



# Software Product Description

---

**PRODUCT NAME:** AltaVista Search CD-ROM Toolkit 98

**SPD:** 70.24.00

## DESCRIPTION

The AltaVista Search CD-ROM Toolkit brings the power of the popular AltaVista Search technology to CD-ROM distributions. Publishers of information on CD-ROM use the AltaVista Search CD-ROM Toolkit to build a complete, full-text index of every file distributed on the CD-ROM. End users are able to quickly and efficiently search the contents of the CD-ROM using the same powerful, high speed AltaVista Search engine technology they use when searching the WorldWide Web.

The AltaVista Search CD-ROM Toolkit offers the following benefits:

- Fast, complete information retrieval using proven AltaVista Search technology. Every word and every number is indexed.
- Support for over 200 different file types allows you to put just about any type of document on your CD-ROM and make it searchable.
- Familiar AltaVista Search syntax and semantics makes it easy to submit queries to the search engine.
- Hit highlighting in HTML and text documents makes it easy to find the query terms for which you're searching.
- Wizard-based index creation process, making it extremely simple to build a searchable CD-ROM.
- Customizable, browser-based user interface for submitting queries and presenting results – add your own logo and tailor the UI to fit your corporate standards or enforce your corporate branding. The search function can be integrated into your existing CD-ROM user interface to make accessing the search interface an easy and seamless operation.

- Small disk and memory “footprint” during searching, resulting in minimal impact on the end user's PC.
- Efficient indexing, resulting in an index that can be as small as 5% of the size of the source data, but usually no more than 30%.
- A broad range of search capabilities give end users a variety of options in how to search for information. These include phrase searching, numeric searching, case-sensitive and accent-sensitive searching and field/keyword searching (by title, by location, by file extension).

The AltaVista Search CD-ROM Toolkit provides all the tools necessary to build a self-contained full-text search and retrieval capability for any CD-ROM. Publishers are guided through the process by a wizard, which provides easy-to-follow step-by-step instructions on building an index. Much of the process is automated and requires no intervention on the part of the publisher.

Using the AltaVista Search CD-ROM Toolkit, the basic process for building a searchable CD-ROM is quite simple:

1. Collect all of the documents to be published on the CD-ROM into a directory structure on the hard drive.
2. Configure the indexer, specifying what directory structure to index, what subdirectories or specific files (if any) that should not be indexed, and what file types to index.
3. Run the indexer and build the index. The index and the search engine components needed by the end user to search the index are automatically placed in the directory structure ready to be published.

4. Customize the query pages that the end user will use to submit queries and review results of searches, and integrate search into your CD-ROM layout.
5. Copy the directory structure (which now includes the source documents, the index and the search engine components) to the master CD-ROM.
6. Reproduce the required number of CD-ROM copies, up to the quantity limit allowed by your AltaVista Search CD-ROM Deployment License.

## Features

- Same, powerful search technology used at the popular AltaVista Search site on the WorldWide Web.
- Ranked search results – results are ranked based on a series of algorithms related to: (1) the rarity (frequency of occurrence) of the query words or phrases in the total index; (2) whether the query words or phrases are found in the title or first few lines of a document; (3) whether all of the specified words or phrases appear in a document; and (4) whether multiple query words or phrases are found close to each other in a document.
- Hit highlighting – Query terms found in HTML and text documents are highlighted during document display, making it easy to find the words or phrases you are interested in.
- Support for over 200 different file types, including Microsoft Office 97 applications, Adobe PDF, PostScript, Microsoft Access, Lotus 1-2-3, HTML, text and many other popular file types.
- Customizable user interface – tailor the query page and add your own logo to the results page to match your corporate standards.
- Keyword searching – search files based on keywords such as *title*, *location*, and *file extension*.
- Phrase searching – because AltaVista Search indexes every word and every number, searches for phrases such as “to be or not to be” are possible.
- Wildcard (pattern matching) searching – use the asterisk to search for word derivatives and spelling variants of query terms.
- Case-sensitive searching – specifying uppercase characters in the query string will narrow the search to exact case matches.
- Multinational searching – specifying accented characters in the query string will narrow the

search to words with those special multinational characters.

