



Internet Search Services

Version 2.04

Search Engine Interface Description

18 July, 2000

Copyright:
AltaVista Company
1825 S. Grant Street, Suite 410
San Mateo, California 94402-2648
<http://solutions.altavista.com>
Phone: +1(650) 295-2500
Fax: +1(650) 295-2532

COPYRIGHT INFORMATION

The information in this document is subject to change without notice and should not be construed as a commitment by AltaVista Company. AltaVista Company assumes no responsibility for any errors that may appear in this document.

The software described in this document is furnished under a license and may be used or copied only in accordance with the terms of such license. Restricted Rights: Use, duplication, or disclosure by the U.S. Government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013.

AltaVista® is a registered trademark and Smart is Beautiful and the AltaVista logo are trademarks of AltaVista Company

Digital Unix, Compaq Tru64 Unix and Alpha are trademarks of Compaq Computer Corporation.

Microsoft®, Windows®, and Windows NT® are registered trademarks of Microsoft Corporation in the United States and/or other countries.

Intel® is a registered trademark of Intel Corporation.

IBM is a registered trademark of International Business Machines Corporation.

Corel and WordPerfect are registered trademarks, and Corel is a trademark of Corel Corporation.

Lotus is a registered trademark and Domino is a trademark of Lotus Development Corporation.

Novell and NetWare are registered trademarks of Novell, Inc.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd. UNIX and XWindow System are registered trademarks of X/Open Company Ltd.

Apple™ and QuickTime™ are trademarks of Apple Computer, Inc., registered in the U.S. and other countries.

RealNames, the RealNames Logo and the Internet Keywords by RealNames logo are service marks of RealNames Corporation.

Sun, Sun Microsystems, Solaris, Java™, JavaServer Pages™ are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries.

All other trademarks and service marks are the property of their respective companies.

© AltaVista Company. All Rights Reserved.

Refer to Section 12 for information on contacting AltaVista.

Contents

1	Introduction.....	1
1.1	ISS Product Overview	1
1.2	ISS Release 2.0	2
2	ISS Architecture	3
3	ISS Input Format.....	4
3.1	Text Search Input Parameters	4
3.1.1	Customer ID Parameter — who.....	5
3.1.2	Output Format Parameter — dp.....	6
3.1.3	Family Friendly Filter Parameter — FFF.....	6
3.1.4	Search Type Parameter — stype	6
3.1.5	Simple/Advanced Search Selection Parameter — pg.....	6
3.1.6	Search String Parameter — q.....	6
3.1.7	Ranking String Parameter — r.....	9
3.1.8	Number of Results per Query Parameter — nbq	10
3.1.9	Results Ordinal Parameter — stq.....	10
3.1.10	Start Date Parameter — d0.....	10
3.1.11	End Date Parameter — d1	11
3.1.12	Language Selection Parameters — kl and KL.....	12
3.1.13	Encoding Selection Parameter — enc	13
3.1.14	Site Compression Parameter — sc.....	14
3.1.15	Site Identification Parameter — si.....	14
3.2	Multimedia Search Input Parameters.....	14
3.2.1	Customer ID Parameter — who.....	16
3.2.2	Output Format Parameter — dp.....	16
3.2.3	Family Friendly Filter Parameter — FFF.....	17
3.2.4	Search Type Parameter — stype	17
3.2.5	Search Selection Parameter — pg.....	17
3.2.6	Search String Parameter — q.....	17
3.2.7	Number of Results per Query Parameter — nbq.....	19
3.2.8	Results Ordinal Parameter — stq.....	19
3.2.9	Multimedia Search Type Options Parameter — mmdo.....	19
3.2.10	Object Identification Parameter — objid.....	20
3.2.11	Multimedia Object Type — objtype	20
3.2.12	Page Identification Parameter — pagefp	21
3.2.13	Audio Specific Parameters — (stype=saudio)	21
3.2.13.1	Audio Search Selection Parameter - audset.....	21
3.2.13.2	Audio File Type Parameters.....	21
3.2.13.3	Audio Source Parameters.....	22
3.2.14	Image Specific Parameters — (stype=simage).....	22
3.2.14.1	Image Search Selection Parameter - imgset.....	22
3.2.14.2	Image Characteristics Parameters.....	22
3.2.14.3	Image Source Parameters.....	23
3.2.15	Video Specific Parameters — (stype=svideo)	23
3.2.15.1	Video Search Selection Parameter - vidset.....	23
3.2.15.2	Video File Type Parameters.....	23
3.2.15.3	Video Source Parameters.....	24
4	ISS VAL Output Format	25
4.1	VAL Text Search Output	25
4.1.1	VAL Text Search Output Header	25
4.1.1.1	Related Search Output.....	27
4.1.1.2	RealNames Output.....	27
4.1.2	VAL Text Search Results Output.....	28
4.1.3	Example VAL Text Search without Related Search and without RealNames	28
4.1.4	Example VAL Text Search with Related Search and without RealNames.....	29

4.1.5	<i>Example VAL Text Search without Related Search and with RealNames</i>	29
4.1.6	<i>Example VAL Text Search with Related Search and with RealNames</i>	30
4.1.7	<i>Example VAL Text with &enc=utf8</i>	31
4.2	VAL Multimedia Search Output	31
4.2.1	<i>VAL Multimedia Search Output Header</i>	31
4.2.1.1	<i>Related Search Output</i>	31
4.2.2	<i>VAL Multimedia Search Results Output</i>	32
4.2.2.1	<i>MmType Definition</i>	32
4.2.2.2	<i>MmClass Values</i>	33
4.2.3	<i>Example VAL Multimedia Audio Search</i>	33
4.2.4	<i>Example VAL Multimedia Image Search without Related Search</i>	34
4.2.5	<i>Example VAL Multimedia Image Search with Related Search</i>	35
4.2.6	<i>Example VAL Multimedia Video Search</i>	36
5	ISS XML Output Format	37
5.1	XML Text Search Output	38
5.1.1	<i>XML Text Search Output Header</i>	38
5.1.1.1	<i>Related Search Output</i>	40
5.1.1.2	<i>RealNames Output</i>	40
5.1.2	<i>XML Text Search Results Output</i>	41
5.1.3	<i>XML Text Search DTD Grammar</i>	41
5.1.4	<i>Example XML Text Search without Related Search and without RealNames</i>	42
5.1.5	<i>Example XML Text Search with Related Search and without RealNames</i>	42
5.1.6	<i>Example XML Text Search without Related Search and with RealNames</i>	43
5.1.7	<i>Example XML Text Search with Related Search and with RealNames</i>	44
5.2	XML Multimedia Search Output	45
5.2.1	<i>XML Multimedia Search Output Header</i>	45
5.2.1.1	<i>Related Search Output</i>	45
5.2.2	<i>XML Multimedia Search Results Output</i>	45
5.2.2.1	<i>MmType Definition</i>	46
5.2.2.2	<i>MmClass Values</i>	47
5.2.3	<i>XML Multimedia DTD Grammar</i>	47
5.2.4	<i>Example XML Multimedia Audio Search in XML</i>	47
5.2.5	<i>Example XML Multimedia Image Search without Related Search</i>	48
5.2.6	<i>Example XML Multimedia Image Search with Related Search</i>	49
5.2.7	<i>Example XML Multimedia Video Search</i>	50
6	Complete DTD File	52
7	Example ISS Query Via TELNET to Port 80	54
8	Input Parameter Validity	56
9	Glossary	58
10	AltaVista Search String Syntax	63
11	URL Encoded Representation	65
12	Contact Information	66

Internet Search Services Revision History

Version	Date	Description
1.0	1996 – 1999	<ul style="list-style-type: none">• VAL product
2.0	13 March 2000	<ul style="list-style-type: none">• XML format• Multimedia Search• RealNames and Related Search available for Text Search• Enhanced security support
2.01	13 April 2000	<ul style="list-style-type: none">• RealNames and Related Search available for Multimedia Search• Modifications to Multimedia parameters
2.02	11 May 2000	<ul style="list-style-type: none">• DTD file clarifications
2.03	29 June 2000	<ul style="list-style-type: none">• Enhancements to Multimedia input and output• Added the KL parameter and clarified the kl parameter description about language selection• Added new SiteName and SiteID keys to both VAL and XML output• Updated DTD file
2.04	18 July, 2000	<ul style="list-style-type: none">• Document enhancements only

1 Introduction

AltaVista's Internet Search Services (ISS) program is designed to provide global search capabilities for Internet portals. ISS allows portal operators to provide their end users with access to AltaVista's world-class collection of images, audio clips, video clips, and general text search results. AltaVista search results are easily integrated into the look and feel of a web site, enabling portal operators to use Internet Search Services to improve the experience of their end users while retaining traffic on their site and generating additional revenues.

1.1 ISS Product Overview

The ISS product consists of two separately purchasable search functions: *Text Search* and *Multimedia Search*. The sections of this document that describe the ISS input format (Section 3) and the VAL and XML output formats (Sections 4 and 5) are divided into a separate subsection for each of these functions.

Text Search allows ISS customers to search AltaVista's extensive *index* of Web pages from around the world. Search engine indexes are generated automatically, based on the words and phrases that are found on Web pages. AltaVista indexes every word of every page found by its numerous *crawlers*. It remembers not just the words, but the order of the words, so users can search for phrases or entire sentences. The index attempts to include all Web pages that contain real and unduplicated information, without making value judgements about that information. The main crawler named *Scooter*, sends out thousands of HTTP requests simultaneously grabbing text, pulling it back, and sending it into the indexing machines so that text can be placed in the index. *Scooter* also has *cousins*, other crawlers that do specialized jobs to help keep the index current, such as checking for and removing dead links (pages that have been moved or deleted). In a typical day *Scooter* and its *cousins* visit over 10 million pages.

Multimedia Search allows ISS customers to search AltaVista's massive index of audio clips (MP3, WAV, streaming audio, and other audio), video clips (avi, mpeg, QuickTime™, streaming video, and other video), and images (photos, graphics, buttons and in color and in black and white). The AltaVista index features premium content from a wide variety of world-class sources. Through premium content partnerships, these multimedia indexes include information that is not normally available on the web. Premium content partners are subject to change. Currently the premium content partners include:

- ABC, Inc. (video) is available only by prior arrangement with AltaVista.
- CDnow Online, Inc. (audio and images)
- Corbis.com (images)
- EMusic.com, Inc. (audio)
- Epitonic.com (audio)
- FasTV, Inc. (video)
- Getty Images, Inc. (images)
- LAUNCH Media, Inc. (video)
- Merrill Lynch & Co., Inc. (video)
- MJUICE.com (audio)
- MSNBC.com (video)
- On the Scene Productions, Inc. (video)
- ON24 Inc. (audio and video)
- Riffage.com, Inc. (audio)
- Vidnet.com (video)

In addition to the two separate search functions, ISS customers may also separately purchase two search enhancements: *Related Search* and *RealNames Search*.

Related Search supplements *Text Search* and *Multimedia Image Search*. ISS will return a list of search terms similar to the search query just submitted. This list helps your users decide upon the search term that provides the most relevant results.

RealNames Search supplements *Text Search*. ISS provides a link to registered pages of products, brands, and companies. *RealNames* results help end users rapidly find the web pages that are most relevant to their search query.

1.2 ISS Release 2.0

The Internet Search Services program is the renamed, updated version of VAL (AltaVista's earlier Value-Added Link offering). AltaVista's Internet Search Services Version 2.0 and subsequent updates are backward compatible with VAL Version 1.0. New additions to the ISS program in the version 2.0 base release include several optional capabilities. A new XML output format supplements the former VAL or ASCII text format, and a new security system has also been implemented.

Three new additions to ISS in the Version 2.0 base release:

1. **XML:** The new XML output allows ISS customers to process ISS search results using a standard XML parser, such as Sun Microsystems' Project X XML Parser, Microsoft's XML Parser, or IBM's Xerces Java parser. Using ISS Version 2.0, customers can switch between VAL and XML formats for different searches. For example: XML output may be selected for a search using *Multimedia Search*, while results for *Text Search* may be returned in VAL format.

Note: The output format is determined by the customer's usage of the **dp** parameter. When *val* is specified within the **dp** parameter the output is in VAL format; conversely when *xml* is specified, the search results are returned in XML format.

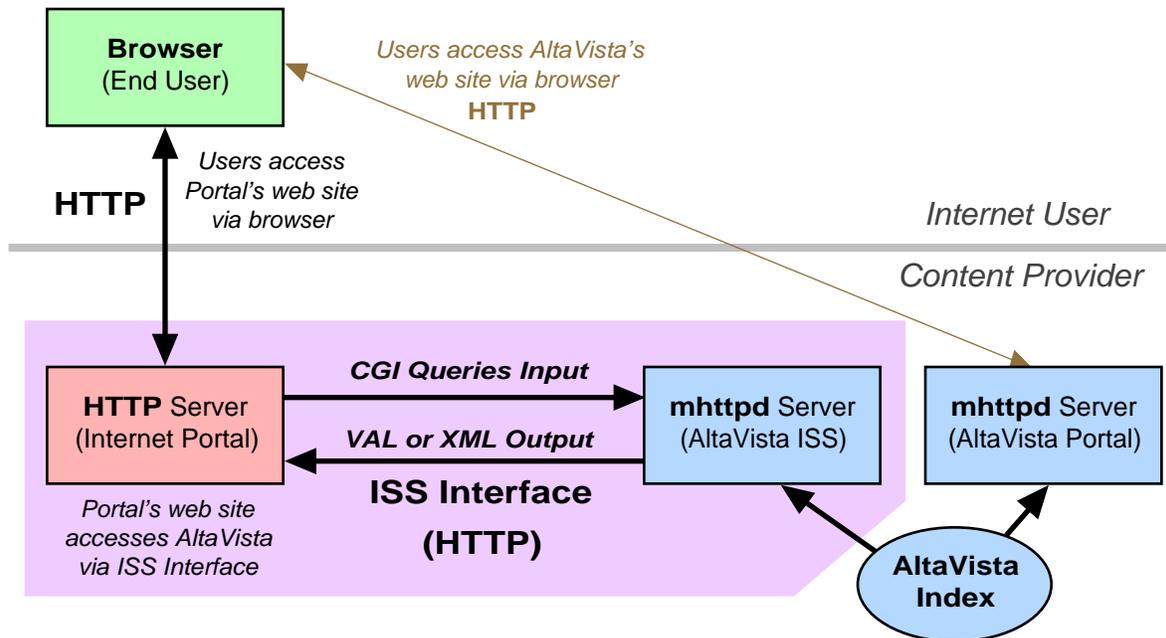
2. **Text Search and Multimedia Search:** Version 2.0 further customizes the search function with the separately purchasable additions of *Text Search* and *Multimedia Search* (image, audio, and video). *Text Search* and *Multimedia Search* can be used independently or in conjunction with the purchasable service, *Related Search*. *Text Search* can also be used with the purchase of *RealNames*. Any of these features can be used in any format, according to the **dp** parameter's specification of *val* or *xml*.
3. **Security Support:** has been implemented with the expansion of searches and the addition of XML to the ISS search interface. All ISS queries must now be authenticated. The **who** parameter provides service only to registered customers. For example: if *Text Search*, *RealNames*, and *Multimedia Search* are purchased but *Related Search* is not purchased, a *Text Search* query will return *RealNames* but not *Related Search* information. New security error messages have been added.

Refer to the Internet Search Services Revision History on page v to review any updates to the base 2.0 release.

Please refer to Section 12 for contact information on ISS products.

2 ISS Architecture

ISS architecture is implemented by providing a simple HTTP interface for the ISS customer to the normal HTTP server of the AltaVista search engine, **mhttpd**. Normally, **mhttpd** responds directly to end user browser HTTP requests over the Internet and serves as a web page interface to the search engine. In the ISS model, an ISS customer content-provider sets up its own web server (Internet portal) that interfaces to **mhttpd**. The figure below illustrates this architecture.



The ISS customer Internet portal must also accept queries from the end user browser. Typically these queries consist of CGI, JavaServer Pages™, Microsoft Active Server Pages, or Apache Modules scripts, one or more of which are supported by most commercial sites. The script then formulates the query according to the ISS interface specification, and forwards the query to the AltaVista HTTP server **mhttpd**. The results are then returned to the Internet portal over the ISS interface in an easily parsed HTTP protocol format. The content-provider's script then reformats the results and displays them to the user.

3 ISS Input Format

The ISS input format consists of an HTTP GET request containing a URL that addresses the AltaVista **mhttpd** server **partners.altavista.com**. The GET request should include the version specification. An example is provided below.

```
GET URL HTTP/1.0
```

This GET request will cause the results to be returned in a form compatible with HTTP version 1.0, and HTTP header information is returned. If the version specification is omitted, the **mhttpd** server assumes that the request is from an obsolete browser that cannot handle HTTP header information, and the HTTP header information is not returned.

The typical format of the URL (using CGI script) is as follows:

```
http://mhttpdhost/cgi-bin/query?querystring
```

- *mhttpdhost*: is the host where the **mhttpd** HTTP server is running. It may include a port component. For example: `partners.altavista.com:80`.
- *querystring*: is a string in CGI QUERY_STRING format, as described for CGI/1.1. This string has a number of *parameter=value* pairs, each separated by an ampersand (&), and can be in any order. These input parameter pairs are described below. The *querystring* is case sensitive and cannot contain any blanks or tabs. Any characters in the *querystring* that are not understood by the **mhttpd** server are simply ignored. Non-alphanumeric characters must be specially coded so that they are not interpreted by the CGI protocol. Refer to Section 11 for a URL encoded representation table.

An example of a URL containing a *querystring* that looks for web pages about baseball might be:

```
http://partners.altavista.com/cgi-  
bin/query?dp=xml&who=issid&pg=q&stype=stext&q=baseball
```

Note: A unique *issid* is given to each AltaVista ISS customer and must be provided in every query request.

As described in Section 1.1, the ISS product consists of two separately purchasable search functions: *Text Search* and *Multimedia Search*. The following subsections describe the input parameters for each of these functions. Some of the parameters are used by more than one function. Sometimes these parameters have the same value in different functions (the **dp** and **who** parameters for example), and sometimes the parameters have different values in different functions (the **stype** and **FFF** parameters for example). It is therefore important to use the description of the parameter that applies to the function being performed by the query. Since the defaults for optional parameters may change it is a good idea to always specify values for every parameter.

Note: All of the parameters and their values are case sensitive. Failure to match the required case will result in the parameter and its value being ignored.

3.1 Text Search Input Parameters

The following quick reference table provides an alphabetic list of the *Text Search* parameters, their values and ranges, whether they are required or are optional and their default value if appropriate. Within the table the user-supplied values are displayed in the *Italics* typeface, and values that must be specified exactly as shown are displayed in normal typeface. Values separated by the bar character (|) represent a list from which one value is to be chosen. Values within brackets ([]) are optional. Values within parentheses represent a range of values. Values within braces ({ }) are repeatable. A detailed description of each of the parameters follows this table.

