UNCLASSIFIED#FOR OFFICIAL USE ONLY

Online Videos and Video Search

Online video of all types accelerated greatly over the past year. During the first half of 2006 it became apparent that online video/video search was going to be one of the year's biggest topics and developments, with new video sharing and video search sites gearing up and established companies such as Google, MICROSOFT, and Yahoo jumping on this particular bandwagon. "The market share of Internet visits to the 10 leading online video sites has increased by 164 percent in the past three months (week ending May 20, 2006 versus week ending February 25, 2006)," according to Hitwise, a competitive intelligence company.¹¹⁵

Furthermore, some in the "old media," e.g., ABC, CBS, BBC, AP, and Fox, saw this as a wave of the future they wanted to ride and began or expanded their video offerings via the Internet. To put the video revolution in perspective, "Apple Computer Inc. sold 12 million video clips at \$1.99 each from its popular iTunes Music Store in just a few months" and more than 40 million videos were viewed per day at the YouTube website.

Because we usually need to be able to find and get our hands on video quickly, there is a lot of interest in online video and video search. As with any new venture or technology, there are problems. In the case of Internet video, the main problems at the moment are:

- Quality—lots and lots of really awful, silly stuff.
- Quantity—still only a small number of videos are available compared to webpages, images, music.
- Format—there is no standard format in which videos are offered; some sites require you to download and install their proprietary software while some use MacroMedia Flash, QuickTime, RealPlayer, Windows Media Player, or something else.

¹¹⁵ "Hitwise Data Shows Overall Visits to Video Search Sites Up 164%," Hitwise News Release, May 24, 2006, http://www.hitwise.com/press-center/hitwiseHS2004/videosearch.php (November 28, 2006).

¹¹⁶ Walter Mossberg and Katherine Boehret, "Searching the Web for Video Clips," The Mossberg Solution, The Wall Street Journal, http://ptech.wsj.com/archive/solution-20060301.html (November 28, 2006).

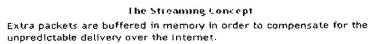
UNCLASSIFIED//FOR OFFICIAL USE ONLY

What blogs are to traditional news, online videos are to television and, to a lesser extent, movies. The spread of broadband connections, cheap data storage, free or inexpensive video technology have all contributed to this boom. In a perverse way, all Qaeda in Iraq was an early adapter of Internet video, making effective if horrendous use of videos to spread its terror message, recruit new members, communicate across its terror network, and even offer training and support.

And, of course, online video has given rise to its own neologisms: **vlogs** (weblogs containing video), vloggers, vlogcasts, vlogcasting, and vlogcasters, as well as <u>vodcasts/vidcasts</u>. I suppose there are even vlogmasters, though that word conjures up images of a character in a Wagnerian opera.

Given the relative ease with which Internet videos can be produced and spread by anyone anywhere, we must take this technology very seriously and learn how to find videos of interest quickly and efficiently. Here are the some of the major players in online video and video search, excluding companies that are offering solely entertainment-related video uploads and downloads. Also, most news sites do not make video clips available for very long, usually no more than a week, before they are archived. Once archived, the clips may require registration and/or payment to view them.

The two basic types of video on line today are downloadable and streaming video. Some of the video sites I discuss below offer downloadable video, which may be saved and played later; this type of video may have a format that requires a specific type of video player. Virtually all the video sites below offer some form of streaming video, that is, one-way video transmissions over the Internet to a compatible media player. These do not require the user to download the video, nor can the video be saved for later viewing. You've seen this "buffering" message before, I'm sure, when requesting video be streamed to your computer:





"Buffering 70% complete" means 70% of a reserved area in memory is filled. When it gets to 100%, the software (Windows Media Player in this example) will start "playing" the video.

¹¹⁷ "Streaming video," Answers.com, Computer Desktop Encyclopedia, Computer Language Company Inc., 2005. http://www.answers.com/topic/streaming-video (15 November 2006).

117

UNCLASSIFIED#FOR OFFICIAL USE ONLY

What you are witnessing is streaming media. While not the smoothest, cleanest technology, it is cheap (the most popular software used for streaming media is free), and with broadband connections, it is relatively fast if sometimes choppy. The major free media players (all handle audio and video) are the following:

- Windows Media Player—Microsoft's proprietary software that plays many audio, video, and streaming formats, including MP3, WMA, CD audio, and MIDI.
- MacroMedia Flash Player—a browser plug-in that, according to the company "is installed on 98% of Internet-enabled desktops worldwide and on a wide range of popular devices." The Flash Player displays content created with MacroMedia Flash and plays files with the .SWF extension.
- MacroMedia Shockwave Player—The Shockwave Player displays content created with MacroMedia Director and plays files with the .DCR extension.

Because there is so much confusion about these two products, MacroMedia created a webpage that compares these two products:

http://www.adobe.com/products/director/resources/integration/

- QuickTime—the only one of these free media players not originally developed for the Windows environment, QuickTime is an Apple Computers' product capable of handling various formats of digital video, audio, text, animation, music, and more. The QuickTime media player is bundled with Apple computers but free downloads are available for Windows as well.
- ➤ RealPlayer—RealNetworks' media player that plays its own proprietary RealAudio and RealVideo as well as other formats including MP3, MPEG-4, and QuickTime. It was one of the first media players capable of handling streaming media over the Internet.

