,520
<u>Advanced search screen</u>	
low fat food	1,745
low OR fat OR food	13,219,681

Web pages that contain your terms or phrases in the title seem to appear higher in the list of results.

In addition to the standard set of results, there is an automatic link to the Ask Jeeves service which checks to see if any of its questions and answers matches all or part of your strategy.

A relatively new feature is the “Paid for links”. According to AltaVista, the "Paid Link Program represents a new highly relevant search feature and an advertising program rolled into one. Web masters will have the opportunity, through auction, to bid on getting their sites placed at the top of the search results." Paid for links are clearly marked and labelled, and surrounded by a box so that they can be distinguished from regular search results. For example, a search on Linux gives the following:

▶ **AltaVista found 3,515,170 Web pages for you.**

[linux](#) - List of near matches related to **linux** provided by RealNames.

AV Relevant Paid Links <a href="#">Learn More</a>	Books at Amazon.com
<p><b>A. <a href="#">WebTrends Corporation</a></b>                      The leader and pioneer in enterprise management and reporting for Internet-based systems, offering a comprehensive set of solutions that are integrated, scaleable, modular, and easy-to-use.                      URL: <a href="http://www.webtrends.com/redirect/webtrends.htm">http://www.webtrends.com/redirect/webtrends.htm</a></p>	<p><a href="#">Search: Linux</a>  <a href="#">Save up to 50%</a></p>
<p><b>B. <a href="#">linux</a></b>                      Searching for useful information? We bring you the best of the Web and save you time finding exactly what you need. It's only at ZineZone.com.                      URL: <a href="http://ads.i33.com/redirect/0030/001357d/0001/CCC/A15/01/">http://ads.i33.com/redirect/0030/001357d/0001/CCC/A15/01/</a></p>	<p><b>AltaVista Shopping</b>  <a href="#">Visit our Online Shopping Guide</a></p>

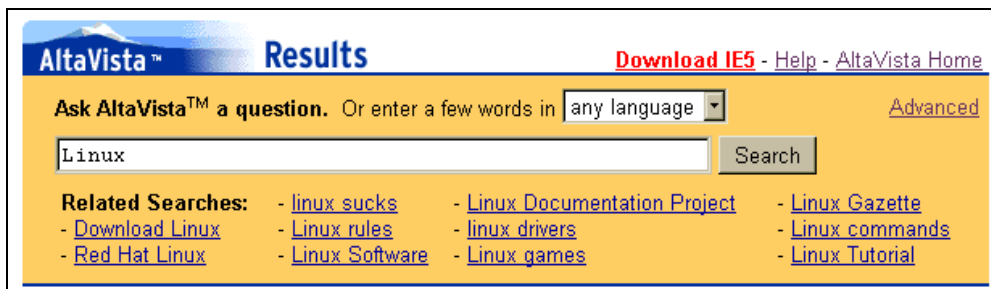
**1. [MkLinux: Linux for the Power Macintosh](#)**  
 What is MkLinux? Apple Computer is supporting a project with The Open Group - Research Institute to port Linux, a freely distributed UNIX-like operating...  
 URL: [www.mklinux.apple.com/](http://www.mklinux.apple.com/)  
 Last modified 17-Mar-99 - page size 4K - in English [ [Translate](#) ]

Two “Paid Links” for the search term Linux

Standard relevance ranked results start here

### 8.3 Refining your search

AltaVista no longer offers a Refine option (it was always erratic and not often very successful) but if you type in a very broad search strategy or use just one search term, for example Linux, it will suggest more specific search strategies called **Related Searches**.



The list does change and it appears that the suggestions are examples of searches that other people have recently carried out (there are often typographical errors).

### 8.4 Real Names

Real Names is a service which is supposed to be able to take you to the official site related to your search terms. The idea behind it is that a company or owner of a brand name can pay to register that name with Real Names and have that name directly linked to their Web page. For every search that you do in AltaVista, the Real Names database is also searched and either a direct match or related names found for you.



This only seems to work satisfactorily with very well known International names such as Disneyland. Real Names will come up straightaway with a link to the official Disneyland site. Furthermore, the service is dependent on the companies concerned registering their sites with Real Names.

## 8.5 Advanced Search

The Advanced Search has its own page and this is where you can either carry out a simple search combined with a date range, or use a Boolean search strategy, with or without a date restriction or range. You can make this screen your default logon screen by using [Set your Preferences](#), but this does require cookies to work.

The Boolean operators are:

- AND to find documents that contain all of your terms (you can use the ampersand (&) symbol instead of AND)
- OR to find documents that contain one or more of your terms
- AND NOT to exclude documents containing a particular term (you can use the exclamation mark (!) instead)
- NEAR (~) documents that contain your words within 10 words of each other and in either order

For complex searches you can use parentheses to ensure that the search is executed in the correct order, for example:

chocolate AND (production OR manufacture) AND (Austria OR Belgium OR Switzerland)