Text Search Parameters Quick Reference				
Parameter	Description	Value (Range)	Required / Optional	Default Value
d0	Start Date	<i>dd%2FMmm%2Fyy[yy] dd%2Fmm%2Fyy[yy] dd</i>	Optional	oldest date
d1	End Date	<i>dd%2FMmm%2Fyy[yy] dd%2Fmm%2Fyy[yy] dd</i>	Optional	today
dp	Output Format	val xml	Required	
enc	Encoding Option	ascii big5 cp1250 cp1251 cp1252 cp1253 cp1254 cp1255 cp1256 cp1257 cp1258 eucjn eucjp euckr iso88591 iso88592 iso88593 iso88594 iso88595 iso88596 iso88597 iso88598 iso88599 iso885910 jis koi8r sjis utf8	Optional	iso88591
FFF	Family Friendly Filter	<i>(present or not present)</i>	Optional	off
kl	Single Language Selection	XX zh cs da nl en et fi fr de el he hu is it ja ko lv lt no pl pt ro ru es sv	Optional	XX
KL	Multiple Language Selection	<i>cc{cc}, cc=zh cs da nl en et fi fr de el he hu is it ja ko lv lt no pl pt ro ru es sv</i>	Optional	kl setting
nbq	Number of Results	<i>number (1 to 100)</i>	Optional	10
pg	Query Type	q aq	Optional	q
q	Query String	<i>CGI query string</i>	Required	
r	Advanced Search Ranking String	<i>CGI ranking string</i>	Optional	automatic for pg=q , random for pg=aq
sc	Site Compression	on off	Optional	on for pg=q , off for pg=aq
si	Site Identification Code	<i>number</i>	Optional	
stq	Ordinal Of First Results Entry	<i>number</i>	Optional	0
stype	Search Type	stext	Optional	stext
who	Customer Id Code	<i>issid</i>	Required	

The following example query illustrates the use of some of the *Text Search* input parameters. This query looks for English web pages that contain the phrase “Portuguese Water Dogs”.

```
http://partners.altavista.com/cgi-bin/query?pg=q&q=%22Portuguese+Water+Dogs%22&kl=en&stype=stext&dp=xml&who=issid
```

The following subsections describe each of the *Text Search* parameters in more detail.

3.1.1 Customer ID Parameter — who

The required **who** parameter identifies the particular AltaVista customer that is accessing the ISS link. This parameter is a necessary component of AltaVista’s ISS security.

Value	Description
<i>issid</i>	The <i>issid</i> is a unique token assigned to each particular ISS customer by AltaVista.

3.1.2 Output Format Parameter — *dp*

The required **dp** parameter indicates which ISS output format, VAL or XML, is to be selected for all content returned by **mhttpd**.

Value	Description
val	Requests VAL format (refer to Section 4.1).
xml	Requests XML format (refer to Section 5.1).

3.1.3 Family Friendly Filter Parameter — *FFF*

The optional **FFF** parameter controls the level of family friendly filtering enforced by the search engine. If used, the returned search results will be filtered for offensive *English* and only pages containing non-offensive *English* will be returned. If the parameter is *not* used, all matching results will be returned. Unlike other parameters that take the *parameter=value* pairs form, the **FFF** parameter consists of only its existence or non-existence in the *querystring*.

Note: If the family friendly filter is on, ISS results will not contain *Related Search* and *RealNames* items (if you have purchased these search options).

3.1.4 Search Type Parameter — *stype*

The optional **stype** parameter is used to select the type of web search sent to ISS: text pages, audio files, video files, or image files. The default value is *stext*.

Value	Description
stext	Text Search

3.1.5 Simple/Advanced Search Selection Parameter — *pg*

The optional **pg** parameter is used to specify which kind of query is to be executed: simple or advanced. If omitted, a simple query is the default.

Value	Description
q	Used for submitting a Simple Search query.
aq	Used for submitting an Advanced Search query.

3.1.6 Search String Parameter — *q*

The required **q** parameter specifies the search string. Simple and Advanced Searches have different search string formats as described below.

Simple Search String

For Simple Searches (**pg=q**), the search string consists of one or more search terms separated by the CGI word concatenation operator (+). Each search term consists of an optional search operator (%2B or -), a search term, and an optional wildcard.

Value	Description
[\diamond]n*]{+[\diamond]n*}	<p>\diamond is one of two operators %2B or -. The %2B operator means that web pages that contain the following term are included in the search results. The - operator means that web pages that contain the following term are excluded from the search results. Note: %2B is the URL encoded representation for the plus sign (+) and is not to be confused with the CGI word concatenation operator (+).</p> <p><i>n</i> is a string of characters that is a word, phrase, or keyword. Refer to Section 10 for a discussion of AltaVista search string syntax. Refer to the table below for the Text Search special keywords.</p> <p>Non-alphanumeric characters must be specially coded so that they are not interpreted by the CGI protocol. Refer to Section 11, for a URL encoded representation table.</p> <p>* is a wildcard. There must be at least 3 characters before the *.</p> <p>+ is the CGI word concatenation operator used to separate search terms.</p> <p>Items within square brackets [] are optional.</p> <p>Items within braces { } are repeatable.</p>

Advanced Search String

For Advanced Searches (**pg=aq**), the search string consists of one or more search terms separated by the CGI word concatenation operator (+). Each search term consists of an optional Boolean search operator (**AND**, **OR**, **AND NOT**, or **NEAR**), a search term, and an optional wildcard.

Value	Description
[\diamond]n*]{+[\diamond]n*}	<p>\diamond is one of the Boolean operators AND, OR, AND NOT, or NEAR. The Boolean operators and their action are described in the Boolean Expression Table below. Either the expression or the symbol is valid. For example: <i>&q=peanut+AND+butter</i> and <i>&q=peanut+%26+butter</i> are equivalent. Note: %26 is the URL encoded representation for ampersand (&)</p> <p><i>n</i> is a string of characters that is a word, phrase, or keyword. Refer to Section 10 for a discussion of AltaVista search string syntax. Refer to the table below for the Text Search special keywords. Non-alphanumeric characters must be specially coded so that they are not interpreted by the CGI protocol. Refer to Section 11, for a URL encoded representation table.</p> <p>* is a wildcard. There must be at least 3 characters before the *.</p> <p>+ is the CGI word concatenation operator used to separate search terms.</p> <p>Items within square brackets [] are optional.</p> <p>Items within braces { } are repeatable.</p>

Boolean Operator Table

Operator	Symbol	Action
AND	%26	Finds documents containing all of the specified words or phrases. <i>peanut+AND+butter</i> finds documents with both the word peanut and the word butter. Note: %26 is the URL encoded representation for the ampersand (&)
OR	%7C	Finds documents containing at least one of the specified words or phrases. <i>peanut+OR+butter</i> finds documents containing either peanut or butter. The found documents could contain both items, but not necessarily. Note: %7C is the URL encoded representation for the bar ()

Operator	Symbol	Action
AND NOT	%21	Excludes documents containing the specified word or phrase. <i>Peanut+AND+NOT+butter</i> finds documents with peanut but not containing butter. NOT must be used with another operator, like AND. AltaVista does not accept <i>peanut+NOT+butter</i> . Note: %21 is the URL encoded representation the exclamation mark (!)
NEAR	%7E	Finds documents containing both of the specified words or phrases within 10 words of each other. <i>peanut+NEAR+butter</i> would find documents with peanut butter, but probably not any other kind of butter. Note: %7E is the URL encoded representation the tilde (~)
	%28 %29	Use parentheses to group complex Boolean phrases. For example, %28 <i>peanut+AND+butter</i> %29+AND+%28 <i>jelly+OR+jam</i> %29 finds web pages with the words peanut butter and jelly or peanut butter and jam or both. Note: %28 and %29 are the URL encoded representation for the left and right parenthesis ()

Text Search Special Keywords

The following special keywords are valid for both *Simple Search* and *Advanced Search* queries:

Keyword	Function
anchor%3A <i>text</i>	Finds pages that contain the specified word or phrase in the text of a hyperlink. anchor%3A %22 <i>Click here to visit garden.com</i> %22 would find pages with "Click here to visit garden.com" as a link. Note: %3A is the URL encoded representation for colon (:) and %22 is the URL encoded representation for double-quote (").
applet%3A <i>class</i>	Finds pages that contain a specified Java applet. Use applet%3A <i>morph</i> to find pages using applets called morph. Note: %3A is the URL encoded representation for colon (:))
domain%3A <i>domainname</i>	Finds pages within the specified domain. Use domain%3A <i>uk</i> to find pages from the United Kingdom, or use domain%3A <i>com</i> to find pages from commercial sites. Note: %3A is the URL encoded representation for colon (:))
host%3A <i>name</i>	Finds pages on a specific computer. The search host%3A <i>www.shopping.com</i> would find pages on the Shopping.com computer, and host%3A <i>dilbert.unitedmedia.com</i> would find pages on the computer called dilbert at unitedmedia.com. Note: %3A is the URL encoded representation for colon (:))
image%3A <i>filename</i>	Finds pages with images having a specific filename. Use image%3A <i>beaches</i> to find pages with images called beaches. Note: %3A is the URL encoded representation for colon (:))
like%3A <i>URLtext</i>	Finds pages similar to or related to the specified URL. For example, like%3A <i>www.abebooks.com</i> finds Web sites that sell used and rare books, similar to the www.abebooks site. like%3A <i>spl.lib.ca.us%2F</i> finds public and university library sites. like%3A <i>http%3A%2F%2Fwww.indiaks.com%2F</i> finds sites about culture on the Indian subcontinent. Note: %3A is the URL encoded representation for colon (:) and %2F is the URL encoded representation for forward slash (/)
link%3A <i>URLtext</i>	Finds pages with a link to a page with the specified URL text. Use link%3A <i>www.zip2.com</i> to find all pages linking to Zip2.com. Note: %3A is the URL encoded representation for colon (:))

Keyword	Function
text%3A <i>text</i>	Finds pages that contain the specified text in any part of the page other than an image tag, link, or URL. The search text%3A <i>graduation</i> would find all pages with the term graduation in them. Note: %3A is the URL encoded representation for colon (:)
title%3A <i>text</i>	Finds pages that contain the specified word or phrase in the page title (which appears in the title bar of most browsers). The search title%3A <i>sunset</i> would find pages with sunset in the title. Note: %3A is the URL encoded representation for colon (:)
url%3A <i>text</i>	Finds pages with a specific word or phrase in the URL. Use url%3A <i>zip2</i> to find all pages on all servers that have the word zip2 in the host name, path, or filename — the complete URL, in other words. Note: %3A is the URL encoded representation for colon (:)

Note: The query string must be in the correct format for the type of query selected; simple query if **pg=q** is used, or advanced query if **pg=aq** is used.

For example, a Simple Search query for pages with a link to a page with *.com* that contain *stuff* and *not pizza* or "*foo bar*" would be:

```
&pg=q&q=stuff+-pizza+%2B%22foo+bar%22+link%3Acom
```

An Advanced Search query for approximately the same thing would be:

```
&pg=aq&q=stuff+and+not+pizza+or+%28%22foo+bar%22+and+link%3Acom%29
```

3.1.7 Ranking String Parameter — *r*

The optional **r** parameter is only valid for advanced searches (**pg=aq**). The ranking string determines the order of the results for advanced searches. If a ranking string is not provided for an advanced query, the results appear in random order.

Value	Description
<i>n</i> { <i>+n</i> }	<i>n</i> is a string of characters which is a word or phrase. Refer to Section 10 for a discussion of AltaVista search string syntax. Non-alphanumeric characters must be specially coded so that they are not interpreted by the CGI protocol. Refer to Section 11, for a URL encoded representation table. + is the CGI word concatenation operator used to separate search terms. Items within braces { } are repeatable.

Basically, the query string (**q=string**) is used for the terms you wish to find on the web and the ranking sting (**r=string**) is used to establish your preferences within that broad set of results.

If you have no special ranking requirements for your search but don't want the results presented randomly, then simply enter the same terms in the query string and the ranking string (except those which you have excluded by AND NOT). For example:

```
&q=Frankenstein+AND+movie+AND+NOT+%22Mel+Brooks%22&r=frankenstein+movie
```

If you want to find someone named *John Smith*, you can improve your chances that pages relating to the right John Smith will appear high on the list of results. In the query string, enter *John Smith*. Then in the ranking string, enter everything you know about this person — all his interests, everything that's likely to appear on a web page that mentions his name. For example, the following query will sort pages about the Mark McGwire that plays baseball for the Cardinals and hits home runs to the top of the results:

&q=%22Mark+McGwire%22&r=baseball+%22home+run%22+cardinals

Note: Ranking for simple searches (**pg=q**) is determined by a few simple rules. Pages move up the list if they have either more instances of the search terms, search terms that are preceded by the **%2B** operator in the query, search terms that appear in relatively few other pages, and search terms closer together in the text. Pages that have the opposite qualities move down, or off, the list.

3.1.8 Number of Results per Query Parameter — *nbq*

The optional **nbq** parameter specifies the number of results that should be returned in this group of results. Fewer results will be returned if fewer than **nbq** documents satisfy the query. If **nbq** is not specified the default is 10 for Text Searches. The minimum and maximum values are 1 and 100 respectively. If a value that is less than the minimum value is specified, then the default value is used. If a value greater than the maximum value is specified, then the value of 100 is used.

Value	Description
<i>n</i>	Number of results per query (1 to 100).

For example, if **&nbq=20** is added to the query string, the VAL or XML output will contain 20 results rather than the usual 10. Of course, if the number of query results is fewer than **nbq** pages, only the number found will be returned.

3.1.9 Results Ordinal Parameter — *stq*

The optional **stq** parameter specifies the ordinal of the first result that should be returned in this group of results. This parameter is used to select the next range of **nbq** results from the output of the query. Initially, it is not necessary to specify this parameter. A value of 0 is assumed, indicating that the first document among the results is requested.

Value	Description
<i>n</i>	Ordinal, counting from 0, of the first result that should be returned. If a value that is greater than the total number of results available is specified, then no results are returned.

For example, assuming **nbq=10**, **&stq=0** will return the 1st through 10th result entry, and **&stq=147** will return the 148th through 157th result entry. If you specify a value for **stq** that is greater than the total number of result entries, you will not get back any results.

3.1.10 Start Date Parameter — *d0*

The optional **d0** parameter specifies the start date for the query. The **d1** parameter specifies the ending date. If **d0** is not specified, ISS will search all documents older than the ending date of the query. AltaVista's ISS uses the *Last Modified* date of a page when searches specify a date range.

Value	Description
<i>dd%2FMmm%2Fyyyy</i>	<i>dd</i> is the integer day (1–31), <i>Mmm</i> is the month (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec), and <i>yyyy</i> is the integer year. %2F is the URL encoded representation for forward slash (/).
<i>dd%2FMmm%2Fyy</i>	<i>dd</i> is the integer day (1–31), <i>Mmm</i> is the month (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec), and <i>yy</i> is the integer year modulo 100 (00–99). %2F is the URL encoded representation for forward slash (/).

Value	Description
<code>dd%2Fmm%2Fyyyy</code>	<code>dd</code> is the integer day (1–31), <code>mm</code> is the integer month (1–12), and <code>yyyy</code> is the integer year. <code>%2F</code> is the URL encoded representation for forward slash (/).
<code>dd%2Fmm%2Fyy</code>	<code>dd</code> is the integer day (1–31), <code>mm</code> is the integer month (1–12), and <code>yy</code> is the integer year modulo 100 (00–99). <code>%2F</code> is the URL encoded representation for forward slash (/).
<code>dd</code>	<code>dd</code> is the integer day (1–31). The current month and year are assumed.

For example, either of the two parameter strings below would specify the range March 5, 1995 through August 15, 1997.

```
&d0=05%2FMar%2F95&d1=15%2FAug%2F1997
&d0=5%2F3%2F95&d1=15%2FAug%2F1997
```

The following parameter string would specify the range from the first day of the current month until *today*:

```
&d0=1
```

3.1.11 End Date Parameter — `d1`

The optional `d1` parameter specifies the end date for the query. The `d0` parameter specifies the starting date. If `d1` is not specified, ISS will search all documents newer than the starting date of the query. AltaVista's ISS uses the *Last Modified* date of a page when a date range is specified.

Value	Description
<code>dd%2FMmm%2Fyyyy</code>	<code>dd</code> is the integer day (1–31), <code>Mmm</code> is the month (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec), and <code>yyyy</code> is the integer year. <code>%2F</code> is the URL encoded representation for forward slash (/).
<code>dd%2FMmm%2Fyy</code>	<code>dd</code> is the integer day (1–31), <code>Mmm</code> is the month (Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec), and <code>yy</code> is the integer year modulo 100 (00–99). <code>%2F</code> is the URL encoded representation for forward slash (/).
<code>dd%2Fmm%2Fyyyy</code>	<code>dd</code> is the integer day (1–31), <code>mm</code> is the integer month (1–12), and <code>yyyy</code> is the integer year. <code>%2F</code> is the URL encoded representation for forward slash (/).
<code>dd%2Fmm%2Fyy</code>	<code>dd</code> is the integer day (1–31), <code>mm</code> is the integer month (1–12), and <code>yy</code> is the integer year modulo 100 (00–99). <code>%2F</code> is the URL encoded representation for forward slash (/).
<code>dd</code>	<code>dd</code> is the integer day (1–31). The current month and year are assumed.

For example, either of the two parameter strings below would specify the range March 5, 1995 through August 15, 1997.

```
&d0=05%2FMar%2F95&d1=15%2FAug%2F1997
&d0=5%2F3%2F95&d1=15%2FAug%2F1997
```

And, the following parameter string would specify the range of all pages that were last modified on or before the first day of the current month:

```
&d1=1
```

3.1.12 Language Selection Parameters — *kl* and *KL*

The optional **kl** (that's k as in kilo and l as in lima) and **KL** parameters are used to restrict the document search to a particular language or languages. The **kl** parameter is used to select all languages or to restrict the document search to a single language, while the **KL** parameter is used to restrict the document search to multiple languages. Values for the parameters are two-letter ISO 639 standard IDs selecting the particular language. If not specified, **XX** (all languages) is the default. A value of **XX** is not valid for the **KL** parameter.

Sorted by Value	
Value	Language
cs	Czech
da	Danish
de	German
el	Greek
en	English
es	Spanish
et	Estonian
fi	Finnish
fr	French
he	Hebrew
hu	Hungarian
is	Icelandic
it	Italian
ja	Japanese
ko	Korean
lt	Lithuanian
lv	Latvian
nl	Dutch
no	Norwegian
pl	Polish
pt	Portuguese
ro	Romanian
ru	Russian
sv	Swedish
XX	all languages
zh	Chinese

Sorted by Language	
Value	Language
XX	all languages
zh	Chinese
cs	Czech
da	Danish
nl	Dutch
en	English
et	Estonian
fi	Finnish
fr	French
de	German
el	Greek
he	Hebrew
hu	Hungarian
is	Icelandic
it	Italian
ja	Japanese
ko	Korean
lv	Latvian
lt	Lithuanian
no	Norwegian
pl	Polish
pt	Portuguese
ro	Romanian
ru	Russian
es	Spanish
sv	Swedish

For example, to restrict results to pages in Spanish, add the following to the query string:

```
&kl=es
```

If more than one language is required, multiple two-letter IDs may be used with the **KL** parameter. For example, to restrict results to pages in Spanish, Portuguese, and Italian, add the following to the query string:

```
&KL=esptit
```

3.1.13 Encoding Selection Parameter — *enc*

The optional **enc** parameter is used to determine the way a query string is parsed by the web server and it directs the web server to return query results using the selected encoding. The parameter takes an alphanumeric string that uniquely identifies a particular encoding. If not specified, the default value is *iso88591*.

