



**MicroModule: formats**

**REVIEW Page**

Below is the entire module on one page.



**How Can You Search For Documents In Non-HTML Formats?**

Search engines were first built to find HTML pages for keyword searching. HTML pages are often rich in text and follow common conventions that are easy to index. At first, search engine crawlers ignored relatively rare non-html files that contained little text. As search technology matured and market pressures created a more competitive environment, the large commercial search engines began to index non-HTML file formats.

Now, search engines are adding PDF, Microsoft Application, image, video, and audio files to their search indexes. Most major search engines make it easy to do simple searches for image files. Additionally, new, specialized search tools have been developed that focus on multimedia formats. Still the results of your searches will usually be HTML web pages, non-HTML files remain a small subset of information on most search engines.

**Which search engines should I use to find non-HTML files?**

Each of the following search engines has strong non-html file searching features.

Search Engine	<a href="#"><u>Google</u></a>	<a href="#"><u>AllTheWeb (FAST)</u></a>	<a href="#"><u>AV (AltaVista)</u></a>
Advanced Search	<a href="#"><u>Advanced</u></a>	<a href="#"><u>Advanced</u></a>	<a href="#"><u>Advanced</u></a>

Image Search	Yes, Thumbnail view of results. any file type, .jpg .gif .png size, color, domain	Yes, Thumbnail view of results. Specify: .jpg .gif .bmp. .pdf .swf	Yes, Thumbnail view of results.No file type filters. Search by photos, graphics, buttons/banners. Specify image sizes, image colors, sources
Video Search	No	Yes. avi, avi/DivX. RealPlayer , mpeg, Quick time Specify streams or download	Yes.avi, mpeg, Quicktime, Windows Media, Real. Specify: Duration, sources
Audio Search	No	Yes. Audio (midi, wav, au) formats can be searched using Advanced Search for web pages.	Yes, mp3, WAV, WindowsMedia, Real Audio, Other, Specify Duration, specify sources
Tabbed Search Categories	Web, Images, Groups, Directories, News	Web, news, Pictures, Video, Audio, Ftp Files.	web, images, mp3/ audio,directory, news
Boolean Options?	Yes	Yes	No
Adult Content Filter	SafeSearch	Offensive Content Filter	Family Filter, Password option  Specify Multimedia or All
View in HTML.	Yes	No	No
FAQ's / Help Files	<a href="#">Help</a>	<a href="#">Advanced Search</a>	<a href="#">Different Types of Searches</a>

### How can I save a file to my local computer?

To save a file from the Internet:

- Point at the desired link, image, or file name, and *right* click your mouse.
- Choose Save Target As or Save Image As

- Direct the file to your desktop, or some designated folder.

Depending on how you configure your Internet browser, simply clicking on a file may start the file-saving procedure. Be sure to note the file name and where you are saving it in your system to avoid the frustration of losing your new file after you've taken the trouble to download it.

### Let's try it!

1. Place your cursor over the picture below of the boy innoculating his computer.
2. Right click the mouse.
3. Choose Save Target As or Save Image As.
4. Select the location in which you would like to save the picture of this boy.
5. Click Save.

### Is it possible to get a virus from a downloaded file?



**Yes, you can 'catch a virus' from downloaded files.** Search engines can locate web pages and files for you, but they do not check for viruses. Any file you download from the Internet could carry a virus. Keep your virus scanning software current, and updated. While most viruses are attached to email or Word documents, it pays to scan all files before opening them.

Alternately, Google's HTML view feature allows you to read many file formats without downloading them, thus protecting you from virus problems.

### What are file extensions?

Different computer applications use different file types. The name of the file is followed by a period (sometimes referred to as "dot") and the file extension. Files are differentiated by a two to four letter *file extension*. These file extensions are used by computer applications to identify useable files. You can usually change a file's name, but it is best not to change the file extension.

# image.jpg

filename

period  
aka dot

file extension

## Other than HTML, which file formats are most common on the Internet?

HTML files predominate. Adobe's PDF (Portable Document Format) files are the second most common type of file found on the Web. Microsoft Word documents, and PostScript files are also common. Image, video, and audio files remain relatively rare when compared to other file formats on the Internet.

