Tech Info Library

AppleSearch 1.5 Quick Reference (11/94)

```
Revised:        11/7/94
Security:       Everyone
```

AppleSearch 1.5 Quick Reference (11/94)

==

Article Created: 7 November 1994

TOPIC ---

This article is an AppleSearch 1.5 quick reference.

DISCUSSION --

Overview

• Full-text search and retrieval
• Search results are ranked based on relevance to search criteria
• WAIS gateway feature enables searches on the Internet
• Scheduled searches automatically deliver new or changed information

Server

Hardware
  AppleSearch server runs on a Power Macintosh, or any Macintosh computer
  with a 68040 processor, including Apple Workgroup Server, Quadra,
  Centris, and some Performa and PowerBook models

System Software
• Requires System 7.0 or later
• Compatible with System 7 Pro and System 7.5
• Requires Macintosh File Sharing or one of these versions of AppleShare:
  • AppleShare Pro (on an Apple Workgroup Server 95)
  • AppleShare 4.0.2
  • AppleShare 3.0.2 or 3.0.3
• Requires MacTCP 2.0.6 (included with Installer)

Network Connections
• Requires AppleTalk connection to client
• Requires IP connection to the Internet for WAIS gateway to operate. This
  connection can be native IP or "IPTalk" (IP encapsulated in AppleTalk)
  through a gateway.

- AppleTalk connection and IP connection may or may not be the same physical network

Disk Space
- Occupied by index files: 1-2 times the text size of all information sources combined
- Used temporarily by the indexing process: an amount equal to the text size of the largest information source
- Index and temporary files may be put on any locally-attached drives

Memory
- Memory partition size of 3500K required for server software; 4000K recommended. 4000K supports up to 10 local information sources. Each information source beyond 10 requires an additional 100K.
- An additional 1000K is used transparently by the WAIS gateway.

Operations
- Program linking must be started on the server.
- Program linking must be enabled for each AppleSearch user.
- Local information sources must be shared via AppleShare or File Sharing. Access is controlled by Users & Groups.
- AppleSearch users can be registered users or guests.
- A server can support up to 50 simultaneously connected users.

Text Translation
The AppleSearch server uses XTND translators to extract text from the files in its information sources. The following DataViz XTND translators are included with the AppleSearch product:

- MacWrite versions 4.5, 5.0, and II
- WordPerfect for Macintosh versions 1.0, 2.0, and 2.1
- Microsoft Word for Macintosh versions 3.0, 4.0, 5.0, and 5.1
- Microsoft Works (word processing) versions 2.0 and 3.0
- FrameMaker MIF versions 2.0 and 3.0
- PageMaker version 4.0
- Ragtime version 3.1
- AppleWorks
- WriteNow versions 2.0 and 3.0
- Nisus version 3.0
- PICT (text extraction only)
- Excel (tab text only, no formula)
- Microsoft Word for Windows versions 1.0 and 2.0
- WordPerfect for Windows version 5.1

Stop Words
Stop words are conjunctions, pronouns, and other words that are not meaningful in a search. These are filtered out during indexing and are not used in any search. The list of stop words is contained in the editable text file "AppleSearch Stopwords," which is in the AppleSearch folder inside the Preferences folder. Examples of stop words are "and," "by," "in," "there," and "with."

Client

```
------
```

Hardware

AppleSearch client requires Macintosh Plus or better, with 4 MB of memory and a hard disk

System Software

Client requires System 7.0 or later

Network Connections

- Requires AppleTalk connection to server
- Client may be connected to one server at a time, and to any number of information sources on that server
- User may choose to be automatically connected to a server upon system startup
- User remains connected to the server upon quitting the client application if any scheduled searches are pending

Searching

Search requests, called "reporters," can be written in natural language—the user's own words, phrases, sentences or key words—or they may use search operators. When search requests use natural language, AppleSearch translates them into Boolean operations by inserting an implied "or" between all the words (except stop words).

Search requests are not case-sensitive and should not use punctuation, since some punctuation marks are taken as adjacency operators.

Saved reporters reside on the server and may be exported to other users.

Search results are ranked according to their relevance to the user's search criteria, using these parameters: frequency of query words, matches on rare words, proximity of query words, and co-occurrence of multiple query words.

Relevance ranking makes natural language searches effective by sorting out what may be a large number of documents returned by the search. The user can concentrate on the highest-ranked documents instead of needing to narrow down the search by respecifying the search criteria.

Using "stemming," AppleSearch finds the root word of a query word, and all words that can be formed from the root word.

Search expressions are evaluated left to right, and follow this order of precedence:

1) Wild characters, wild strings, and word expansion

  ? = single wild character, for example, s?ste? finds system, sister, and so on.
  * = wild string, for example, micro* finds microbe, microbiology, micron,and so on.
  ! = word expansion, causes search to also return articles containing most common co-occurring words.

2) Parentheses

3) Proximity operators

   W/n = within n words
   Hyphen (-), period (.), apostrophe ('), comma (,) and "adj"
         are all equivalent adjacency operators, which allow a user to
         find a specific phrase.

4) Boolean operators

   or   An article will match the search criteria if either of the terms
        separated by "or" occurs in the article.
   and An article will match if the terms separated by "and" both occur
        in the article.
   not An article will match only if the term preceding "not" occurs in
        the article, but the term following "not" does not occur.

   Note: This list of operators applies to local information sources.
   For a list of operators applicable to WAIS searches, please refer to
   the AppleSearch Interface to WAIS section in this card.

Scheduled Searches
  Scheduled searches automatically deliver new or changed information to a
  user's hard disk at a predetermined time, in the form of a
  newspaper-like "update." An update contains any articles from the
  information source that match the search criteria and were added or
  changed between the last delivery of the update (or the creation of the
  reporter) and the most recent indexing.

  To receive the update, the user must be connected at the time of the
  scheduled search. If the user is not connected, the update will not be
  delivered and any new information will appear in the next update.

  Scheduled searches may be done only by registered users, not by guests.

  In the case of WAIS databases, scheduled searches can only be carried
  out on databases that support date information.

AppleSearch Interface to WAIS

AppleSearch uses the WAIS Directory Servers file to find directory servers on
the Internet. Using information from them, it builds and maintains the WAIS Data
File. WAIS Directory Servers is an editable text file which lists the following
servers by default:

  210@wais.com
  210@quake.think.com
  210@cnidr.org

WAIS Searches
  Searching WAIS databases is similar to searching local information

sources in AppleSearch. All WAIS databases support the "or" operator. In addition, some WAIS databases support the following AppleSearch operators:

- Boolean operators "and," and "not"
- Wild string (*)
- Parentheses
- Adjacency ("ADJ" in upper case only). In WAIS, this will locate words not only directly adjacent to each other, but near each other as well. The closer the words, the higher the relevance score.

Scheduled Searches
  Similar to scheduled searches of local information sources in AppleSearch, but applicable only to WAIS databases that support date information.

Custom Databases
  Databases that are not listed in the directory server can be added by specifying the name of the database, port ID, and IP address or host name in the WAIS Custom Databases file. The custom databases will be available after the next scan for new sources.

Support Information Services

Copyright 1994, Apple Computer, Inc.

Tech Info Library Article Number:16704