Value	Encoding	Primary Language Group
ascii	US-ASCII	English
big5	Big5	Chinese (Traditional)
cp1250	Windows-1250	Central and Eastern European (Windows Latin 2)
cp1251	Windows-1251	Russian (Windows Cyrillic)
cp1252	Windows-1252	Western European (Windows Latin 1)
cp1253	Windows-1253	Greek (Windows)
cp1254	Windows-1254	Turkish (Windows)
cp1255	Windows-1255	Hebrew (Windows)
cp1256	Windows-1256	Arabic (Windows)
cp1257	Windows-1257	Baltic (Windows)
cp1258	Windows-1258	Vietnamese (Windows)
euccn	GB2312	Chinese (Simplified)
euclp	EUC-JP	Japanese
euclkr	EUC-KR	Korean
iso88591	ISO-8859-1	Western European (Latin 1)
iso88592	ISO-8859-2	Central and Eastern European (Latin 2)
iso88593	ISO-8859-3	Southern European (Latin 3)
iso88594	ISO-8859-4	Northern European (Latin 4)
iso88595	ISO-8859-5	Russian (Cyrillic)
iso88596	ISO-8859-6	Arabic
iso88597	ISO-8859-7	Greek
iso88598	ISO-8859-8	Hebrew
iso88599	ISO-8859-9	Turkish (Latin 5)
iso885910	ISO-8859-10	Nordic (Latin 6)
jis	ISO-2022-JP	Japanese
koi8r	KOI8-R	Russian
sjis	Shift_JIS	Japanese
utf8	UTF-8	All languages. Unicode (ISO 10646)

Note: The Encoding column above contains the official names of the character sets as used on the Internet as defined by IANA. When XML output is requested (**dp=xml**), the returned XML will include this value in the XML encoding declaration. For example: `<?xml version="1.0" encoding="ISO-8859-1" ?>`

Note: To achieve maximum benefit from the **enc** parameter, the end user's browser must be set appropriately. End users should consult their browser's help pages.

3.1.14 Site Compression Parameter — *sc*

The optional *sc* parameter controls site compression. When site compression is off, AltaVista's ISS service will return more than one URL per site in each set of query results if appropriate. When site compression is on, AltaVista's ISS returns only one result per web site. If not specified, site compression is *on* for Simple Searches and it is *off* for Advanced Searches.

Value	Description
on	Turns site compression on.
off	Turns site compression off.

3.1.15 Site Identification Parameter — *si*

The optional *si* parameter identifies the compressed site for which all of the results are requested. Only results from this site are returned. The value for this parameter is returned as the **SiteID** key from a previous query. Refer to Sections 4.1.2 and 5.1.2 for the **SiteID** key. If the *si* parameter is present, the *sc* parameter is ignored.

3.2 Multimedia Search Input Parameters

The following three quick reference tables provide an alphabetical list of the *Multimedia Audio Search*, *Multimedia Image Search*, and *Multimedia Video Search* parameters, their values and ranges, whether they are required or are optional, and their default value if appropriate. Within the table the user-supplied values are displayed in the *Italics* typeface, and values that must be specified exactly as shown are displayed in normal typeface. Values separated by the bar character (|) represent a list from which one value is to be chosen. Values within parentheses represent a range of values. A detailed description of each of the parameters follows these tables.

Multimedia Audio Search Parameters Quick Reference				
Parameter	Description	Value (Range)	Required / Optional	Default Value
amp3	Include Mp3 Files Flag	0 1	Optional	0
aother	Include Other Audio Files Flag	0 1	Optional	0
astrm	Include Streaming Audio Flag	0 1	Optional	0
audset	Multimedia Audio Search Flag	1	Required	
awav	Include WAV File Flag	0 1	Optional	0
dp	Output Format	val xml	Required	
FFF	Family Friendly Filter Option	0 1 on off	Optional	on
macat	Multimedia Audio Partner Selection	0 1 CDNOW EMUSIC EPITONIC MJUICE ON24 RIFFAGE	Optional	0
mmdo	Multimedia Search Subtype	3 4 5 6	Optional	3
mmF	Multimedia Source FTP Flag	0 1	Optional	0
mmW	Multimedia Source Web Flag	0 1	Optional	0
nbq	Number Of Entries In Results	<i>number</i> (1 to 20)	Optional	12
objid	Multimedia About/Similar Object Id	<i>number</i>	Optional	
objtype	Multimedia Object Type	taudio	Optional	
pagefp	Multimedia Same Page Id	<i>number</i>	Optional	
pg	Query Type	q	Optional	q
q	Query String	<i>CGI query string</i>	Required	
stq	Ordinal Of First Results Entry	number	Optional	0

Multimedia Audio Search Parameters Quick Reference				
Parameter	Description	Value (Range)	Required / Optional	Default Value
stype	Search Type	saudio	Required	
who	Customer Id Code	<i>issid</i>	Required	

The following example query illustrates the use of some of the Audio *Multimedia Search* input parameters. This query looks for songs about “trees” in any format from all audio sources.

```
http://partners.altavista.com/cgi-
bin/query?stype=saudio&pg=q&audset=1&q=trees&amp3=1&awav=1&astrm=1&aotr=1&mmW=
1&mmF=1&macat=0&dp=xml&who=issid
```

Multimedia Image Search Parameters Quick Reference				
Parameter	Description	Value (Range)	Required / Optional	Default Value
dp	Output Format	val xml	Required	
FFF	Family Friendly Filter Option	0 1 on off	Optional	on
ibb	Include Buttons And Banners Flag	0 1	Optional	0
ibw	Include Black And White Image Flag	0 1	Optional	0
iclr	Include Color Images Flag	0 1	Optional	0
igrph	Include Graphics Flag	0 1	Optional	0
imgset	Multimedia Image Search Flag	1	Required	
ipht	Include Photos Flag	0 1	Optional	0
micat	Multimedia Image Partner Selection	0 1 CDNOW CORBIS GETTY	Optional	0
mmdo	Multimedia Search Subtype	3 4 5 6	Optional	3
mmW	Multimedia Source Web Flag	0 1	Optional	0
nbq	Number Of Entries In Results	<i>number</i> (1 to 20)	Optional	12
objid	Multimedia About/Similar Object Id	<i>number</i>	Optional	
objtype	Multimedia Object Type	timage	Optional	
pagefp	Multimedia Same Page Id	<i>number</i>	Optional	
pg	Query Type	q	Optional	q
q	Query String	<i>CGI query string</i>	Required	
stq	Ordinal Of First Results Entry	number	Optional	0
stype	Search Type	simage	Required	
who	Customer Id Code	<i>issid</i>	Required	

The following example query illustrates the use of some of the Image *Multimedia Search* input parameters. This query looks for color photographs of “cats” from all image sources.

```
http://partners.altavista.com/cgi-
bin/query?stype=simage&pg=q&imgset=1&q=cats&ipht=1&igrph=0&ibb=0&iclr=1&ibw=0&mmW=1
&micat=0&dp=xml&who=issid
```

Multimedia Video Search Parameters Quick Reference				
Parameter	Description	Value (Range)	Required / Optional	Default Value
dp	Output Format	val xml	Required	
FFF	Family Friendly Filter Option	0 1 on off	Optional	on
mmdo	Multimedia Search Subtype	3 4 5 6	Optional	3
mmW	Multimedia Source Web Flag	0 1	Optional	0
mvcat	Multimedia Video Partner Selection	0 1 ABC FASTV LAUNCH MSNBC ML ON24 OTS VIDNET	Optional	0
nbq	Number Of Entries In Results	<i>number</i> (1 to 20)	Optional	12
objid	Multimedia About/Similar Object Id	<i>number</i>	Optional	
objtype	Multimedia Object Type	tvideo	Optional	
pagefp	Multimedia Same Page Id	<i>number</i>	Optional	
pg	Query Type	q	Optional	q
q	Query String	<i>CGI query string</i>	Required	
stq	Ordinal Of First Results Entry	<i>number</i>	Optional	0
stype	Search Type	svideo	Required	
vavi	Include Avi Files Flag	0 1	Optional	0
vidset	Multimedia Video Search Flag	1	Required	
vmov	Include QuickTime Files Flag	0 1	Optional	0
vmpeg	Include MPEG Files Flag	0 1	Optional	0
vothr	Include Other Video Files Flag	0 1	Optional	0
vstrm	Include Streaming Video Files Flag	0 1	Optional	0
who	Customer Id Code	<i>issid</i>	Required	

The following example query illustrates the use of some of the Video *Multimedia Search* input parameters. This query looks for videos of “sailboats” in all formats from all video sources.

```
http://partners.altavista.com/cgi-bin/query?stype=svideo&pg=q&vidset=1&q=sailboats&vavi=1&vmpeg=1&vmov=1&vstrm=1&vothr=1&mmW=1&mvcat=0&dp=xml&who=issid
```

The following subsections describe each of the *Multimedia Search* parameters in more detail.

3.2.1 Customer ID Parameter — who

The required **who** parameter identifies the particular AltaVista customer that is accessing the ISS link. This parameter is a necessary component of AltaVista’s ISS security.

Value	Description
<i>issid</i>	The <i>issid</i> is a unique token assigned to each particular ISS customer by AltaVista.

3.2.2 Output Format Parameter — dp

The required **dp** parameter indicates which ISS output format, VAL or XML, is to be selected for all content returned by **mhttpd**.

Value	Description
val	Requests VAL format (refer to Section 4.2).
xml	Requests XML format (refer to Section 5.2).

3.2.3 Family Friendly Filter Parameter — *FFF*

The optional **FFF** parameter controls the level of family friendly filtering enforced by the search engine. If family friendly filtering is on, the returned search results will be filtered for offensive *English* and only pages containing non-offensive *English* will be returned. If family friendly filtering is off, all matching results will be returned. If the query string does not contain a valid **FFF=value** the default is that family friendly filtering is *on*.

Value	Description
on	Family friendly filtering is on
off	Family friendly filtering is off
1	Family friendly filtering is on
0	Family friendly filtering is off

Note: If the family friendly filter is on, ISS results will not contain *Related Search* items (if you have purchased this search option).

3.2.4 Search Type Parameter — *stype*

The required **stype** parameter is used to select the type of web search sent to ISS: text pages, audio files, video files, or image files. For multimedia searches there are three values.

Value	Description	Format
saudio	Multimedia Search (audio)	MP3, WAV, and/or streaming
svideo	Multimedia Search (video)	AVI, MPEG, MOV, and/or streaming
simage	Multimedia Search (image)	JPEG, and/or GIF

3.2.5 Search Selection Parameter — *pg*

The optional **pg** parameter is used to specify which kind of query is to be executed. For Multimedia searches only simple queries are allowed. If omitted, a simple query is the default.

Value	Description
q	Used for submitting a simple query.

3.2.6 Search String Parameter — *q*

The required **q** parameter specifies the search string. The search string consists of one or more search terms separated by the CGI word concatenation operator (+). Each search term consists of an optional search operator (%2B or -), a search term, and an optional wildcard.

Value	Description
[\diamond]n*]{+[\diamond]n*}	<p>\diamond is one of two operators %2B or -. The %2B operator means that multimedia objects that contain the following term are included in the search results. The - operator means that multimedia objects that contain the following term are excluded from the search results. Note: %2B is the URL encoded representation for plus sign (+) and is not to be confused with the CGI word concatenation operator (+).</p> <p>n is a string of characters which is a word, phrase, or keyword. Refer to Section 10 for a discussion of AltaVista search string syntax. Refer to the table below for the Multimedia Search special keywords. Non-alphanumeric characters must be specially coded so that they are not interpreted by the CGI protocol. Refer to Section 11, for a URL encoded representation for table.</p> <p>* is a wildcard. There must be at least 3 characters before the *.</p> <p>+ is the CGI word concatenation operator used to separate search terms.</p> <p>Items within square brackets [] are optional.</p> <p>Items within braces { } are repeatable.</p>

For example, the following search string will look for images of the Bluebird's land speed record project:

```
&q=bluebird+%2Bland+%2Bspeed+%2Brecord
```

Multimedia Search Special Keywords

The following special keywords are valid for Multimedia searches:

Keyword	Function
domain%3A <i>domainname</i>	Finds pages within the specified domain. Use domain%3Auk to find pages from the United Kingdom, or use domain%3Acom to find pages from commercial sites. Note: %3A is the URL encoded representation for colon (:).
host%3A <i>name</i>	Finds pages on a specific computer. The search host%3Awww.shopping.com would find pages on the Shopping.com computer, and host%3Adilbert.unitedmedia.com would find pages on the computer called dilbert at unitedmedia.com. Note: %3A is the URL encoded representation for colon (:).
media%3A <i>text</i>	Finds files containing the specified text in their filename or URL. Use media%3Abeaches to find pages with images, audios or videos named beaches. Note: %3A is the URL encoded representation for colon (:).
title%3A <i>text</i>	Finds pages that contain the specified word or phrase in the page title (which appears in the title bar of most browsers). The search title%3Asunset would find pages with sunset in the title. Note: %3A is the URL encoded representation for colon (:).
url%3A <i>text</i>	Finds pages with a specific word or phrase in the URL. Use url%3ACorbis to find all pages that have the word Corbis in the host name, path or file name — the complete URL, in other words. Note: %3A is the URL encoded representation for colon (:).
artist%3A <i>text</i>	Finds audio files of the specified artist. Use artist%3AMetallica to find audio files of Metallica. Note: %3A is the URL encoded representation for colon (:).

Keyword	Function
song%3A <i>text</i>	Finds audio files whose title includes your text, for example, song%3A <i>vogue</i> . Note: %3A is the URL encoded representation for colon (:).
genre%3A <i>text</i>	Finds audio files within the specified genre, for example, genre%3A <i>jazz</i> . Note: %3A is the URL encoded representation for colon (:).

For example, the following query string will look for all songs by Metallica that contain the word *turn* in the title.

```
&q=%2Bartist%3Ametallica+%2Bsong%3Aturn
```

3.2.7 Number of Results per Query Parameter — *nbq*

The optional **nbq** parameter specifies the number of results that should be returned in this group of results. Fewer results will be returned if fewer than **nbq** documents satisfy the query. If **nbq** is not specified the default is 12 for Multimedia searches. The minimum and maximum values are 1 and 20 respectively. If a value that is less than the minimum value is specified, then the default value is used. If a value greater than the maximum value is specified, then the value of 20 is used.

Value	Description
<i>n</i>	Number of results per query (1 to 20).

For example, if **&nbq=20** is added to the query string, the VAL or XML output will contain 20 results rather than the usual 12. Of course, if the number of query results is fewer than **nbq** pages, only the number found will be returned.

3.2.8 Results Ordinal Parameter — *stq*

The optional **stq** parameter specifies the ordinal of the first result that should be returned in this group of results. This parameter is used to select the next range of **nbq** results from the output of the query. Initially, it is not necessary to specify this parameter. A value of 0 is assumed, indicating that the first document among the results is requested.

Value	Description
<i>n</i>	Ordinal, counting form 0, of the first result that should be returned. If a value that is greater than the total number of results available is specified, then no results are returned.

For example, assuming **nbq=10**, **&stq=0** will return the 1st through 10th result entry, and **&stq=147** will return the 148th through 157th result entry. If you specify a value for **stq** that is greater than the total number of result entries, you will not get back any results.

3.2.9 Multimedia Search Type Options Parameter — *mmdo*

The optional **mmdo** parameter determines which page is being displayed. If the **mmdo** parameter is not present the default value is 3.

Value	Description
3	Regular Multimedia Search. Regular results based on the terms typed in by the user.

Value	Description
4	About this query. Queries for the objid and returns information about its occurrences on the web and other data specific to that object. The objid parameter must be specified for this query. Its value is returned as MmID from a previous query (mmdo=3 for example).
5	Similar Query. Currently applies to images only (stype=simage). Given an image, objid , display other images from the database that are similar in appearance. Similarity is based on visual characteristics such as dominant colors, shapes and textures. The objid parameter must be specified for this query. Its value is returned as MmID from a previous query (mmdo=3 for example).
6	Same Page. Return results about other images (stype=simage), audios (stype=saudio), or videos (stype=svideo) that appear on the same page, pagefp as the selected object. The pagefp parameter must be specified for this query. Its value is returned as MmPage from a previous query (mmdo=3 for example).

The following is an example of a **mmdo=3** query. The following parameters will look for color photos that have *ford* in the name of the file.

```
&stype=simage&imgset=1&q=ford&mmdo=3&ipht=1&iclr=1&mmW=1&micat=0
```

3.2.10 Object Identification Parameter — *objid*

The optional **objid** parameter identifies the object for which additional information is desired. It must be specified for **mmdo=4** and **mmdo=5** queries. The value for this parameter is returned as the **MmID** key from a previous query (**mmdo=3** for example). Refer to Sections 4.2.2 and 5.2.2 for the **MmID** key.

The following is an example of a **mmdo=4** query. If one of the results from the example **mmdo=3** query shown in Section 3.2.9 contained a **MmID** value of 14247144, then the following query would request additional information about the object. Only results for the object are returned.

```
&stype=simage&imgset=1&q=ford&mmdo=4&objid=14247144&objtype=timage&ipht=1
&iclr=1&mmW=1&micat=0
```

Note: Currently, the results from a **mmdo=4** query of an **objid** does not return any additional information beyond what is returned in the **mmdo=3** query. This may change in the future.

The following is an example of a **mmdo=5** query. If one of the results from the **mmdo=3** example query shown in Section 3.2.9 contained **Similar=y** and a **MmID** value of 14247144, then the following query would request results for objects that are similar to the current object. Refer to Sections 4.2.2 and 5.2.2 for information about the Similar and **MmID** output keys.

```
&stype=simage&imgset=1&q=ford&mmdo=5&objid=14247144&ipht=1&iclr=1&mmW=1&micat=0
```

Note: Currently **mmdo=5** queries are only valid for image searches (**stype=simage**).

3.2.11 Multimedia Object Type — *objtype*

The **objtype** parameter is required for **mmdo=4** queries only and its value identifies the multimedia object type. Refer to the **mmdo=4** example in Section 3.2.10 for an example of the **objtype** parameter's usage.

Value	Description
taudio	The query is for an audio object.
timage	The query is for an image object.
tvideo	The query is for a video object.

3.2.12 Page Identification Parameter — *pagefp*

The optional **pagefp** parameter identifies the page for which a *Same Page* (**mmdo=6**) query is made. It must be specified for **mmdo=6** queries. The value for this parameter is returned as the **MmPage** key from a previous query (**mmdo=3** for example). Refer to sections 4.2.2 and 5.2.2 for the **MmPage** key.

The following is an example of a **mmdo=6** query. If one of the results from the **mmdo=3** example query shown in Section 3.2.9 contained **MmPage** value of 17459586915029445966, then the following query would request results for all of the other objects on the same page as the current object.

```
&stype=simage&imgset=1&q=ford&mmdo=6&pagefp=17459586915029445966&ipht=1&i
clr=1&mmW=1&micat=0
```

3.2.13 Audio Specific Parameters — (*stype=audio*)

3.2.13.1 Audio Search Selection Parameter - *audset*

The **audset** parameter is required for all Multimedia Audio Search queries.

Value	Descriptions
1	Audio Search Flag

3.2.13.2 Audio File Type Parameters

Audio searches are done by file type. While each of the four parameters (**amp3**, **awav**, **astrm**, and **aothr**) is optional, every audio query must specify at least one of them with a value of "1" or no output will be returned. Each parameter specifies a file type: mp3, WAV, streaming audio, other audio. Streaming audio consists of RealMedia and Microsoft Streaming Media. Any audio file not falling into mp3, WAV, or streaming audio file types is classified as *other*.