## What are some common multimedia file types?

Image, video, and audio files are available in a number of different file types. Some multimedia file types work with both Windows and Macintosh computers. Others are platform specific.

File Extension	Type of File / Type of Computer
.aiff	Sound File Mac
.au	Sound File Mac
.avi	Video / Windows
.bmp	Image file (Bit map) Windows
.gif	Image file (Graphic Interchange Format) Mac & Windows
.jpg . jpeg .jpe	Image file a common graphics file found on the Internet
.mid	Sound File for use with MIDI sound mixing systems Mac & Windows
.mpeg . mpg	Video format (Motion Pictures Expert Group) Mac & Windows
.mp3	Music file, requires .mp3 player Mac & Windows

.ppt	Powerpoint Presentation Mac & Windows
.ram .ra	Real Audio file Mac & Windows
.swf	Macromedia Flash Mac & Windows

### Can a search engine's advanced search features help me find multimedia files?

Each search engine takes a unique approach to indexing and retrieving multimedia files. Google includes all file types in their main results. Google also has a separate search engine for static images. Alta Vista (<http://altavista.com>) and AlltheWeb (FAST) (<http://alltheweb.com>) provide a rich set of file type options.



Web Images MP3/Audio **Video** Directory News Family Filter: on Help

Video Search Find video files throughout the Web!

**FIND**

Find:  Avi  MPEG  Quicktime  Windows Media  Real  Other

Duration:  Sources:

[Business Services](#) [Submit a Site](#) [Advertise](#) [About AltaVista](#) [Help](#)

© 2003 AltaVista



Search for

all of the words  any of the words  the exact phrase [Simple Search](#)

### Advanced Image Search Features

Transparent Backgrounds  Formats:  All  JPEG  GIF  BMP  
Types:  All  Color  Gray  Line Art

### Advanced Audio Search Features

Streams or Downloads  Formats:  All  MP3  Wave  AIFF  
 RealAudio  MIDI  
Channels:  Both  Stereo  Mono

### Advanced Video Search Features

Streams or Downloads  Formats:  All  AVI  AVI - DivX  QuickTime  
 MPEG

### Offensive Content Reduction

Offensive Content Reduction:

Turning the Offensive Content Reduction filter "on" reduces the amount of offensive material displayed in the search results. Please note that the filter is not 100% accurate, and offensive material will in some cases slip through the filter and be displayed in the search results. Some pages which are not offensive may also be incorrectly filtered out. The filter only supports English documents.

## How can I read a file, if I don't have the proper application software?

If you do not have the right application software, there is often a free, stand-alone "reader" available on the Internet. Additionally, Google *attempts* to translate all files into html, or text formats. Google will highlight the search terms you used making it easier to scan the file.

[\[PDF\] Presentations - ICE Cold Mini Conference October 1, 2005 Mimio Xi ...](#)

File Format: PDF/Adobe Acrobat - [View as HTML](#)

Finding the Good stuff on the Internet, 21st Century **Information Fluency ...**

Bob joined the **IMSA** (Illinois Math and Science Academy) 21CIF Team in July 2002 ...

[chapters.iceberg.org/icecold/mini/p.pdf](#) - [Similar pages](#) - [Remove result](#)

You will see [View as HTML](#) or [View as Text](#) in the snippets describing the file. This allows you to read the file online.

## What about Copyright?



All files and web pages on the Internet are subject to copyright laws. Image, video, and audio files are often specifically copyrighted and cannot be used without explicit permission by the author. Search engines leave it up to the user to determine copyright details and take the necessary steps to insure fair use. (For more, see the IMSA Module: Copyright.)

## Are there specialized search engines for uncommon file types?

Yes, a number of specialized search engines have been developed to help you find non-html files on the Internet. Here are a few:

### Audio and Video:

- **Speechbot Audio** <http://speechbot.research.compaq.com>
- **Singing Fish** <http://www.singingfish.com>

### Images:

- **Ditto.com** <http://www.ditto.com>
- **Pictures Now** <http://www.picturesnow.com/html/index.htm>
- **YotoPhoto** <http://yotophoto.com>

Authored by Dennis O'Connor 2003



[1](#) ... [5](#) [6](#) [7](#) [8]

*End of Micromodule - formats.*

*Return to [Micromodule List](#)*