The screenshot shows the AltaVista search interface. At the top, it says 'AltaVista Results' with links for 'Download IE5', 'Help', and 'AltaVista Home'. Below this is a search bar with a dropdown menu set to 'any language' and a 'Simple Search' link. A search button is present. Below the search bar, there is a section for 'Enter boolean expression:' with a text area containing the query 'chocolate near (manufacture or production) and (Belgium or Austria or Switzerland)'. To the right of this is a 'Range of dates:' section with 'From:' and 'To:' input fields, and an example 'e.g.: 21/Mar/96'. There is a checkbox labeled 'Count documents matching the boolean expression.' Below the search area is an advertisement for 'manufacturing.net' with the tagline 'Industry's Voice' and a search box for the site. The search results section shows 'AltaVista found about 298 Web pages for you.' The first result is '1. Prefamac NV - Chocolate machines - Cooling tunnels - Conveyor belts' with a description, URL 'www.prefamac.com/', and a 'Translate' link. The second result is '2. Tea & Coffee Trade Journal On the Continent' with a description, URL 'teacofmag.com/9710continent.html', and a 'Translate' link. On the right side of the results, there are two advertisement boxes: 'Books at Amazon.com' with a search link for 'chocolate ne...' and 'Amazon.com Bestsellers', and 'AltaVista Shopping' with links for 'Visit our Online Shopping Guide' and 'Shop at Shopping.com' with a search link for 'chocolate ne...' and a 'FREE PC Contest Daily' link.

The AND operator should have the same effect as using the + symbol in the Simple Search screen. Thus:

+functional +foods +labelling +regulations

should, and does, give the same number of results (263) as:

functional AND foods AND labelling AND regulations

With some searches, though, the number of documents found does differ. No-one has yet been able to come up a satisfactory answer as to why this happens. But even when the number of documents found is the same, they are displayed in a different order. With the Simple Search, relevance ranking is used and some of the search terms are in the title of the page:

▶ **AltaVista found 263 Web pages.** [Family Filter is off.](#)  
[Take our one minute survey!](#)

1. **[Press Release - Functional foods make fools of the law](#)**  
 Functional foods make fools of the law. Food companies that make health-boosting claims for their products may be breaking the labelling regulations, ...  
**URL:** [www.gurney.co.uk/foodcomm/](http://www.gurney.co.uk/foodcomm/)  
 Last modified 29-Apr-96 - page size 3K - in English [ [Translate](#) ]
2. **[Food Today #6 - New impetus for Functional Foods](#)**  
 SPOTLIGHT. New impetus for Functional Foods. Diet-related health problems may be modified through the introduction of functional foods and ingredients...  
**URL:** [www.eufic.org/foodtoday/eng/ft6/spotlight.html](http://www.eufic.org/foodtoday/eng/ft6/spotlight.html)  
 Last modified 10-Jun-98 - page size 5K - in English [ [Translate](#) ]
3. **[The Nutraceuticals and Functional Foods Industry in Canada: Issues, Challenges](#)**  
 The Nutraceuticals and Functional Foods Industry in Canada: Issues, Challenges and Opportunities. The purpose of this document is to identify the issues...  
**URL:** [www.agr.ca/cb/min/nutraceuticalse.html](http://www.agr.ca/cb/min/nutraceuticalse.html)  
 Last modified 30-Apr-99 - page size 15K - in English [ [Translate](#) ]

Within the Advanced Search screen relevance ranking is not used:

▶ **AltaVista found about 263 Web pages.** [Family Filter is off.](#)  
[Take our one minute survey!](#)

1. **[FOOD SCIENCE AUSTRALIA: Research, technology, infrastructure at your service..](#)**  
 Food Science Australia. Research, technology, infrastructure... at your service .. but not on your balance sheet. left: microtesting salad products....  
**URL:** [srcmain.dfst.csiro.au/capabilities.htm](http://srcmain.dfst.csiro.au/capabilities.htm)  
 Last modified 28-Sep-98 - page size 11K - in English (Win-1252) [ [Translate](#) ]
2. **[Bioresco Bioresearch Management and Consulting Ltd.](#)**  
 Bioresco Bioresearch Management and Consulting Ltd. is a consulting company that advises a broad range of clients on scientific and regulatory issues...  
**URL:** [www.bioresco.ch/bio-e-1.html](http://www.bioresco.ch/bio-e-1.html)  
 Last modified 30-Nov-98 - page size 4K - in English [ [Translate](#) ]
3. **[Bioresco Bioresearch Management and Consulting Ltd.](#)**  
 Bioresco Bioresearch Management and Consulting Ltd. is a consulting company that advises a broad range of clients on scientific and regulatory...  
**URL:** [www.bioresco.ch/bio-projects.html](http://www.bioresco.ch/bio-projects.html)  
 Last modified 30-Nov-98 - page size 6K - in English [ [Translate](#) ]

## ALTAVISTA

In fact the top results from the Advanced Search were less useful than those found using the Simple Search. This is because the AND operator will look for words anywhere in the Web page and there is no relevance ranking involved. In this case a better approach would have been to use the NEAR operator:

```
(functional NEAR foods) AND (labelling NEAR regulations)
```

which gave 48 results, or even better still:

```
(functional foods) AND (labelling regulations)
```

which gave 12 results and is the same as the Simple Search +“functional foods” +“labelling regulations”. (Note that if you omit an operator in the Advanced Search screen AltaVista assumes that the words form a phrase, for example `functional` immediately followed by `foods`.)

## 8.6 Translation Service

Next to each of the entries in your results list you will find a [Translate](#) hypertext link. Clicking on this takes you to straight to <http://babelfish.altavista.com/>. This is the machine translation option that translates pages to or from English into French, German, Italian, Portuguese or Spanish. Please remember that it is *machine* translation so do not expect perfection and under no circumstances use the results as the finished article; pass it to a human translator first. It is a useful way, though, of getting a rough idea of the content of the page

You do not have to carry out a search before you can use the service. If you go direct to the Babelfish pages, you can enter directly the URL of a page for translation, or type or paste some text into the box.