Parameter	Value and Description
amp3	if set to 1, include mp3 files
awav	if set to 1, include WAV
astrm	if set to 1, include streaming files
aothr	if set to 1, include <i>other</i> files

Note: Setting a parameter's value to "1" turns it on. Omitting a parameter, or setting its value to "0" turns it off. Values other than "1" or "0" are considered off.

For example, adding **&amp3=1&awav=0&astrm=0&aothr=0** or **&amp3=1** to the query string will search only for mp3 files.

3.2.13.3 Audio Source Parameters

Audio searches can be done from the web and/or ftp and/or AltaVista Partner sites. While neither the **mmW** nor the **mmF** nor the **macat** parameter is required, at least one must be specified with a value of "1" in order to receive any query results. The following parameters are used to specify sources:

Parameter	Value and Description
mmW	if set to 1, search web
mmF	if set to 1, search FTP sites
macat	If set to 1, search no audio partner sites If set to 0, search all audio partner sites If set to one of the following, search for results on that partner's site: CDNOW, EMUSIC, EPITONIC, MJUICE, ON24, RIFFAGE. These settings correspond to the following AltaVista partners: CDnow Online, Inc., EMusic.com, Inc., Epitonic.com, MJUICE.com, ON24 Inc., and Riffage.com, Inc., respectively.

Note: Omitting a parameter, or setting its value to "0" turns it off. Setting a parameter to any other value than specified above is considered off.

For example, to obtain from ISS a list of the MP3 files on the CDNOW site related to *jazz*, your program could send the following query string to partners.altavista.com:

```
&dp=xml&who=issid&stype=audio&pg=q&amp3=1&macat=CDNOW&q=jazz
```

3.2.14 Image Specific Parameters — (stype=simage)

3.2.14.1 Image Search Selection Parameter - *imgset*

The **imgset** parameter is required for all Multimedia Audio Search queries.

Value	Descriptions
1	Image Search Flag

3.2.14.2 Image Characteristics Parameters

Image searches are done according to image characteristics. Each image query must specify color (**iclr=1**) and/or black and white (**ibw=1**). Each image query must also specify photos (**ipht=1**) and/or graphics (**igrph=1**). If *buttons and banners* are to be included in the image query results, specify **ibb=1** in addition to the other *required* specifications.

Parameter	Value and Description
ipht	if set to 1, include photos
igrph	if set to 1, include graphics
iclr	if set to 1, include color images
ibw	if set to 1, include black and white images
ibb	if set to 1, include buttons and banners in photo and/or graphics results

Note: Setting a parameter's value to "1" turns it on. Omitting a parameter, or setting its value to "0" turns it off. Values other than "1" or "0" are considered off.

For example, adding `&iclr=1&ibw=0&ipht=1&igrph=0&ibb=0` or `&iclr=1&ipht=1` to the query string will search only for color photos.

3.2.14.3 Image Source Parameters

Image searches can be done from the web and/or AltaVista Partner sites. Either the `mmW` and/or the `micat` parameter must be specified in order to receive any query results. The following parameters are used to specify sources:

Parameter	Value and Description
<code>mmW</code>	if set to 1, search web
<code>micat</code>	If set to 1, don't search any image partner sites If set to 0, search all image partner sites If set to one of the following, search for results on that partner's site: CDNOW, CORBIS, GETTY. These settings correspond to the following AltaVista partners: CDnow Online, Inc., Corbis.com, and Getty Images, Inc., respectively.

Note: Omitting a parameter, or setting its value to "0" turns them off. Setting a parameter to any other value than specified above is considered off.

For example, adding `&mmW=1&micat=GETTY` to the query string will search the web and the GETTY site and not the CDNOW nor the CORBIS sites.

3.2.15 Video Specific Parameters — (styp=svideo)

3.2.15.1 Video Search Selection Parameter - vidset

The `vidset` parameter is required for all Multimedia Audio Search queries.

Value	Descriptions
1	Audio Search Flag

3.2.15.2 Video File Type Parameters

Video searches are done by file type. Each video query must specify one or more of the following file types: avi, mpeg, QuickTime™, streaming video, other video. Streaming video consists of RealMedia and Microsoft Streaming Media. Any video file not falling into avi, mpeg, QuickTime™, or streaming video file types is classified as *other*.

Parameter	Value and Description
<code>vavi</code>	if set to 1, include avi files
<code>vmpeg</code>	if set to 1, include mpeg files
<code>vmov</code>	if set to 1, include QuickTime™ files
<code>vstrm</code>	if set to 1, include streaming files
<code>vothr</code>	if set to 1, include <i>other</i> files

Note: Setting a parameter's value to "1" turns it on. Omitting a parameter, or setting its value to "0" turns it off. Values other than "1" or "0" are considered off.

For example, adding `&vavi=0&vmpeg=1&vmov=1&vstrm=0&vothr=0` or `&vmpeg=1&vmov=1` to the query string will search only mpeg and QuickTime™ video files.

3.2.15.3 Video Source Parameters

Video searches can be done from the web and/or AltaVista Partner sites. Either the `mmW` and/or the `mvcat` parameter must be specified in order to receive any query results. The following parameters are used to specify sources:

Parameter	Value and Description
<code>mmW</code>	if set to 1, search web
<code>mvcat</code>	If set to 0, search all video partner sites
	If set to one of the following, search for results on that partner's site: <i>ABC, FASTV, LAUNCH, ML, MSNBC, OTS, ON24, VIDNET</i> . These settings correspond to the following AltaVista partners: ABC, Inc., FasTV, Inc., LAUNCH Media, Inc., Merrill Lynch & Co., Inc., MSNBC.com, On the Scene Productions, Inc., ON24 Inc., and Vidnet.com, respectively. Note: ABC, Inc. video is available only by prior arrangement with AltaVista.

Note: Omitting a parameter, or setting its value to "0" turns them off. Setting the parameter to any other value than specified above is considered off.

For example, to obtain from ISS a list of "election 2000" video files from both the web and from the MSNBC.com site:

```
&pg=q&stype=svideo&q=election+2000&vavi=1&vmpeg=1&vmov=1&vstrm=1&vothr=1&
mmW=1&mvcat=MSNBC&dp=xml&who=issid
```

4 ISS VAL Output Format

ISS supports two different output formats for query results: VAL (the original format created for the ISS product) and XML (a format introduced in the 2.0 ISS release). Customers specify the desired output format for each individual query via the **dp** parameter (see Sections 3.1.2 and 3.2.2). The same information is returned regardless of output format type selection. This section describes the VAL output format. See Section 5 for a description of the XML output format.

VAL output consists of three parts: the HTTP header, the VAL header, and the optional VAL search results. Each part is separated from the next part by a single blank line.

The HTTP header identifies the HTTP version, the HTTP status, the result date, the AltaVista server version, the MIME version, the content length, and the content type (text/plain). The following is an example of the HTTP header:

```
HTTP/1.0 200 OK
Date: Fri, 02 Jun 2000 22:42:17 GMT
Server: AV/1.0.1
MIME-Version: 1.0
Content-Length: 2953
Content-Type: text/plain
```

The VAL header and VAL search results consist of multiple lines each containing a *key=value* pair. The keys are case sensitive. The VAL search results may be repeated more than once with each result separated from the next by a single blank line.

Four important points should be followed when implementing a VAL parser:

- Each section of the body is separated from the next one by a single blank line. The first section is the header. The second and subsequent sections are the results.
- Ignore any unknown *keys*. This will prevent the code from breaking when a new feature is added.
- *Do not* depend on the ordering of *key=value* pairs within each section.
- In all references, MaxI represents the value $+2^{31-1}$, MaxL the value $\pm 2^{63-1}$, and MinI the value -2^{31-1} .

This section is divided into two subsections, one for each of the ISS search products: *Text* and *Multimedia*.

4.1 VAL Text Search Output

4.1.1 VAL Text Search Output Header

Key	Range	Description
M	0...MaxL	Total number of matches found.
C	0... MaxI	Number of keywords or phrases in the query.
e	1... 8	An error number. Present only if there is an error.

Key	Range	Description
E		An error string associated with the above e=<error number>. Present if there is an error. Possible values are: <ul style="list-style-type: none"> e=1. E=internal error e=2. E=no service available e=3. E=syntax error e=4. E=query too long e=5. E=Incorrect ID for Internet Search Services. e=6. E=Missing ID for Internet Search Services. e=7. E=The service is not available to you; please contact your sales representative. e=8. E=The server is unavailable. Please try again.
B		If present, is a list of words, separated by commas (,), which are spell check suggestions. Both the query string (see Section 3.1.6) and the Advanced Search ranking string (see Section 3.1.7) are spell checked.
SiteName		Site name. Only returned for compressed site expansion requests (see si parameter in Section 3.1.15).
K		A keyword or phrase. For <i>Simple Searches</i> (pg=q) the search string (q=text) is used. For <i>Advanced Searches</i> (pg=aq) the ranking string (r=text) is used. For <i>Advanced Searches</i> (pg=aq) it is not returned if the ranking string (r=text) was omitted. Each K value can have the following options on subsequent lines: (c and u). Not returned if there are no results returned (M=0).
c	0...MaxL	For the preceding keyword or phrase, total number of occurrences of the keyword or phrase in the entire index. If absent, default is 0. Not returned for queries containing the si parameter. For <i>Advanced Searches</i> (pg=aq) it is not returned if the ranking string (r=text) was omitted.
u	y, n	For the preceding keyword or phrase, indicates whether the keyword or phrase was used (y) or ignored (n). For <i>Advanced Searches</i> (pg=aq) it is not returned if the ranking string (r=text) was omitted.
I	1...MaxI	Indicates the range of the document results sent back by the search engine. This is represented by 2 integers separated by a hyphen. For example: <i>I=11-20</i> indicates that there are 10 results returned and that they are results 11 through 20. The number of results returned depends upon the total number of matches found (M above), and the setting of the nbq parameter (see Section 3.1.8). The range of the results returned is determined by the stq parameter (see Section 3.1.9). The I key is not returned if there are no results returned (i.e. M=0).

Examples:

The header returned for a *Simple Search* query of "apple pie" filling,

`&pg=q&q=%22apple+pie%22+filling` is:

```
M=95796
C=2
K=apple pie
c=19284
u=y
K=filling
c=441844
```

```
u=y
I=1-10
```

The header returned for a *Simple Search* query of “serch engines”,
&pg=q&q=%22serch+engines%22 is:

```
M=77
C=1
B=search
K=serch engines
c=234
u=y
I=1-10
```

Note the spell check result B=search above, which indicates that serch might be a misspelling of search.

The header returned for the *Advanced Search* Boolean expression of *cat NOT dog*,
&pg=aq&q=cat+NOT+dog is:

```
M=0
C=0
e=3
E=syntax error
```

Note: This is a syntax error, *cat AND NOT dog* is valid, and will not return an error.

The header returned for the empty query “” &pg=q&q=%22%22 or if no results that satisfy the query were found is:

```
M=0
C=0
```

4.1.1.1 Related Search Output

When a *Simple Search* query (**pg=q**) is issued, and if a customer purchases the *Related Search* service, then *RelatedTopic* and *RelatedOptions* are returned in the Text Header (see the example in Section 4.1.4). *RelatedOptions* are the parameters that can be included in a subsequent query to get results for the *RelatedTopic*. They are not returned for *Advanced Search* queries (**pg=aq**) nor if the family friendly filter (see Section 3.1.3) is on. Note: this is only returned if there are related topics associated with the user’s query.

Key	Description
RelatedTopic	Related Search Topic to be displayed to the user.
RelatedOptions	Related Search query options. This CGI-query string can be concatenated with the user’s dp and who parameters to request results for the related topic.

4.1.1.2 RealNames Output

When a *Simple Search* query (**pg=q**) is issued, and if a customer purchases the *RealNames* service, and *RealNames* are available, the Text Header returns three extra fields: RNKey, RNSite, and RNPerfectMatch (see the example in Section 4.1.5). RNKey is the *RealNames* key word, and RNSite is the URL address containing the *RealNames* information. They are not returned for *Advanced Search* queries (**pg=aq**) nor if the family friendly filter (see Section 3.1.3) is on.

Key	Description
RNKey	RealNames key word
RNSite	the URL address containing RealNames information
RNPerfectMatch	value (y/n) indicates whether or not there is a perfect match

If there is no existing *RealNames* information for a given query, these fields are not returned in the output.

4.1.2 VAL Text Search Results Output

Key	Range	Description
T		Title of the document. It is always present and is listed first, but it may have an empty value
u		The URL for the document. Always present.
a		Abstract of the document.
s	0...MaxI	Size of the document, in bytes.
r	MinI...MaxI	The relevancy score for the document.
d		Last-modified date of the document in the format <i>dd-Mmm-yy</i> . For example: 21-Mar-96. AltaVista determines a document's last-modified date in the following ways: If the web server, when serving the document to AltaVista's web page gatherer provides a <i>last-modified</i> date in the header information, AltaVista uses this. If not, AltaVista records the date and time it downloaded the document and stamps it with this last-modified date.
l		Two-letter ISO 639 standard ID of the language used in the document. Possible values are the same as the values for the kl parameter (Refer to Section 3.1.12) excluding XX.
L		The encoding of the document. It is not present if the encoding is the default (iso88591). Possible values are the same as the values for the enc parameter (Refer to Section 3.1.13).
SiteID		Sixteen-character site code. If present, the site was compressed.

4.1.3 Example VAL Text Search without Related Search and without RealNames

An example of the Simple Search query `&pg=q&q=world+%2Bseries` for customers that have not purchased the *Related Search* and *RealNames* search options is:

```
M=1684695
C=2
K=series
c=5907000
u=y
K=world
c=19584640
u=y
I=1-10
```

```
T=World Billfish Series: Presidential Candidates Campaign for Sailfish
u=http://www.worldbillfishseries.com/news/20000103/usdcse/20000103.CLPST.Prelud
e.html
a=Presidential Candidates Campaign for Sailfish. Jill Zima Borski. ISLAMORADA,
Florida, January 3, 2000 -- Anglers Jim E. Fakler, Jr. and Philip...
```

```
s=11938
r=8385
d=17-Jan-00
l=en
SiteID=34c5e71f9ba7dcd3
```

. . .

4.1.4 Example VAL Text Search with Related Search and without RealNames

An example of the Simple Search query `&pg=q&q=world+%2Bseries` for customers that have purchased the *Related Search* option but not the *RealNames* search option is:

```
M=1684695
C=2
K=series
c=5907000
u=y
K=world
c=19584640
u=y
I=1-10
RelatedTopic=world series tickets
RelatedOptions=/cgi-bin/query?q=world+series+tickets&pg=q&qe
RelatedTopic=1999 world series
RelatedOptions=/cgi-bin/query?q=1999+world+series&pg=q&qe
RelatedTopic=world series winners
RelatedOptions=/cgi-bin/query?q=world+series+winners&pg=q&qe
RelatedTopic=World Series Baseball
RelatedOptions=/cgi-bin/query?q=World+Series+Baseball&pg=q&qe
RelatedTopic=world series history
RelatedOptions=/cgi-bin/query?q=world+series+history&pg=q&qe
RelatedTopic=who won the world series
RelatedOptions=/cgi-bin/query?q=who+won+the+world+series&pg=q&qe
RelatedTopic=world series of poker
RelatedOptions=/cgi-bin/query?q=world+series+of+poker&pg=q&qe
RelatedTopic=+title:World Series +1979
RelatedOptions=/cgi-bin/query?q=%2Btitle%3aWorld+Series+%2b1979&pg=q&qe
```

```
T=CNN - Epic World Series matchup has fans fired up - October 18, 1996
u=http://www.cnn.com/US/9610/18/series.fever/
a=Epic World Series matchup has fans fired up. October 18, 1996 Web posted
at:10:00 p.m.EDT. NEW YORK (CNN) -- From the 1994 players' strike to the...
s=8008
r=9865
d=11-Jan-00
l=en
```

. . .

4.1.5 Example VAL Text Search without Related Search and with RealNames

An example of the Simple Search query `&pg=q&q=car` for customers that have not purchased the *Related Search* option but have purchased the *RealNames* search option is:

```
M=1175269
C=1
K=car
c=7313540
u=y
I=1-10
RNKey=car
RNSite=http://jump.altavista.com/rns.go?car+
```

RNPerfectMatch=n

T=Car Clubs World Wide
 u=http://carclubs.com/
 a=Car Clubs World Wide. Click here to Find A Club. or List your Club FREE.
 Click and View our Sponsors. [Club listings] [Web Friends] [Hot Links] [Hot...
 s=16183
 r=3278
 d=05-Dec-99
 l=en

. . .

4.1.6 Example VAL Text Search with Related Search and with RealNames

An example of the Simple Search query `&pg=q&q=car` for customers that have purchased both the *Related Search* and the *RealNames* search options is:

M=1175269
 C=1
 K=car
 c=7313540
 u=y
 I=1-10
 RelatedTopic=car rental
 RelatedOptions=/cgi-bin/query?q=car+rental&pg=q&qe
 RelatedTopic=new cars
 RelatedOptions=/cgi-bin/query?q=new+cars&pg=q&qe
 RelatedTopic=sports car
 RelatedOptions=/cgi-bin/query?q=sports+car&pg=q&qe
 RelatedTopic=used car prices
 RelatedOptions=/cgi-bin/query?q=used+car+prices&pg=q&qe
 RelatedTopic=car audio
 RelatedOptions=/cgi-bin/query?q=car+audio&pg=q&qe
 RelatedTopic=car insurance
 RelatedOptions=/cgi-bin/query?q=car+insurance&pg=q&qe
 RelatedTopic=rc cars
 RelatedOptions=/cgi-bin/query?q=rc+cars&pg=q&qe
 RelatedTopic=car stereo
 RelatedOptions=/cgi-bin/query?q=car+stereo&pg=q&qe
 RelatedTopic=new car prices
 RelatedOptions=/cgi-bin/query?q=new+car+prices&pg=q&qe
 RelatedTopic=kit cars
 RelatedOptions=/cgi-bin/query?q=kit+cars&pg=q&qe
 RelatedTopic=car prices
 RelatedOptions=/cgi-bin/query?q=car+prices&pg=q&qe
 RelatedTopic=car parts
 RelatedOptions=/cgi-bin/query?q=car+parts&pg=q&qe
 RelatedTopic=used car
 RelatedOptions=/cgi-bin/query?q=used+car&pg=q&qe
 RelatedTopic=budget rent a car
 RelatedOptions=/cgi-bin/query?q=budget+rent+a+car&pg=q&qe
 RNKey=car
 RNSite=http://jump.altavista.com/rns.go?car+
 RNPerfectMatch=n

T=Car Clubs World Wide
 u=http://carclubs.com/
 a=Car Clubs World Wide. Click here to Find A Club. or List your Club FREE.
 Click and View our Sponsors. [Club listings] [Web Friends] [Hot Links] [Hot...
 s=16183
 r=3278
 d=05-Dec-99
 l=en

. . .