- Require/exclude specific words/numbers/phrases in searches to quickly narrow results.

## Hardware Requirements

*For creating an index:*

- Intel 486 25MHz processor
- 12MB of memory
- 15MB disk space required for installation
- Additional disk space required for building an index – index size will range from an average of 5% to 30% of the source data

*For searching an indexed CD-ROM:*

- Intel 486 25MHz processor
- 12MB of memory

## Software Requirements

*For creating an index and searching an indexed CD-ROM:*

- Microsoft Windows 95
- or
- Microsoft Windows NT V4.0

and

- Microsoft Internet Explorer V2 or later
- or
- Netscape Navigator V2.x or later

The Windows 95 TCP/IP protocol (including in Windows 95 and Windows NT) must be installed and configured for the browser-based search interface to function. Third party TCP/IP implementations are not supported.

## *Optional Software*

For viewing documents listed in the results of a search, either the associated application is required or one of the following viewers:

- Quick View (included as optional feature of Windows 95)
- Quick View Plus (available separately)

## Software Licensing

Two types of licenses are required for use of this product:

### *Development License*

The AltaVista Search CD-ROM Toolkit Development License gives you the right to use the Toolkit to create

an index of your source files and to test the operation of the search interface against that master index. It does NOT give you the right to reproduce or distribute that index on CD-ROM.

### *Deployment License*

The AltaVista Search CD-ROM Toolkit Redistributable Components Deployment License gives you the right to reproduce and distribute on multiple CD-ROMs the "Redistributable Components" (the search engine and the index created with the Toolkit). The number of copies you may reproduce is limited to the quantity specified in the Deployment License purchased. There are six license tiers:

- 1,000 CD-ROM Impressions
- 5,000 CD-ROM Impressions
- 20,000 CD-ROM Impressions
- 50,000 CD-ROM Impressions
- 100,000 CD-ROM Impressions
- 250,000 CD-ROM Impressions

Each CD-ROM that you produce consumes one unit from the license quantity. When the number of CD-ROMs reaches the license quantity, the license expires. Additional CD-ROMs can only be produced and distributed by purchasing another Deployment License of the appropriate quantity.

The avsIndexer.exe file (which is the actual indexer code used to create an index) and the document converter files may NOT be reproduced or distributed.

For more information about Digital's licensing terms and policies, contact your local Digital office.

### **Growth Considerations**

The minimum hardware and software requirements for any future version of this product may be different from the requirements for the current version.

### **Distribution Media**

CD-ROM

### **Ordering Information**

#### *Development Package*

Development License and Kit: QB-62ZAA-SA

Software Product Services

- Package Service: QT-62ZAA-XA

A Deployment License is required before the index created with the Development Package can be reproduced and distributed.

Software Product Services Package Service includes the right to receive any product license and kit updates released during the contract period.

### *Deployment Licenses*

Deployment Licenses give you the right to reproduce and distribute the Redistributable Components of the Toolkit in the quantities specified by the license.

<u># of CD-ROM Impressions</u>	<u>Order Number</u>
1,000	QM-62YAA-AA
5,000	QM-62YAA-AB
20,000	QM-62YAA-AC
50,000	QM-62YAA-AD
100,000	QM-62YAA-AE
250,000	QM-62YAA-AF

### **Software Product Services**

The service options listed under Ordering Information are available for this product. For more information on Software Product Services, contact your local Digital office.

### **Software Warranty**

A 90-day warranty for this Software Application Product is provided by Digital with the purchase of a license for the product.

The above information was valid at the time of release. Please contact your local Digital office for the most up-to-date information.

---

© 1998 Digital Equipment Corporation  
All rights reserved.

AltaVista, Digital and the Digital logo are trademarks of Digital Equipment Corporation. Quick View Plus is a trademark of Inso Corporation. Microsoft, Windows 95, Windows NT, Microsoft Access, Excel and PowerPoint are registered trademarks of Microsoft Corporation. Lotus 1-2-3 is a trademark of Lotus Development Corporation. Adobe, Acrobat and PostScript are trademarks of Adobe Systems Incorporated. Netscape and Netscape Navigator are trademarks of Netscape Communications Corporation.