4.1.7 Example VAL Text with `&enc=utf8`

This is the result of a query for münchen, with münchen having been sent to partners.altavista.com as `&q=mÃ¼nchen%20&enc=utf8&kl=de` (the UTF-8 encoded form of münchen). This example is without *RealNames* and without *Related Search* options.

```
M=1430720
C=1
K=münchen
c=2409865
u=y
I=1-10
```

```
T=Technische Universität München
u=http://www.tu-muenchen.de/
a=NEU: UnivIS: Das Universitäts-Informationssystem! Vorlesungsverzeichnis,
Stundenpläne, Personen- und Telefonnummernsuche und mehr! Aktuelles. -...
s=2764
r=14316
d=07-Apr-00
l=de
L=cp1252
SiteID=1f8833594b13634e
```

. . .

4.2 VAL Multimedia Search Output

4.2.1 VAL Multimedia Search Output Header

Key	Range	Description
M	0...MaxL	Total number of matches found
Available		Number of entries that are available to be returned
InputParam		The query input parameters

4.2.1.1 Related Search Output

When a *Multimedia Search* query is issued, and if a customer purchases the *Related Search* service, then *RelatedTopic* and *RelatedOptions* are returned (see the example in Section 4.2.5). *RelatedOptions* are the parameters that can be included in a subsequent query to get results for the *RelatedTopic*. The Related Search service is only available for image searches (`stype=simage`). These items are not returned if the family friendly filter (see Section 3.2.3) is on.

Key	Description
RelatedTopic	Related Search Topic to be displayed to the user.
RelatedOptions	Related Search query options. This CGI-query string can be concatenated with the user's dp and who parameters to request results for the related topic.

4.2.2 VAL Multimedia Search Results Output

Key	Description
T	Title
a	Abstract (Provided if available)
s	(Size) Reserved; value is currently set to 0.
u	URL of page that contains the object.
FrameCount	Frame count, 0 if N/A
Similar	Similar images are available in the database. The returned value is <i>y</i> or <i>n</i> for yes or no. Currently applies only to image searches. (stype=simage).
MmPage	Multimedia 20 digit page number. Returned if the page contains additional multimedia objects.
ClipTitle	Clip title (Provided if available)
ClipAuthor	Clip author (Provided if available)
ClipCopyright	Clip copyright (Provided if available)
ClipAbstract	Clip abstract (Provided if available)
ClipFormat	Clip format (for example: jpeg, gif89a, mp3, snd, et cetera.)
ClipDuration	Clip duration in seconds, 0 if N/A
ClipChannels	Clip channels, 0 if N/A
ClipSampleBits	Clip sample bits, 0 if N/A
ClipSampleRate	Clip sample rate in frames per second, 0 if N/A
MmType	Multimedia type (values explained in 4.2.2.1)
MmClass	Multimedia class (values explained in 4.2.2.2)
MmSize	Image or Frame size in bytes
MmHeight	Image or Frame height in pixels
MmWidth	Image or Frame width in pixels
MmFileName	File name
MmID	Multimedia object ID. Concatenate to TNailURL to point to thumbnail image
TNailURL	Base thumbnail URL. Concatenate MmID to it to point to thumbnail image.
MmURL	Multimedia URL (reserved, not currently used)
TNailType	Thumb nail type (values explained in 4.2.2.1)
TNailSize	Thumb nail size in bytes
TNailHeight	Thumb nail height in pixels
TNailWidth	Thumb nail width in pixels
VideoFrameRate	Video frame rate in frames per second
VideoDepth	Video depth in bits, 0 if N/A

4.2.2.1 MmType Definition

Value	File Type
0	Unknown
1	JPEG
2	GIF
3	GIF89a

Value	File Type
4	JPGProg
5	PNG
6	MPEG
7	AVI
8	ASF
9	MOV
10	AU
11	AIFF
12	RealVideo
13	RealAudio
14	MIDI
15	WAV
16	MP3
17	MPEG audio
18	MOV audio
19	AVI audio
20	ASF audio
21	Liquid audio

4.2.2.2 MmClass Values

MmClass uses bits to assign certain characteristics to an image, such as: color, gray, or black and white. For each characteristic, a mask is supplied as well as values for each field. Unless otherwise indicated, these characteristics apply to images (as opposed to video/audio objects). Bit 0 is the least significant bit. The information stored is:

Bit(s)	Characteristic	Value and Description
0, 1	Color	0: not used; 1: Black and White; 2: Gray; 3: Color
2	Medium	0: Photo; 1: Graphic
3	Transparent	0: Not transparent; 1: Is transparent
4	Omit	0: Use image; 1: Omit image from index
5	Animation	0: Not animated; 1: Is animated
6	Very Wide	0: Not wide; 1: Is wide (may be a horizontal line)
7	Very Tall	0: Not tall; 1: Is tall (may be a vertical line)
8	Very Small	0: Not small; 1: Is small (less than 10x10 pixels)
9	Click Through	0: Allow click through to image directly; 1: Don't allow
10, 11	Media Type	0: Image; 1: Audio; 2: Video; 3: Unknown

4.2.3 Example VAL Multimedia Audio Search

An example of the *Multimedia Audio Search* query

```
&pg=q&stype=saudio&pg=q&q=car&amp3=1&awav=1&astrm=1&aotr=1&mmW=1&mmF=1&macat=0
is:
```

```
M=1156
Available=1675
InputParam=pg=q&stype=saudio&q=car&ipt=1&iclr=1&ibw=1&iexc=1&mmW=1&micat=0&dp=
val&who=issid
```

```

T=Analyst Perspective: The Web As Car Dealership
a=So you can buy anything from baby products to perfume on the Web- but what
about cars? Tom Healy, partner at JD Power and Associates, says purchasing a
car is a rather emotionally invested
s=0
u=http://on24.av.com/vuwindow/scripts/vuwin.asp?id=1872&type=av&ref=altam&cb=al
ta
FrameCount=0
ClipTitle=Cars
ClipAuthor=Alona@on24.com
ClipCopyright=©1999
ClipFormat=asp
ClipDuration=98
ClipChannels=1
ClipSampleBits=16
ClipSampleRate=16000
MmType=0
MmClass=1024
MmSize=0
MmHeight=0
MmWidth=0
MmFileName=on24ram.asp?cn=1872&ref=altam&cb=alta
MmID=58405162
TNailURL=http://thumb-2.image.altavista.com/image/
TNailType=1
TNailSize=0
TNailHeight=0
TNailWidth=0
VideoFrameRate=0
VideoDepth=0

```

4.2.4 Example VAL Multimedia Image Search without Related Search

An example of the *Multimedia Image Search* query

```
&pg=q&stype=simage&q=cars&ipht=1&iclr=1&ibw=1&iexc=1&mmW=1&micat=0
```

for customers that have not purchased the *Related Search* option is:

```

M=11444
Available=1000
InputParam=pg=q&stype=simage&q=cars&ipht=1&iclr=1&ibw=1&iexc=1&mmW=1&micat=0&dp
=val&who=issid

```

```

T=Need for Speed: High Stakes
a=
s=0
u=http://www.needforspeed.com/hs_psx/showroom.asp?3
FrameCount=0
ClipFormat=jpeg
ClipDuration=0
ClipChannels=0
ClipSampleBits=0
ClipSampleRate=0
MmType=1
MmClass=3
MmSize=8413
MmHeight=90
MmWidth=160
MmURL=
MmFileName=cars_f550_01.jpg
MmID=42942603
TNailURL=http://thumb-2.image.altavista.com/image/

```

```

TNailType=1
TNailSize=3376
TNailHeight=115
TNailWidth=115
VideoFrameRate=0
VideoDepth=0

```

4.2.5 Example VAL Multimedia Image Search with Related Search

An example of the *Multimedia Image Search* query

```
&pg=q&stype=simage&q=house&ipht=1&igrph=1&iclr=1&ibw=1&mmW=1&micat=0&FFF=0
```

for customers that have purchased the *Related Search* option is:

```

M=95129
Available=1000
InputParam=pg=q&stype=simage&q=house&ipht=1&igrph=1&iclr=1&ibw=1&mmW=1&micat=0&
FFF=0&dp=val&who=issid
RelatedTopic=white house
RelatedOptions=/cgi-bin/query?q=white+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=bill gates house
RelatedOptions=/cgi-bin/query?q=bill+gates+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=sydney opera house
RelatedOptions=/cgi-
bin/query?q=sydney+opera+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=full house
RelatedOptions=/cgi-bin/query?q=full+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=haunted house
RelatedOptions=/cgi-bin/query?q=haunted+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=house plants
RelatedOptions=/cgi-bin/query?q=house+plants&mmdo=3&stype=simage&origin=qe
RelatedTopic=light house
RelatedOptions=/cgi-bin/query?q=light+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=house wives
RelatedOptions=/cgi-bin/query?q=house+wives&mmdo=3&stype=simage&origin=qe
RelatedTopic=old house
RelatedOptions=/cgi-bin/query?q=old+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=house plans
RelatedOptions=/cgi-bin/query?q=house+plans&mmdo=3&stype=simage&origin=qe
RelatedTopic=house fire
RelatedOptions=/cgi-bin/query?q=house+fire&mmdo=3&stype=simage&origin=qe
RelatedTopic=victorian houses
RelatedOptions=/cgi-bin/query?q=victorian+houses&mmdo=3&stype=simage&origin=qe
RelatedTopic=gingerbread house
RelatedOptions=/cgi-bin/query?q=gingerbread+house&mmdo=3&stype=simage&origin=qe
RelatedTopic=animal house
RelatedOptions=/cgi-bin/query?q=animal+house&mmdo=3&stype=simage&origin=qe

s=0
u=http://aurora.phys.utk.edu/~twebber/house.html
FrameCount=0
Similar=n
MmPage=2604673609426170320
ClipFormat=gif
ClipDuration=0
ClipChannels=0
ClipSampleBits=0
ClipSampleRate=0
MmType=2
MmClass=15
MmSize=35918
MmHeight=298
MmWidth=216

```

```
MmFileName=house9.gif
MmID=54395291
TNailURL=http://thumb-2.image.altavista.com/image/
TNailType=1
TNailSize=3439
TNailHeight=115
TNailWidth=115
VideoFrameRate=0
VideoDepth=0
```

4.2.6 Example VAL Multimedia Video Search

An example of the *Multimedia Video Search* query

```
&stype=svideo&pg=q&q=car&vavi=1&vmpeg=1&vmov=1&vstrm=1&vothr=1&mmW=1&mvcat=0
is:

M=4592
Available=1331
InputParam=pg=q&stype=svideo&q=car&ipht=1&iclr=1&ibw=1&iexc=1&mmW=1&micat=0&dp=
val&who=issid

T=CNN 4/27/99 3:48:42 PM 0:36
a=A car theft in progress. Every day hundreds of cars are stolen in the united
states, 1.4 Million in 1997 alone. Of those, 450,000 were never recovered.
Tuesday the fbi announced a nationwide
s=0
u=http://videocast.av.com/default.cfm?playclip=cnn_1999_0427_15484236
FrameCount=0
ClipTitle=CNN
ClipAuthor=FastTV
ClipCopyright=(c)1998, 1999 FastTV Inc.
ClipFormat=ram
ClipDuration=37
ClipChannels=1
ClipSampleBits=16
ClipSampleRate=8000
MmType=1
MmClass=2051
MmSize=6442
MmHeight=180
MmWidth=240
MmFileName=15484236.rm
MmID=58467792
TNailURL=http://thumb-2.image.altavista.com/image/
TNailType=1
TNailSize=2904
TNailHeight=115
TNailWidth=115
VideoFrameRate=0
VideoDepth=12
```

5 ISS XML Output Format

ISS supports two different output formats for query results: VAL (the original format created for the ISS product) and XML (a format introduced in the 2.0 ISS release). Customers specify the desired output format for each individual query via the **dp** parameter (see Sections 3.1.2 and 3.2.2). The same information is returned regardless of output format type selection. This section describes the XML output format. See Section 4 for a description of the VAL output format.

XML output consists of three parts: the HTTP header, the XML header, and the optional XML search results. Each part is separated from the next part by a single blank line.

The HTTP header identifies the HTTP version, the HTTP status, the result date, the AltaVista server version, the MIME version, the content length, and the content type (text/xml). The following is an example of the HTTP header:

```
HTTP/1.0 200 OK
Date: Fri, 02 Jun 2000 22:41:54 GMT
Server: AV/1.0.1
MIME-Version: 1.0
Content-Length: 4841
Content-Type: text/xml
```

The format of the XML headers and the XML search results are defined by ISS's SGML Document Type Definition (DTD) file. Excerpts from the DTD file are included in this section. The complete DTD file is listed in Section 6. This section addresses the values and meanings of each of the *key=value* pairs returned in the XML output. The keys are case sensitive.

The optional XML search results may be repeated more than once.

The AltaVista front-end servers return the XML response for a search result. The XML format is designed to be very similar to the VAL output format (see Section 4). Within the AltaVista XML response, built-in XML entities (&, <, >, ") are replaced as follows:

Character	Replaced with
&	&
"	"
<	<
>	>

Customers may elect to utilize XML parsers at their discretion and any of the various XML parsers may be used as desired. AltaVista has run tests of our ISS XML output with the following XML parsers and lists them here for informational purposes only. Refer to <http://xml.com> for information about parser conformance and support of industry standards like DOM Level 1, Level 2, SAX1, and SAX2.

Vendor	XML Parser	Type	Support
IBM	Xerces Java 1.0.4	Validation	DOM Level 2 and SAX2
Sun Microsystems	Java Project X Release 2	Validation	DOM Level 1 and SAX1
Microsoft	MSXML3	Validation	DOM Level 1.

The following important points should be remembered when processing XML output:

- Ignore any unknown *keys*. This will prevent the code from breaking when a new feature is added.
- *Do not* depend on the ordering of *key=value* pairs within each section.
- In all references, MaxI represents the value $+2^{31-1}$, MaxL the value $\pm 2^{63-1}$, and MinI the value -2^{31-1} .

This section is divided into two subsections, one for each of the ISS search products: *Text* and *Multimedia*.

5.1 XML Text Search Output

5.1.1 XML Text Search Output Header

The header is separated from the results by a single blank line. In addition, the lines for each result are separated from the other results by a blank line. All header information is formatted in standard XML tags.

Key	Range	Description
Match	0...MaxL	Total number of matches found
NumberOfKW	0...MaxI	Number of keywords or phrases in the query.
ErrorNumber	1...8	An error number. Present if there is an error
ErrorString		An error string associated with the above e=<error number>. Present if there is an error. Possible values are: <ul style="list-style-type: none"> e=1. E=internal error e=2. E=no service available e=3. E=syntax error e=4. E=query too long e=5. E=Incorrect ID for Internet Search Services. e=6. E=Missing ID for Internet Search Services. e=7. E=The service is not available to you; please contact your sales representative. e=8. E=The server is unavailable. Please try again.
SpellCheck		If present, is a list of words, separated by commas (,), which are spell check suggestions. Both the query string (see Section 3.1.6) and the Advanced Search ranking string (see Section 3.1.7) are spell checked.
SiteName		Site name. Only returned for compressed site expansion requests (see si parameter in Section 3.1.15).
Keywords		A keyword or phrase. For <i>Simple Searches</i> (pg=q) the search string (q=text) is used. For <i>Advanced Searches</i> (pg=aq) the ranking string (r=text) is used. For <i>Advanced Searches</i> (pg=aq) the keyword or phrase is not returned if the ranking string (r=text) was omitted. Each Keyword value can have the following options on the same line: Occurrences and KWUsed . Not returned if there are no results returned (i.e. M=0)
Occurrences	0...MaxL	For the preceding keyword or phrase, total number of occurrences of the keyword or phrase in the entire index. If absent, default is 0. Not returned for queries containing the si parameter. For <i>Advanced Searches</i> (pg=aq) parameter is not returned if the ranking string (r=text) was omitted.
KWUsed	y, n	For the preceding keyword or phrase, indicates whether the keyword or phrase was used (y) or ignored (n). For <i>Advanced Searches</i> (pg=aq) parameter is not returned if the ranking string (r=text) was omitted.

Key	Range	Description
Range	1...MaxI	Indicates the range of the document results sent back by the search engine. This is represented by 2 integers separated by a hyphen (-). For example: <Range>11-20</Range> indicates that there are 10 results returned and that they are results 11 through 20. The number of results returned depends upon the total number of matches found (M above), and the setting of the nbq parameter (see Section 3.1.8). The range of the results is determined by the stq parameter (see Section 3.1.9). The Range key is not returned if there are no results returned (i.e. M=0).

For example, the header returned for a *Simple Search* query of “apple pie” filling,

&pg=q&q=%22apple+pie%22+%2Bfilling is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<WebSearch>
<WebHeader>
  <Match>75734</Match>
  <NumberOfKW>2</NumberOfKW>
  <KWOccurrences Keywords="apple pie" Occurrences="19284" KWUsed="y" />
  <KWOccurrences Keywords="filling" Occurrences="441844" KWUsed="y" />
  <Range>1-10</Range>
</WebHeader>
. . .
</WebSearch>
</Val>
```

For example, the header returned for a *Simple Search* query of “serch engines”,

&pg=q&q=%22serch+engines%22 is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<WebSearch>
<WebHeader>
  <Match>650</Match>
  <NumberOfKW>1</NumberOfKW>
  <SpellCheck>search</SpellCheck>
  <KWOccurrences Keywords="serch engines" Occurrences="824" KWUsed="y" />
  <Range>1-10</Range>
</WebHeader>
. . .
</WebSearch>
</Val>
```

Note: The spell check <SpellCheck>search</SpellCheck> above.

For example, the header returned for the *Advanced Search* Boolean expression of *cat NOT dog*,

&pg=aq&q=cat+NOT+dog is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<VALError>
  <Match>0</Match>
  <NumberOfKW>0</NumberOfKW>
  <ErrorNumber>3</ErrorNumber>
  <ErrorString>syntax error</ErrorString>
</VALError>

</Val>
```

Note: This is a syntax error, cat AND NOT dog is valid, and will not return an error.

The header returned for the empty query “” is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<VALError>
  <Match>0</Match>
  <NumberOfKW>0</NumberOfKW>
</VALError>

</Val>
```

5.1.1.1 Related Search Output

When a *Simple Search* query (**pg=q**) is issued, and if a customer purchases *Related Search*, then *RelatedTopic* and *RelatedOptions* are returned in the Text Header (see the example in Section 5.1.5). *RelatedOptions* are the parameters that can be included in a subsequent query to get results for the *RelatedTopic*. Not returned for *Advanced Search* queries (**pg=aq**). Not returned if the family friendly filter (see Section 3.1.3) is on.

Key	Description
RelatedTopic	Related Search Topic to be displayed to the user.
RelatedOptions	Related Search query options. This CGI-query string can be concatenated with the user's dp and who parameters to request results for the related topic.

5.1.1.2 RealNames Output

When a *Simple Search* query (**pg=q**) is issued, and if a customer purchases the *RealNames* service and *RealNames* are available, the Text Header returns three extra fields: RNKey, RNSite, and RNPerfectMatch (see the example in Section 5.1.6). RNKey is the *RealNames* key word, RNSite is the URL address containing the *RealNames* information. Not returned for *Advanced Search* queries (**pg=aq**). Not returned if the family friendly filter (see Section 3.1.3) is on.

Key	Description
RNKey	RealNames Key
RNSite	the URL address containing RealNames information
RNPerfectMatch	value (y/n) indicates whether or not there is a perfect match

If there is no existing *RealNames* information for a given query, these keys are not returned in the output.

5.1.2 XML Text Search Results Output

Key	Range	Description
Title		Title of the document. It is always present and is listed first, but it may have an empty value
URL		The URL for the document. Always present
Abstract		Abstract of the document.
Size	0...MaxI	Size of the document, in bytes
Ranking	MinI... MaxI	The relevancy score for the document
Date		Last-modified date of the document in the format <i>dd-Mmm-yy</i> . For example: 21-Mar-96. AltaVista determines a document's last-modified date in the following ways: If the web server, when serving the document to AltaVista's web page gatherer (Scooter), provides a <i>last-modified</i> date in the header information, AltaVista uses this. If not, AltaVista records the date and time it downloaded the document and stamps it with this last-modified date.
Language		Two-letter ISO 639 standard ID of the language used in the document. Possible values are the same as the values for the kl parameter (Refer to Section 3.1.12) excluding XX.
Encoding		The encoding of the document. It is not present if the encoding is the default (iso88591). Possible values are the same as the values for the enc parameter. Refer to Section 3.1.13.
SiteID		Sixteen-character site code. If present, the site was compressed.

5.1.3 XML Text Search DTD Grammar

The following extracted portion of the DTD file shows the grammar of the *Text Search XML* output. Refer to Section 6 for a complete listing of the file.

```
<!ELEMENT WebHeader (Match?, NumberOfKW?, SpellCheck?, SiteName?,
    KWOccurrences*, Range?, RelatedSearch*, AskJeeves*, RealName*)>

<!ELEMENT KWOccurrences EMPTY>
<!ELEMENT RelatedSearch EMPTY>
<!ELEMENT AskJeeves EMPTY>
<!ELEMENT RealName EMPTY>

<!ELEMENT WebEntry (Title?, URL?, Abstract?, Size?, Ranking?, Date?, Language?,
    Encoding?, SiteID?, FactsheetID?)>

<!ATTLIST KWOccurrences Keywords CDATA #REQUIRED>
<!ATTLIST KWOccurrences Occurrences CDATA #REQUIRED>
<!ATTLIST KWOccurrences KWUsed (y|n) "y">

<!ATTLIST RelatedSearch RelatedTopic CDATA #REQUIRED>
<!ATTLIST RelatedSearch RelatedOptions CDATA #REQUIRED>

<!ATTLIST AskJeeves AJQuestion CDATA #REQUIRED>
<!ATTLIST AskJeeves AJAnswer CDATA #REQUIRED>

<!ATTLIST RealName RNKey CDATA #REQUIRED>
<!ATTLIST RealName RNSite CDATA #REQUIRED>
<!ATTLIST RealName RNPerfectMatch (y|n) "y">
```

5.1.4 Example XML Text Search without Related Search and without RealNames

An example of the *Simple Search* query `&pg=q&q=world+%2Bseries` for customers that have not purchased the *Related Search* and *RealNames* search options is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<WebSearch>
<WebHeader>
  <Match>1646127</Match>
  <NumberOfKW>2</NumberOfKW>
  <KWOccurrences Keywords="series" Occurrences="2953500" KWUsed="y" />
  <KWOccurrences Keywords="world" Occurrences="9792320" KWUsed="y" />
  <Range>1-10</Range>
</WebHeader>

<WebEntry>
  <Title>CNN - Epic World Series matchup has fans fired up - October 18,
1996</Title>
  <URL>http://www.cnn.com/US/9610/18/series.fever/</URL>
  <Abstract>Epic World Series matchup has fans fired up. October 18, 1996 Web
posted at:10:00 p.m.EDT. NEW YORK (CNN) -- From the 1994 players' strike to
the...</Abstract>
  <Size>8008</Size>
  <Ranking>7998</Ranking>
  <Date>11-Jan-00</Date>
  <Language>en</Language>
</WebEntry>

. . .

</WebSearch>

</Val>
```

5.1.5 Example XML Text Search with Related Search and without RealNames

An example of the *Simple Search* query `&pg=q&q=car` for customers that have purchased the *Related Search* option but not the *RealNames* search option is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<WebSearch>
<WebHeader>
  <Match>1414798</Match>
  <NumberOfKW>1</NumberOfKW>
  <KWOccurrences Keywords="car" Occurrences="3656770" KWUsed="y" />
  <Range>1-10</Range>
  <RelatedSearch RelatedTopic="car rental" RelatedOptions="/cgi-
bin/query?q=car+rental&pg=q&qe" />
  <RelatedSearch RelatedTopic="new cars" RelatedOptions="/cgi-
bin/query?q=new+cars&pg=q&qe" />
  <RelatedSearch RelatedTopic="sports car" RelatedOptions="/cgi-
bin/query?q=sports+car&pg=q&qe" />
  <RelatedSearch RelatedTopic="used car prices" RelatedOptions="/cgi-
bin/query?q=used+car+prices&pg=q&qe" />
  <RelatedSearch RelatedTopic="car audio" RelatedOptions="/cgi-
bin/query?q=car+audio&pg=q&qe" />
  <RelatedSearch RelatedTopic="car insurance" RelatedOptions="/cgi-
bin/query?q=car+insurance&pg=q&qe" />
```

```

    <RelatedSearch RelatedTopic="rc cars" RelatedOptions="/cgi-
bin/query?q=rc+cars&pg=q&qe" />
    <RelatedSearch RelatedTopic="car stereo" RelatedOptions="/cgi-
bin/query?q=car+stereo&pg=q&qe" />
    <RelatedSearch RelatedTopic="new car prices" RelatedOptions="/cgi-
bin/query?q=new+car+prices&pg=q&qe" />
    <RelatedSearch RelatedTopic="kit cars" RelatedOptions="/cgi-
bin/query?q=kit+cars&pg=q&qe" />
    <RelatedSearch RelatedTopic="car prices" RelatedOptions="/cgi-
bin/query?q=car+prices&pg=q&qe" />
    <RelatedSearch RelatedTopic="car parts" RelatedOptions="/cgi-
bin/query?q=car+parts&pg=q&qe" />
    <RelatedSearch RelatedTopic="used car" RelatedOptions="/cgi-
bin/query?q=used+car&pg=q&qe" />
    <RelatedSearch RelatedTopic="budget rent a car" RelatedOptions="/cgi-
bin/query?q=budget+rent+a+car&pg=q&qe" />

<WebEntry>
  <Title>Car Clubs World Wide</Title>
  <URL>http://carclubs.com/</URL>
  <Abstract>Car Clubs World Wide. Click here to Find A Club. or List your Club
FREE. Click and View our Sponsors. [Club listings] [Web Friends] [Hot Links]
[Hot...</Abstract>
  <Size>16183</Size>
  <Ranking>2647</Ranking>
  <Date>05-Dec-99</Date>
  <Language>en</Language>
</WebEntry>

. . .

</WebSearch>

</Val>

```

5.1.6 Example XML Text Search without Related Search and with RealNames

An example of the *Simple Search* query `&pg=q&q=car` for customers that have not purchased the *Related Search* option but have purchased the *RealNames* search option is:

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<WebSearch>
<WebHeader>
  <Match>1414798</Match>
  <NumberOfKW>1</NumberOfKW>
  <KWOccurrences Keywords="car" Occurrences="3656770" KWUsed="y" />
  <Range>1-10</Range>
  <RealName RNKey="car " RNSite="http://jump.altavista.com/rns.go?car+"
RNPerfectMatch="n" />
</WebHeader>

<WebEntry>
  <Title>Car Clubs World Wide</Title>
  <URL>http://carclubs.com/</URL>
  <Abstract>Car Clubs World Wide. Click here to Find A Club. or List your Club
FREE. Click and View our Sponsors. [Club listings] [Web Friends] [Hot Links]
[Hot...</Abstract>
  <Size>16183</Size>
  <Ranking>2647</Ranking>
  <Date>05-Dec-99</Date>
  <Language>en</Language>
</WebEntry>

```

```

<WebEntry>
  <Title>New car prices, used cars, auto insurance quotes, auto classifieds,
factory re</Title>
  <URL>http://www.carbuyer.com/</URL>
  <Abstract>New car prices, auto insurance quotes, truck, van, motorcycle
classified listing service...</Abstract>
  <Size>9681</Size>
  <Ranking>2613</Ranking>
  <Date>24-Dec-97</Date>
  <Language>en</Language>
</WebEntry>

. . .

</WebSearch>

</Val>

```

5.1.7 Example XML Text Search with Related Search and with RealNames

An example of the *Simple Search* query `&pg=q&q=apple` for customers that have purchased both the *Related Search* and *RealNames* search options is:

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<WebSearch>
<WebHeader>
  <Match>346180</Match>
  <NumberOfKW>1</NumberOfKW>
  <KWOccurrences Keywords="apple" Occurrences="910726" KWUsed="y" />
  <Range>1-10</Range>
  <RelatedSearch RelatedTopic="fiona apple" RelatedOptions="/cgi-
bin/query?q=fiona+apple&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="Apple Computer" RelatedOptions="/cgi-
bin/query?q=Apple+Computer&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple vacations" RelatedOptions="/cgi-
bin/query?q=apple+vacations&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="candy apples" RelatedOptions="/cgi-
bin/query?q=candy+apples&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="Apple Macintosh" RelatedOptions="/cgi-
bin/query?q=Apple+Macintosh&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple pie" RelatedOptions="/cgi-
bin/query?q=apple+pie&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple cider vinegar" RelatedOptions="/cgi-
bin/query?q=apple+cider+vinegar&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple g4" RelatedOptions="/cgi-
bin/query?q=apple+g4&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple tree" RelatedOptions="/cgi-
bin/query?q=apple+tree&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple crisp" RelatedOptions="/cgi-
bin/query?q=apple+crisp&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="Apple iMac" RelatedOptions="/cgi-
bin/query?q=Apple+iMac&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple newton" RelatedOptions="/cgi-
bin/query?q=apple+newton&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="apple cider" RelatedOptions="/cgi-
bin/query?q=apple+cider&amp;pg=q&amp;qe" />
  <RelatedSearch RelatedTopic="Apple Quick Time" RelatedOptions="/cgi-
bin/query?q=Apple+Quick+Time&amp;pg=q&amp;qe" />
  <RealName RNKey="Apple" RNSite="http://jump.altavista.com/rns.go?Apple"
RNPerfectMatch="y" />
</WebHeader>

```

```

<WebEntry>
  <Title>Apple Computer Australia</Title>
  <URL>http://www.apple.com.au/</URL>
  <Abstract>11 January 2000. AppleStore Australia | Hot News | Corporate | Sign
Me Up! | Products | Support | Design & Publishing | Education | About
Apple |...</Abstract>
  <Size>9618</Size>
  <Ranking>7979</Ranking>
  <Date>01-Jan-70</Date>
  <Language>en</Language>
  <SiteID>4a228f033a5e79ee</SiteID>
</WebEntry>

. . .

</WebSearch>

</Val>

```

5.2 XML Multimedia Search Output

5.2.1 XML Multimedia Search Output Header

Key	Range	Description
Match	0...MaxL	Total number of matches found
Available		Number of entries that are available to be returned
InputParam		The query input parameters

5.2.1.1 Related Search Output

When a *Multimedia Search* query is issued, and if a customer purchases *Related Search*, *RelatedTopic* and *RelatedOptions* are returned (see the example in section 5.2.6). *RelatedOptions* are the parameters that can be included in a subsequent query to get results for the *RelatedTopic*. These are not returned if the family friendly filter (see Section 3.2.3) is on. These are only returned if there is a related search relevant to the user's query.

Key	Description
RelatedTopic	Related Search Topic to be displayed to the user.
RelatedOptions	Related Search query options. This CGI-query string can be concatenated with the user's dp and who parameters to request results for the related topic.

5.2.2 XML Multimedia Search Results Output

Key	Description
Title	Title
Abstract	Abstract (Provided if available)
Size	(Size) Reserved; value is currently set to 0.
URL	URL of page that contains the object.
FrameCount	Frame count, 0 if N/A
Similar	Similar images are available in the database. The returned value is <i>y</i> or <i>n</i> for yes or no. Currently applies only to image searches. (stype=simage).

Key	Description
MmPage	Multimedia 20 digit page number. Returned if the page contains additional multimedia objects.
ClipTitle	Clip title (Provided if available)
ClipAuthor	Clip author (Provided if available)
ClipCopyright	Clip copyright (Provided if available)
ClipAbstract	Clip abstract (Provided if available)
ClipFormat	Clip format (for example: jpeg, gif89a, mp3, snd, et cetera.)
ClipDuration	Clip duration in seconds, 0 if N/A
ClipChannels	Clip channels, 0 if N/A
ClipSampleBits	Clip sample bits, 0 if N/A
ClipSampleRate	Clip sample rate in frames per second, 0 if N/A
MmType	Multimedia type (values explained in 5.2.2.1)
MmClass	Multimedia class (values explained in 5.2.2.2)
MmSize	Image or Frame size in bytes
MmHeight	Image or Frame height in pixels
MmWidth	Image or Frame width in pixels
MmURL	File name
MmFileName	Multimedia object ID. Concatenate to TNailURL to point to thumbnail image
MmID	Base thumbnail URL. Concatenate MmID to it to point to thumbnail image.
TNailURL	Multimedia URL (reserved, not currently used)
TNailType	Thumb nail type (values explained in 5.2.2.1)
TNailSize	Thumb nail size in bytes
TNailHeight	Thumb nail height in pixels
TNailWidth	Thumb nail width in pixels
VideoFrameRate	Video frame rate in frames per second
VideoDepth	Video depth in bits, 0 if N/A

5.2.2.1 MmType Definition

Value	File Type
0	Unknown
1	JPEG
2	GIF
3	GIF89a
4	JPGProg
5	PNG
6	MPEG
7	AVI
8	ASF
9	MOV
10	AU
11	AIFF

Value	File Type
12	RealVideo
13	RealAudio
14	MIDI
15	WAV
16	MP3
17	MPEG audio
18	MOV audio
19	AVI audio
20	ASF audio
21	Liquid audio

5.2.2.2 MmClass Values

MmClass uses bits to assign certain characteristics to an image, such as: color, gray, or black and white. For each characteristic, a mask is supplied as well as values for each field. Unless otherwise indicated, these characteristics apply to images (as opposed to video/audio objects). Bit 0 is the least significant bit. The information stored is:

Bit(s)	Characteristic	Value and Description
0,1	Color	0: not used; 1: black and white; 2: gray; 3: color
2	Medium	0: Photo; 1: Graphic
3	Transparent	0: Not transparent; 1: Is transparent
4	Omit	0: Use image; 1: Omit image from index
5	Animation	0: Not animated; 1: Is animated
6	Very Wide	0: Not wide; 1: Is wide (may be a horizontal line)
7	Very Tall	0: Not tall; 1: Is tall (may be a vertical line)
8	Very Small	0: Not small; 1: Is small (less than 10x10 pixels)
9	Click Through	0: Allow click through to image directly; 1: Don't allow
10,11	MM type	0: Image; 1: Audio; 2: Video; 3: Unknown

5.2.3 XML Multimedia DTD Grammar

The following extracted portion of the DTD file shows the grammar of the *Multimedia Search* XML output. Refer to Section 6 for a complete listing of the file.

```
<!ELEMENT MmHeader (Match?, Available?, InputParam?, RelatedSearch*)>
<!ELEMENT MmEntry (Title?, Abstract?, Size?, URL?, FrameCount?, Similar?,
MmPage?,
  ClipTitle?, ClipAuthor?, ClipCopyright?,
  ClipAbstract?, ClipFormat?, ClipDuration?, ClipChannels?,
  ClipSampleBits?, ClipSampleRate?,
  MmType?, MmClass?, MmSize?, MmHeight?,
  MmWidth?, MmURL?, MmFileName?, MmID?,
  TNailURL?, TNailType?, TNailSize?, TNailHeight?, TNailWidth?,
  VideoFrameRate?, VideoDepth?
)>
```

5.2.4 Example XML Multimedia Audio Search in XML

An example of the *Multimedia Audio Search* query:

```
&pg=q&stype=saudio&pg=q&q=car&amp3=1&awav=1&astrm=1&aotr=1&mmW=1&mmF=1&macat=0
```

is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<MmSearch>
<MmHeader>
  <Match>2763</Match>
  <Available>1704</Available>

<InputParam>stype=saudio&amp;pg=q&amp;q=car&amp;amp3=1&amp;awav=1&amp;astrm=1&a
mp;aotr=1&amp;mmW=1&amp;mmF=1&amp;macat=0&amp;dp=xml&amp;who=issid</InputParam
>
</MmHeader>

<MmEntry>
  <Title>Getaway Car - The Silence</Title>
  <Size>0</Size>
  <URL>http://www.emusic.com/albums/3692/</URL>
  <FrameCount>0</FrameCount>
  <Similar>n</Similar>
  <ClipFormat>mp3</ClipFormat>
  <ClipDuration>29</ClipDuration>
  <ClipChannels>2</ClipChannels>
  <ClipSampleBits>16</ClipSampleBits>
  <ClipSampleRate>44100</ClipSampleRate>
  <MmType>16</MmType>
  <MmClass>1024</MmClass>
  <MmSize>0</MmSize>
  <MmHeight>0</MmHeight>
  <MmWidth>0</MmWidth>
  <MmFileName>Getaway_Car-Highway_Lost-7-The_Silence-Demo.mp3</MmFileName>
  <MmID>58443757</MmID>
  <TNailURL>http://thumb-2.image.altavista.com/image/</TNailURL>
  <TNailType>1</TNailType>
  <TNailSize>0</TNailSize>
  <TNailHeight>0</TNailHeight>
  <TNailWidth>0</TNailWidth>
  <VideoFrameRate>0</VideoFrameRate>
  <VideoDepth>0</VideoDepth>
</MmEntry>

. . .

</MmSearch>

</Val>
```

5.2.5 Example XML Multimedia Image Search without Related Search

An example of the *Multimedia Image Search* query

```
&pg=q&stype=simage&q=cars&ipt=1&iclr=1&ibw=1&iexc=1&mmW=1&micat=0
```

for customers that have not purchased the *Related Search* option is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<MmSearch>
<MmHeader>
```

```

    <Match>23937</Match>
    <Available>1000</Available>

<InputParam>stype=simage&amp;pg=q&amp;q=car&amp;ipht=1&amp;igrph=1&amp;ibb=1&amp;
p;iclr=1&amp;ibw=1&amp;mmW=1&amp;micat=0&amp;dp=xml&amp;who=issid</InputParam>
</MmHeader>

<MmEntry>
  <Title>Target Racing</Title>
  <Size>0</Size>
  <URL>http://beepbeep.target.com/news_results/results/index.html</URL>
  <FrameCount>0</FrameCount>
  <Similar>n</Similar>
  <MmPage>15643662249791225180</MmPage>
  <ClipFormat>gif</ClipFormat>
  <ClipDuration>0</ClipDuration>
  <ClipChannels>0</ClipChannels>
  <ClipSampleBits>0</ClipSampleBits>
  <ClipSampleRate>0</ClipSampleRate>
  <MmType>2</MmType>
  <MmClass>7</MmClass>
  <MmSize>4221</MmSize>
  <MmHeight>128</MmHeight>
  <MmWidth>58</MmWidth>
  <MmFileName>car_01b2.GIF</MmFileName>
  <MmID>67252800</MmID>
  <TNailURL>http://thumb-2.image.altavista.com/image/</TNailURL>
  <TNailType>1</TNailType>
  <TNailSize>2256</TNailSize>
  <TNailHeight>115</TNailHeight>
  <TNailWidth>115</TNailWidth>
  <VideoFrameRate>0</VideoFrameRate>
  <VideoDepth>0</VideoDepth>
</MmEntry>

. . .

</MmSearch>

</Val>

```

5.2.6 Example XML Multimedia Image Search with Related Search

An example of the *Multimedia Image Search* query

```
&pg=q&stype=simage&q=house&ipht=1&igrph=1&iclr=1&ibw=1&mmW=1&micat=0&FFF=0
```

for customers that have purchased the *Related Search* option is:

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<MmSearch>
<MmHeader>
  <Match>95129</Match>
  <Available>1000</Available>

<InputParam>stype=simage&amp;pg=q&amp;q=house&amp;ipht=1&amp;igrph=1&amp;ibb=1&amp;
&amp;iclr=1&amp;ibw=1&amp;mmW=1&amp;micat=0&amp;FFF=0&amp;dp=xml&amp;who=issid</
InputParam>
  <RelatedSearch RelatedTopic="white house" RelatedOptions="/cgi-
bin/query?q=white+house&amp;mmdo=3&amp;stype=simage&amp;origin=qe" />
  <RelatedSearch RelatedTopic="bill gates house" RelatedOptions="/cgi-
bin/query?q=bill+gates+house&amp;mmdo=3&amp;stype=simage&amp;origin=qe" />

```

```

    <RelatedSearch RelatedTopic="sydney opera house" RelatedOptions="/cgi-
bin/query?q=sydney+opera+house&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="full house" RelatedOptions="/cgi-
bin/query?q=full+house&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="haunted house" RelatedOptions="/cgi-
bin/query?q=haunted+house&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="house plants" RelatedOptions="/cgi-
bin/query?q=house+plants&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="light house" RelatedOptions="/cgi-
bin/query?q=light+house&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="house wives" RelatedOptions="/cgi-
bin/query?q=house+wives&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="old house" RelatedOptions="/cgi-
bin/query?q=old+house&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="house plans" RelatedOptions="/cgi-
bin/query?q=house+plans&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="house fire" RelatedOptions="/cgi-
bin/query?q=house+fire&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="victorian houses" RelatedOptions="/cgi-
bin/query?q=victorian+houses&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="gingerbread house" RelatedOptions="/cgi-
bin/query?q=gingerbread+house&mmdo=3&stype=simage&origin=qe" />
    <RelatedSearch RelatedTopic="animal house" RelatedOptions="/cgi-
bin/query?q=animal+house&mmdo=3&stype=simage&origin=qe" />
</MmHeader>

<MmEntry>
  <Size>0</Size>
  <URL>http://aurora.phys.utk.edu/~twebber/house.html</URL>
  <FrameCount>0</FrameCount>
  <Similar>n</Similar>
  <MmPage>2604673609426170320</MmPage>
  <ClipFormat>gif</ClipFormat>
  <ClipDuration>0</ClipDuration>
  <ClipChannels>0</ClipChannels>
  <ClipSampleBits>0</ClipSampleBits>
  <ClipSampleRate>0</ClipSampleRate>
  <MmType>2</MmType>
  <MmClass>15</MmClass>
  <MmSize>35918</MmSize>
  <MmHeight>298</MmHeight>
  <MmWidth>216</MmWidth>
  <MmFileName>house9.gif</MmFileName>
  <MmID>54395291</MmID>
  <TNailURL>http://thumb-2.image.altavista.com/image/</TNailURL>
  <TNailType>1</TNailType>
  <TNailSize>3439</TNailSize>
  <TNailHeight>115</TNailHeight>
  <TNailWidth>115</TNailWidth>
  <VideoFrameRate>0</VideoFrameRate>
  <VideoDepth>0</VideoDepth>
</MmEntry>

. . .

</MmSearch>

</Val>

```

5.2.7 Example XML Multimedia Video Search

An example of the *Multimedia Video Search* query

```
&stype=svideo&pg=q&q=car&vavi=1&vmpeg=1&vmov=1&vstrm=1&vothr=1&mmW=1&mvcat=0
```

is:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
<Val>
<MmSearch>
<MmHeader>
  <Match>5884</Match>
  <Available>2000</Available>

<InputParam>stype=svideo&amp;pg=q&amp;q=car&amp;vavi=1&amp;vmpeg=1&amp;vmov=1&a
mp;vstrm=1&amp;vothr=1&amp;mmW=1&amp;mvcat=0&amp;dp=xml&amp;who=issid</InputPar
am>
</MmHeader>

<MmEntry>
  <Title>CNN 3/28/00 7:29:52 AM 0:26</Title>
  <Abstract>In car-crazy california, gas prices are said to be the most
expensive in the nation. Look at these self-serve prices. There are rumblings
in the california legislature to repeal or temporarily</Abstract>
  <Size>0</Size>

<URL>http://fastv.altavista.com/default.cfm?playclip=cnn_2000_0328_07295226</UR
L>
  <FrameCount>0</FrameCount>
  <Similar>n</Similar>
  <ClipTitle>cnn</ClipTitle>
  <ClipAuthor>FastV</ClipAuthor>
  <ClipCopyright>(c)1998, 1999 FasTV Inc.</ClipCopyright>
  <ClipFormat>ram</ClipFormat>
  <ClipDuration>28</ClipDuration>
  <ClipChannels>1</ClipChannels>
  <ClipSampleBits>16</ClipSampleBits>
  <ClipSampleRate>16000</ClipSampleRate>
  <MmType>12</MmType>
  <MmClass>2051</MmClass>
  <MmSize>7147</MmSize>
  <MmHeight>180</MmHeight>
  <MmWidth>240</MmWidth>
  <MmFileName>07295226.rm</MmFileName>
  <MmID>70419431</MmID>
  <TNailURL>http://thumb-2.image.altavista.com/image/</TNailURL>
  <TNailType>1</TNailType>
  <TNailSize>2794</TNailSize>
  <TNailHeight>115</TNailHeight>
  <TNailWidth>115</TNailWidth>
  <VideoFrameRate>0</VideoFrameRate>
  <VideoDepth>12</VideoDepth>
</MmEntry>

. . .

</MmSearch>

</Val>
```

6 Complete DTD File

The URL of the DTD is returned in the second line of the XML header for use by XML parsers. For example:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE Val SYSTEM "http://iss.svc.altavista.com/iss.dtd">
```

The complete DTD file is displayed below.

```
<!-- DTD file content. Define to support web and multimedia search. -->

<!-- Document Type Definition for VAL. Web and Multimedia for now -->
<!-- A VAL document contains a header following by many search entries. -->
<!ELEMENT Val (WebSearch|MmSearch | NewsSearch | VALError)+>
<!ELEMENT WebSearch (WebHeader, WebEntry*)>
<!ELEMENT MmSearch (MmHeader, MmEntry*)>
<!ELEMENT NewsSearch (NewsHeader, NewsEntry*)>

<!-- Web Search Elements -->
<!ELEMENT WebHeader (Match?, NumberOfKW?, SpellCheck?, SiteName?,
    KWOccurrences*, Range?, RelatedSearch*, AskJeeves*, RealName*)>

<!ELEMENT KWOccurrences EMPTY>
<!ELEMENT RelatedSearch EMPTY>
<!ELEMENT AskJeeves EMPTY>
<!ELEMENT RealName EMPTY>

<!ELEMENT WebEntry (Title?, URL?, Abstract?, Size?, Ranking?, Date?, Language?,
    Encoding?, SiteID?, FactsheetID?)>

<!ATTLIST KWOccurrences Keywords CDATA #REQUIRED>
<!ATTLIST KWOccurrences Occurrences CDATA #REQUIRED>
<!ATTLIST KWOccurrences KWUsed (y|n) "y">

<!ATTLIST RelatedSearch RelatedTopic CDATA #REQUIRED>
<!ATTLIST RelatedSearch RelatedOptions CDATA #REQUIRED>

<!ATTLIST AskJeeves AJQuestion CDATA #REQUIRED>
<!ATTLIST AskJeeves AJAnswer CDATA #REQUIRED>

<!ATTLIST RealName RNKey CDATA #REQUIRED>
<!ATTLIST RealName RNSite CDATA #REQUIRED>
<!ATTLIST RealName RNPerfectMatch (y|n) "y">

<!ELEMENT Match (#PCDATA)>
<!ELEMENT NumberOfKW (#PCDATA)>
<!ELEMENT SpellCheck (#PCDATA)>
<!ELEMENT SiteName (#PCDATA)>
<!ELEMENT Range (#PCDATA)>

<!ELEMENT Title (#PCDATA)>
<!ELEMENT URL (#PCDATA)>
<!ELEMENT Abstract (#PCDATA)>
<!ELEMENT Size (#PCDATA)>
<!ELEMENT Ranking (#PCDATA)>
<!ELEMENT Date (#PCDATA)>
<!ELEMENT Language (#PCDATA)>
<!ELEMENT Encoding (#PCDATA)>
<!ELEMENT SiteID (#PCDATA)>
<!ELEMENT FactsheetID (#PCDATA)>
```

```
<!-- Multimedia Search Elements -->

<!ELEMENT MmHeader (Match?, Available?, InputParam?, RelatedSearch*)>
<!ELEMENT MmEntry (Title?, Abstract?, Size?, URL?, FrameCount?, Similar?, MmPage?,
  ClipTitle?, ClipAuthor?, ClipCopyright?,
  ClipAbstract?, ClipFormat?, ClipDuration?, ClipChannels?,
  ClipSampleBits?, ClipSampleRate?,
  MmType?, MmClass?, MmSize?, MmHeight?,
  MmWidth?, MmURL?, MmFileName?, MmID?,
  TNailURL?, TNailType?, TNailSize?, TNailHeight?, TNailWidth?,
  VideoFrameRate?, VideoDepth?
)>

<!ELEMENT Available (#PCDATA)>
<!ELEMENT InputParam (#PCDATA)>
<!ELEMENT TNailURL (#PCDATA)>
<!ELEMENT MmURL (#PCDATA)>
<!ELEMENT MmFileName (#PCDATA)>
<!ELEMENT ClipTitle (#PCDATA)>
<!ELEMENT ClipAuthor (#PCDATA)>
<!ELEMENT ClipCopyright (#PCDATA)>
<!ELEMENT ClipAbstract (#PCDATA)>
<!ELEMENT ClipFormat (#PCDATA)>
<!ELEMENT ClipDuration (#PCDATA)>
<!ELEMENT ClipChannels (#PCDATA)>
<!ELEMENT ClipSampleBits (#PCDATA)>
<!ELEMENT ClipSampleRate (#PCDATA)>
<!ELEMENT MmType (#PCDATA)>
<!ELEMENT MmClass (#PCDATA)>
<!ELEMENT MmSize (#PCDATA)>
<!ELEMENT MmHeight (#PCDATA)>
<!ELEMENT MmWidth (#PCDATA)>
<!ELEMENT TNailType (#PCDATA)>
<!ELEMENT TNailSize (#PCDATA)>
<!ELEMENT TNailHeight (#PCDATA)>
<!ELEMENT TNailWidth (#PCDATA)>
<!ELEMENT VideoFrameRate (#PCDATA)>
<!ELEMENT VideoDepth (#PCDATA)>
<!ELEMENT FrameCount (#PCDATA)>
<!ELEMENT Similar (#PCDATA)>
<!ELEMENT MmPage (#PCDATA)>
<!ELEMENT MmID (#PCDATA)>

<!-- News Search Element -->
<!ELEMENT NewsHeader (Match?, Available?, InputParam?)>
<!ELEMENT NewsEntry (Title?, Abstract?, Source?, Size?, Ranking?,
  URL?, Date?, Time?, NewsCategory?)>
<!ELEMENT Source (#PCDATA)>
<!ELEMENT Time (#PCDATA)>
<!ELEMENT NewsCategory (#PCDATA)>

<!-- VAL Error -->
<!ELEMENT VALError (Match, NumberOfKW, ErrorNumber?, ErrorString?)>
<!ELEMENT ErrorNumber (#PCDATA)>
<!ELEMENT ErrorString (#PCDATA)>
```

7 Example ISS Query Via TELNET to Port 80

XML or VAL queries can be made via a telnet connection to the http port, 80. The following example makes a simple query, **pg=q**, requests all languages be returned, **kl=XX**, queries for pages that contain capybara, **q=capybara**, asks for VAL (ASCII text) output, **dp=val**, and identifies the customer making the query, **who=issid**.

Note: The **who** parameter, *issid* in this example, is a unique id given by AltaVista to each ISS customer.

```
> telnet partners.altavista.com 80
Trying 204.152.190.65...
Connected to altavista.com.
Escape character is '^]'.
GET /cgi-bin/query?pg=q&kl=XX&q=capybara&dp=val&who=issid HTTP/1.0

HTTP/1.0 200 OK
Date: Fri, 12 Feb 1999 19:19:36 GMT
Server: AV/1.0.1
MIME-Version: 1.0
Content-Length: 2621
Content-Type: text/plain

M=1869
C=1
K=capybara
c=3141
u=y
I=1-10

T=Capybara-33 Specifications
u=http://www.symbolicsound.com/kyma-capyinfo.html
a=Capybara-66. The Capybara-66 is the hardware component of the Kyma System.
Processing Capacity. 2 to 8 66 MHz Motorola 56002 digital signal processors. 67.
s=2228
r=1126
d=31-Oct-97
l=en

T=The Tapir Gallery - A capybara in South Africa
u=http://www.tapirback.com/tapirgal/capy002.htm
a=A capybara in South Africa. Photo by Marlene Spies, October 1996. Isn't the Web
wonderful? This JPEG photo was originally sent from South Africa by...
s=2021
r=1126
d=07-Apr-98
l=en

T=Capybara Club
u=http://www.bekkoame.or.jp/~ttakasu/capybara.html
a=Sorry, Capybara Club has been moved to
http://www.bekkoame.or.jp/~ttakasu/capybara/capybara.html.
s=248
r=1125
d=03-Nov-97
l=xx

T=Welcome to capybara.org!
u=http://www.capybara.org/
a=4:13PM up 17 days, 6 mins, 4 users, load averages: 0.25, 0.16, 0.11. All
Beorn.Capybara.Org users should read and follow the system's Acceptable Usage...
s=1350
```

r=1125
d=18-Sep-98
l=en

T=Capybara, Netherlands
u=http://utx500.civ.utwente.nl:8888/Rcn%3dCapybara,%20c%3dNL
a=Move upwards to. The World. Netherlands. Capybara. cn. capybara. l. Nijmegen. o. Katholieke Universiteit Nijmegen. description. Voedsel: waterplanten,...
s=1301
r=1124
d=03-Nov-95
l=nl

T=Capybara, Netherlands
u=http://iliad.lib.mq.edu.au:8888/Rcn%3dCapybara,%20c%3dNL
a=Capybara. Name. capybara. Locality. Nijmegen. Organization. Katholieke Universiteit Nijmegen. Description. Voedsel: waterplanten, gras. Lengte tot 120...
s=1391
r=1124
d=03-Nov-95
l=nl

T=Networked capybara!
u=http://capybara.sk-pttsc.lj.edus.si/
a=Say hello to our little friend! Welcome to Capybara, a Linux-addicted 486 on a 128K line. It bears the proud name of the world's largest rodent, Capybara,.
s=2657
r=1112
d=08-Jan-99
l=en

T=J-WAVE Website: KENWOOD CAPYBARA RESTAURANT: 019
u=http://www.j-wave.co.jp/capybara/index.htm
a=ISSUE 019 SEP.,1998.
8...9...4.....
.....
s=15691
r=1112
d=08-Jan-99
l=ja
L=sjis

T=Chester Zoo's Bush Dogs, Camels and Capybara
u=http://www.chesterl.demon.co.uk/
a=Bush Dogs. These stocky, terrier-like wild dogs from South America often live in packs of up to 15 animals, hunting and eating anything that they can...
s=2555
r=1111
d=29-Mar-97
l=en

T=Capybara, Netherlands
u=http://wp1.rutgers.edu:8888/Lcn%3dCapybara,%20c%3dNL
a=Capybara. objectClass. top applicationEntity dSA pilotObject quipuDSA quipuObject. Name. capybara. Locality. Nijmegen. Organization. Katholieke...
s=2581
r=1109
d=03-Nov-95
l=nl
Connection closed by foreign host.>

8 Input Parameter Validity

This quick reference table provides the input parameters valid for each search type.

Parameter	Description	Used in Search Type:				
		Text Search		Multimedia Search		
		Simple	Advanced	Audio	Image	Video
amp3	Include MP3 Files Flag	No	No	Yes	No	No
aotr	Include Other Audio Files Flag	No	No	Yes	No	No
astrm	Include Streaming Audio Files Flag	No	No	Yes	No	No
audset	Audio Search Selection	No	No	Yes	No	No
awav	Include WAV Files Flag	No	No	Yes	No	No
d0	Start Date	Yes	Yes	No	No	No
d1	End Date	Yes	Yes	No	No	No
dp	Output Format: VAL or XML	Yes	Yes	Yes	Yes	Yes
enc	Encoding Option	Yes	Yes	No	No	No
FFF	Family Friendly Filter Option	Yes	Yes	Yes	Yes	Yes
ibb	Include Black And White Images Flag	No	No	No	Yes	No
ibw	Include Color Images Flag	No	No	No	Yes	No
iclr	Include Graphics Flag	No	No	No	Yes	No
igrph	Include Buttons And Banners Flag	No	No	No	Yes	No
imgset	Image Search Selection	No	No	No	Yes	No
ipht	Include Photos Flag	No	No	No	Yes	No
kl	Single Language Selection	Yes	Yes	No	No	No
KL	Multiple Language Selection	Yes	Yes	No	No	No
macat	Multimedia Audio Partner Selection	No	No	Yes	No	No
micat	Multimedia Image Partner Selection	No	No	Yes	Yes	No
mmdo	Multimedia Search Subtype	No	No	Yes	Yes	Yes
mmF	Multimedia Source FTP Flag	No	No	Yes	No	No
mmW	Multimedia Source Web Flag	No	No	Yes	Yes	Yes
mvcat	Multimedia Video Partner Selection	No	No	Yes	No	Yes
nbq	Number Of Entries In Results	Yes	Yes	Yes	Yes	Yes
objid	Multimedia About/Similar Object Id	No	No	Yes	Yes	Yes
objtype	Multimedia Object Type	No	No	Yes	Yes	Yes
pagefp	Multimedia Same Page Id	No	No	Yes	Yes	Yes
pg	Query Type: simple or advanced	Yes	Yes	Yes	Yes	Yes
q	Query String	Yes	Yes	Yes	Yes	Yes
r	Advanced Search Ranking String	No	Yes	No	No	No
sc	Site Compression	Yes	Yes	No	No	No
si	Site Identification Code	Yes	Yes	No	No	No
stq	Ordinal Of First Results Entry	Yes	Yes	Yes	Yes	Yes
stype	Search Type: text audio, video, image	Yes	Yes	Yes	Yes	Yes
vavi	Include Avi Files Flag	No	No	No	No	Yes
vidset	Video Search Selection	No	No	No	No	Yes
vmov	Include Quicktime Files Flag	No	No	No	No	Yes
vmpeg	Include Mpeg Files Flag	No	No	No	No	Yes
vothr	Include Other Video Files Flag	No	No	No	No	Yes

Parameter	Description	Used in Search Type:				
		Text Search		Multimedia Search		
		Simple	Advanced	Audio	Image	Video
vstrm	Include Streaming Video Files Flag	No	No	No	No	Yes
who	Customer Id Code	Yes	Yes	Yes	Yes	Yes

9 Glossary

AIFF: Audio IFF. A format developed by Apple Computer Inc. for storing high-quality sampled audio and musical instrument information. SGI and several professional audio packages also use this format.

Alphanumeric: An item that consists of both letters and numbers.

Apache Modules: Software written using the Apache API to add extension to the public-domain Web server Apache, which was developed by a loosely-knit group of programmers. The first version of Apache, based on the NCSA httpd Web server, was developed in 1995. Because it was developed from existing NCSA code plus various patches, it was called a patchy server — hence the name Apache Server.

ASCII: This is an acronym for American Standard Code for Information Interchange. This seven-bit code represents most of the basic letters of the alphabet, numbers and characters. The computer interprets these characters as binary and can read, manipulate, store and retrieve files created in ASCII — also known as text files.

ASF: Advanced Streaming Format, developed by Microsoft.

AU: Unix Audio File.

AVI: Audio Video Interleave. An audio-video standard designed by Microsoft. Apparently proprietary and Microsoft Windows-specific.

Boolean logic: In this form of logic answers are either *true* or *false*. Alternatively 0 is *false* and 1 is *true*.

Browser: A program that allows a person to read hypertext. The browser gives some means of viewing the contents of pages and of navigating from one page to another.

CGI: Common Gateway Interface is a standard for running external programs from a World Wide Web HTTP server. CGI specifies how to pass arguments to the executing program as part of the HTTP request. It also defines a set of environment variables. For more information about CGI refer to NCSA's CGI web pages, <http://hoohoo.ncsa.uiuc.edu/cgi/interface.html>.

Client: This is a computer connected to a host or to a server; can also refer to the software that makes this connection possible.

Crawler: A program that automatically explores the World Wide Web by retrieving a document and recursively retrieving some or all the documents that are referenced in it. This is in contrast with a normal web browser operated by a human that doesn't automatically follow links other than inline images and URL redirection. Crawlers are also called spiders or robots.

Default: When a value is left unspecified, this is the value used.

DOM: Document Object Model. A W3C specification for application program interfaces for accessing the content of HTML and XML documents.

Domain: A group of computers whose *hostnames* share a common suffix, the *domain name*.

DTD: Document Type Definition. This is the formal specification of a markup language, written using SGML.

Encode: The process of converting data from one format into a given format.

- FTP:** File Transfer Protocol. A client-server protocol which allows a user on one computer to transfer files to and from another computer over a TCP/IP network. Also the client program the user executes to transfer files. It is defined in STD 9, RFC 959.
- GIF:** Graphics Interchange Format. A standard for digitized images compressed with the LZW algorithm, defined in 1987 by CompuServe.
- GIF89a:** See GIF.
- HTML:** Hypertext Markup Language. A Hypertext document format used on the World Wide Web. Built on top of SGML. *Tags* are embedded in the text. A *tag* consists of a “<”, a *directive*, zero or more parameters and a “>”. Matched pairs of directives, like <TITLE> and </TITLE> are used to delimit text which is to appear in a special place or style. The World Wide Web Consortium (W3C) is the standards body for HTML. Refer to <http://www.w3.org/Markup/> for more information.
- HTTP:** Hypertext Transfer Protocol. This is the method by which files are transferred between computers on the Internet. It conventionally uses port 80. ISS uses version 1.0, refer to RFC 1945.
- HTTPd:** Hypertext transfer protocol daemon is an HTTP/1.0-compatible server, written by Rob McCool of NCSA, for making hypertext and other documents available to World Wide Web browsers. HTTPd is designed to be small and fast and to work with most HTTP/0.9 and HTTP/1.0 browsers. It allows users to customize their server to execute searches and handle HTML forms. It also supports server side include files, allowing users to include the output of commands or other files in HTML documents.
- Hypertext:** A term coined by Ted Nelson around 1965 for a collection of documents containing cross-references or *links* which, with the aid of an interactive browser program, allow the reader to move easily from one document to another. HTML is the hypertext used on the Internet.
- IANA:** Internet Assigned Numbers Authority. The central registry for various “assigned numbers”: Internet protocol parameters, such as port, protocol and enterprise numbers, and options, codes and types. Refer to <http://www.iana.org/>
- Internet:** The Internet is the largest internet (with a lowercase “i”) in the world. It is a three level hierarchy composed of backbone networks (e.g. ARPANet, NSFNet, MILNET), mid-level networks, and stub networks. These include commercial (.com or .co), university (.ac or .edu) and other research networks (.org, .net) and military (.mil) networks and span many different physical networks around the world with various protocols, chiefly the Internet Protocol.
- IP:** Internet Protocol. The network layer for the TCP/IP protocol suite widely used on Ethernet networks and defined in STD 5, RFC 791. IP is a connectionless, best-effort packet switching protocol. It provides packet routing, fragmentation and re-assembly through the data link layer.
- ISO:** International Organization for Standardization. A voluntary, non-treaty organization founded in 1946, responsible for creating international standards in many areas, including computers and communications. Its members are the national standards organizations of 89 countries, including the American National Standards Institute. The ISO web site is <http://www.iso.ch/>.
- Java™:** A simple, object-oriented, distributed, interpreted, robust, secure, architecture-neutral, portable, multithreaded, dynamic, buzzword-compliant, general-purpose programming language developed by Sun Microsystems. Java supports programming for the Internet in the form of platform-independent Java applets.

- JavaServer Pages™ (JSP):** are designed to create interactive pages as part of a Web-based application. JSP enables developers to separate programming logic from page design through the use of components that are called from the page itself. Therefore, JSP provides an alternative to creating CGI scripts that makes page development and deployment easier and faster.
- JPEG:** Joint Photographic Experts Group was the original name of the committee that designed the standard image compression algorithm. JPEG is designed for compressing either full-color or grayscale digital images of real-world scenes. It does not work so well on non-realistic images, such as cartoons or line drawings. JPEG does not handle compression of black-and-white (1 bit-per-pixel) images or moving pictures. Standards for compressing those types of images are being worked on by other committees, named JBIG and MPEG.
- JPGProg:** Progressive JPEG.
- Key:** The flag returned in both VAL and XML output that precedes and defines an output data value.
- Keyword:** A command or a parameter that restricts or enhances a search. Refer to Sections 3.1.6 and 3.2.6.
- Lossless:** A term describing a data compression algorithm that retains all the information in the data, allowing it to be recovered perfectly by decompression.
- Liquid Audio:** Audio file format developed by Liquid Audio, Inc.
- MIDI:** Musical Instrument Digital Interface. Music, hardware, protocol, and file formats. A hardware specification and protocol used to communicate note and effect information between synthesizers, computers, music keyboards, controllers, and other electronic music devices. It is basically a high-speed serial connection with separate connections for MIDI in, MIDI out and MIDI through (to allow devices to be chained). The basic unit of information is a “note on/off” event which includes a note number (pitch) and key velocity (loudness). There are many other message types for events such as pitch bend, patch changes and synthesizer-specific events for loading new patches etc. There is a file format for expressing MIDI data that is like a dump of data sent over a MIDI port.
- MIME:** Multipurpose Internet Mail Extensions. A standard for multi-part, multimedia electronic mail messages and World Wide Web hypertext documents on the Internet. MIME provides the ability to transfer non-textual data, such as graphics, audio and fax. It is defined in RFC 2045, RFC 2046, RFC 2047, RFC 2048, RFC 2049, and BCP0013. It uses mimencode to encode binary data into base 64 using a subset of ASCII.
- MOV:** See QuickTime™
- MP3:** MPEG-1 audio layer 3. MP3 is a digital audio compression algorithm that achieves a compression factor of about twelve while preserving sound quality. It does this by optimizing the compression according to the range of sound that people can actually hear. MP3 is currently (July 1999) the most powerful algorithm in a series of audio encoding standards developed under the sponsorship of the Moving Picture Experts Group (MPEG) and formalized by the International Organization for Standardization (ISO).
- MPEG:** Moving Picture Experts Group. MPEG is an ISO committee that generates standards for digital video compression and audio. It is also the name of their algorithms. MPEG-1 is optimized for CD-ROM and is the basis for MP3. MPEG-2 is aimed at broadcast quality video for applications such as digital television set-top boxes and DVD. MPEG-4 is a standard for low bandwidth video telephony and multimedia on the World Wide Web. MPEG-3 was merged into MPEG-2.

- Microsoft® Active Servers Pages (ASP):** are a language-independent framework designed by Microsoft for efficient coding of server side scripts that are designed to be executed by a Web server in response to a user's request for a URL. ASP scripts are similar to other server-side scripting used on other platforms such as Perl, Python, and so on.
- Multimedia:** refers to the use of audio, video or images.
- NCSA:** National Center for Supercomputing Applications. NCSA is the birthplace of the first version of the Mosaic World Wide Web browser.
- Parameter:** Flag used in the input query string that precedes and defines an input data value.
- Parser:** An algorithm or program to determine the syntactic structure of a sentence or string of symbols in some language.
- Perl:** A high-level programming language, started by Larry Wall in 1987 and developed as an open source project. It has an eclectic heritage, deriving from the ubiquitous C programming language and to a lesser extent from sed, awk, various Unix shell languages, Lisp, and at least a dozen other tools and languages. Originally developed for Unix, it is now available for many platforms.
- PNG:** Portable Network Graphics. An extensible file format for the lossless, portable, well-compressed storage of raster images. PNG provides a patent-free replacement for GIF and can also replace many common uses of TIFF. Indexed-color, grayscale and true color images are supported, plus an optional alpha channel. Sample depths range from 1 to 16 bits.
- Protocol:** specifies how computers will talk to each other on a network.
- Query String:** When a CGI script is called with the GET method, the server places into the *QUERY_STRING* variable the information that the user provided in the FORM.
- Quicktime™:** Apple Computer's standard for integrating full-motion video and digitized sound into application programs.
- RealAudio:** A program for playing audio over the Internet. The system is implemented as a client/server architecture. The RealAudio server incorporates an encoder that compresses sound into RealAudio files. The client side is a web browser plug-in or add-on (a recent version of Internet Explorer apparently has built-in support for RealAudio) which allows the stream of data sent from the server to be uncompressed and output using the normal sound facilities of the computer, such as a sound card. It is a product of RealNetworks, Inc.
- RealVideo:** Similar to RealAudio only for the playing of video over the Internet.
- Script:** A script is a program (or a group of commands) that runs on a web server and processes requests based on input from the browser.
- Search Engine:** A general class of programs that search documents for specific keywords and return results.
- Server:** The computer that handles requests from other computers.
- SGML:** Standard Generalized Markup Language. A generic markup language for representing documents. SGML is an International Standard that describes the relationship between a document's content and its structure. SGML allows document-based information to be shared and re-used across applications and computer platforms in an open, vendor-neutral format. SGML is defined in ISO 8879:1986 Information processing -- Text and office systems - Standard Generalized Markup Language (SGML), an ISO standard produced by JTC 1/SC 18 and amended by Amendment 1:1988.

- Source:** refers to the location from which the results are obtained, i.e.: web, FTP, AltaVista Partners sites.
- Streaming:** Playing audio or video in real time as it is downloaded over the Internet as opposed to storing it in a local file first. Streaming audio or video avoids the delay entailed in downloading an entire file and then playing it later. Streaming requires a fast connection and a computer powerful enough to execute the decompression algorithm in real time.
- Syntax:** The construction and arrangement of words in a language or API.
- TCP:** Transmission Control Protocol. The most common transport layer protocol used on Ethernet and the Internet. It was developed by DARPA. TCP is built on top of Internet Protocol (IP) and is nearly always seen in the combination TCP/IP (TCP over IP). It adds reliable communication, flow-control, multiplexing and connection-oriented communication. It provides full-duplex, process-to-process connections. Defined in STD 7, RFC 793. It is connection-oriented and stream-oriented.
- URL:** Uniform Resource Locator. A standard way of specifying the location of an object, typically a web page, on the Internet. Other types of object are described below. URLs are the form of address used on the World Wide Web. They are used in HTML documents to specify the target of a hyperlink that is often another HTML document (possibly stored on another computer).
- Value:** This is an assigned or returned quality, it follows *parameters* for input and *keys* for output.
- W3C:** World Wide Web Consortium. The main standards body for the World Wide Web. W3C works with the global community to establish international standards for client and server protocols that enable on-line commerce and communications on the Internet. It also produces reference software. The Massachusetts Institute of Technology (MIT) created W3C on 25 October 1994. Netscape Communications Corporation was a founding member. MIT LCS and INRIA run the Consortium, in collaboration with CERN where the web originated. Industrial members fund W3C but its products are freely available to all.
- WAV:** A sound format developed by Microsoft and used extensively in Microsoft Windows.
- XML:** XML is the “Extensible Markup Language” (extensible because it is not a fixed format like HTML). It is designed to enable the use of SGML on the World Wide Web. XML is not a single, predefined markup language: it’s a metalanguage — a language for describing other languages — which lets users design their own markup. (A predefined markup language like HTML defines a way to describe information in one specific class of documents: XML lets you define your own customized markup languages for different classes of document.) XML can do this because it’s written in SGML, the international standard metalanguage for markup. Complete information on XML is available at World Wide Web Consortium (W3C) web site: <http://www.w3.org/TR/1998/REC-xml-19980210>.

10 AltaVista Search String Syntax

This section provides an overview of the search string syntax used by the AltaVista Search engine when processing query (*q=string*) and ranking (*r=string*) strings. Refer to Sections 3.1.6 and 3.2.6 for the definition of the *q* parameter and to Section 3.1.7 for the definition of the *r* parameter.

The table below and the following description explain how the strings are viewed.

Query String	Search Description
mona+lisa	Finds documents that contain either <i>mona</i> or <i>lisa</i> or any capitalized variant (Mona, MONA, liSA, Lisa, <i>et cetera</i>). AltaVista ranks the results to show first the documents containing both words, close together, and near the top of the document.
Mona+Lisa	Finds documents that have either <i>Mona</i> or <i>Lisa</i> but not any other capitalized variation. When you use a capitalized word, AltaVista assumes that you are only interested in an exact match.
%2Bmona+%2Blisa	Finds only documents that contain both <i>mona</i> and <i>lisa</i> . Any capitalized variant (Mona, MONA, liSA, Lisa, <i>et cetera</i>) will be found.
%22Mona+Lisa%22	Finds documents that have the two words capitalized and found right next to each other. Placing double quote (%22) marks around any series of words turns them into a phrase and tells AltaVista that you are only interested in documents that have them in this specific order.
%2B%22Mona+Lisa%22+-Louvre	Finds documents that contain the phrase <i>Mona Lisa</i> but do not contain the word <i>Louvre</i>
%2BMona+%2BLis*	Finds documents that contain <i>Mona</i> and any word starting with <i>Lis</i> . Use this feature if you are not sure how a word is spelled.

Note: In the table above %2B is the URL encoded representation for plus sign (+) and is not to be confused with the CGI word concatenation operator (+). %22 is the URL encoded representation for double quote mark (")

Words and Phrases

A phrase is a group of words or numbers linked together. Phrases are used when you want specific words or numbers to appear together in the results. If you want to find an exact phrase, use double quote marks (%22) around the phrase in your query. You can also create phrases using punctuation or special characters such as dashes, underscore lines, commas, slashes, or dots. For example, by searching for *1-800-555-1212* instead of *1 800 555 1212*, the dashes link the numbers together as a phrase.

Case Sensitivity

AltaVista offers case sensitivity that fits the normal usage paradigm of typical users. When you use lowercase text, the search service finds both uppercase and lowercase results. When you use uppercase text, the search service only finds uppercase results. For example, if you search for *sweden*, you'll find *Sweden*, *sweden*, and *SWEDEN* in your results pages. However, when you search for *Sweden*, you'll only see *Sweden* in the results.

Wildcard

AltaVista offers a powerful wildcard feature by using an asterisk (*) at the end of or within a word with at least three letters of the search term. For example, by typing an asterisk (*) at the end of a keyword, you can search for multiple forms of the word, e.g. use *big**, to find *big*, *bigger*, *biggest*, and *bigwig*. The

wildcard (*) is also very convenient if you are not sure of the spelling of a word, e.g. *alumin*m* will find the (US) word *aluminum* and the (UK) word *aluminium*.

Punctuation and Special Characters

AltaVista offers special characters and punctuation. AltaVista searches define a word as any combination of letters and numbers separated by any of the following: white space (spaces, tabs, line ends, or the start or end of a document), special characters, and punctuation, such as % (%25), \$ (%24), / (%2F), # (%23), and _).

AltaVista interprets punctuation as a separator for words. Placing punctuation or a special character between each word (with no spaces between the characters and the words) is another way to indicate a phrase. Example: Entering *Jean-Luc Picard* is the same as *%22Jean Luc%22 Picard*. Hyphenated words, such as x-files, are also considered phrases because of the hyphen.

11 URL Encoded Representation

The following table can be used to convert non-alphanumeric characters to URL encoded representation for replacement in query (**q=**) and ranking (**r=**) strings. Note that period (.), asterisk (*), minus (-) and underline () do not have to be converted.

ASCII Character	URL Encoded Representation
^@ NUL (Null)	%00
^A SOH (Start of Heading)	%01
^B STX (Start Text)	%02
^C ETX (End Text)	%03
^D EOT (End of Transmission)	%04
^E ENQ (Enquiry)	%05
^F ACK (Acknowledge)	%06
^G BEL (Bell)	%07
^H BS (Backspace)	%08
^I TAB HT (Horizontal Tab)	%09
^J LF (Linefeed, Newline)	%0A
^K VT (Vertical Tab)	%0B
^L FF (Formfeed)	%0C
^M CR (Carriage Return)	%0D
^N SO (Shift Out)	%0E
^O SI (Shift In)	%0F
^P DLE (Data Link Escape)	%10
^Q DC1 (X-ON)	%11
^R DC2	%12
^S DC3 (X-OFF)	%13
^T DC4	%14
^U NAK (Negative Acknowledge)	%15
^V SYN (Synchronous Idle)	%16
^W ETB (End Transmission Blocks)	%17
^X CAN (Cancel)	%18
^Y EM (End of Medium)	%19
^Z SUB (Substitute)	%1A
^[ESC (Escape)	%1B
^ \ FS (File Separator)	%1C
^] GS (Group Separator)	%1D
^ ^ RS (Record Separator)	%1E
^ _ US (Unit Separator)	%1F

ASCII Character	URL Encoded Representation
!	%21
"	%22
#	%23
\$	%24
%	%25
&	%26
'	%27
(%28
)	%29
*	%2A
+	%2B
,	%2C
-	%2D
.	%2E
/	%2F
:	%3A
;	%3B
<	%3C
=	%3D
>	%3E
?	%3F
@	%40
Space	%20
[%5B
\	%5C
]	%5D
^	%5E
_	%5F
`	%60
{	%7B
	%7C
}	%7D
~	%7E
DEL, RUB	%7F

12 Contact Information

If you would like information on becoming a customer of AltaVista Internet Search Services please visit http://doc.altavista.com/business_solutions/index.html. Please specify that you want information on Internet Search Services and provide your contact information.

This latest version of this document and *Powered by AltaVista* logos may be downloaded from http://doc.altavista.com/business_solutions/search_link/iss_download.html.

To provide technical feedback on this document, please send that information to iss-support@av.com.