The Major Online Video and Video Search Sites

I decided to list these alphabetically because there is no good way to rank them: each has some advantages and usually quite a few drawbacks. Some things to keep in mind about the current state of video search:

Most video search sites only search the metadata and/or text associated with a video. TVEyes is an exception; it uses voice recognition technology to index every word spoken in a video, but the technology is imperfect. Most video

¹¹⁸ Macromedia Flash Player FAQ, http://www.adobe.com/products/flashplayer/productinfo/faq/-item-1-1 (November 28, 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY

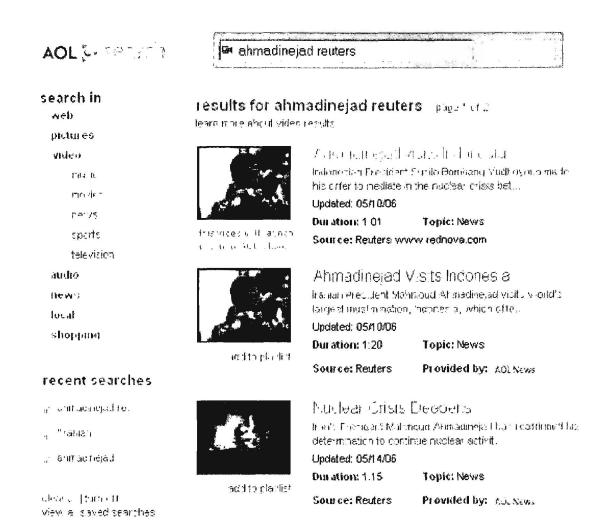
search engines will index the text of television programs with closed captioning.

Many video search sites only search their own video collection and do not search for videos across the web. I will note which sites do this below.

AOL Video Search

http://search.aol.com/aolcom/videohome

AOL Video Search employs the video search engines of not one but two of the companies it owns—Singingfish and Truveo—for its own search site. AOL "Video Search results include multimedia streams and files (Real Media, Windows Media, QuickTime, MP3 and Flash) from AOL and the Web sorted by relevance." Note that you cannot sort the results by date. AOL Video search lets users limit the search to subfields including news, television, music, movies. Phrase searching works well. There is no way to limit your search by source, but if you include a source such as Reuters, AOL Video Search picks up on that keyword:



UNCLASSIFIED//FOR OFFICIAL USE ONLY

Summary:

- Sources of video data: multiple news and entertainment sources from AOL and across the web.
- Media types searched: Real Media, Windows Media, QuickTime, MP3 and Flash.
- Videos viewed: videos at AOL use AOL Video Player; videos from external sites use site's default media player.
- Search options: phrase searching; advanced search; no search by source option.
- Sort options: by relevance only.
- > Upload video? Not directly; only via AOL's companies Singingfish or Truveo.

http://news.bbc.co.uk/

BBC does not make it easy to find its news videos. For example, there is no separate BBC News video page. Instead, to view a BBC video, users must go to any News Online page and click the "Watch BBC news in video" or "Latest news in video and audio" button at the top of the page. This opens a new window where users can select from a list of current videos by news topic: headline, UK, business, politics, health, etc. The videos play in the same window using either RealPlayer (preferred) or Windows Media Player. Another option is BBC News 24, constantly updated national and international headlines. The best way to find BBC videos is using a third-party site such as Yahoo Video Search, which offers a search by source option.

In May the BBC announced it was opening its news archives, but only for users in the UK. If you are in the UK, "You can download nearly 80 news reports covering iconic events of the past 50 years including the fall of the Berlin Wall, crowds ejecting soldiers from Beijing's Tiananmen Square and behind-the-scenes footage of the England team prior to their victory over West Germany in 1966." http://www.bbc.co.uk/calc/news/

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Summary:

- Sources of video data: BBC news video only.
- > Videos viewed: Real Media (preferred) or Windows Media at BBC site.
- Search options: no separate video search.
- > Sort options: by category (e.g., headlines or business news).
- Upload video? No.

Blinkx http://www.blinkx.tv/

Blinkx certainly has one of the snazziest sites, which is not necessarily a good thing because it can be distracting. However, it is a very good video search tool. Blinkx uses voice recognition software to transcribe the content of audio and video material, whether it is commercial television or personal vlogs. Users simply enter search terms in the box on the left of the screen and, once the results are returned, a slider appears that lets the user determine whether to rank the results by date or relevance or somewhere in between. Here is a snapshot on the Blinkx page showing results for [ahmadinejad] sorted by relevance.



UNCLASSIFIED#FOR OFFICIAL USE ONLY

One of the best features in Blinkx is that it shows animated thumbnail images of each video. In some cases, Blinkx plays the full clip in the Blinkx window without users having to go to another website. Here is the key for determining the type of video and where it will play:



The blinkx b and camera icon located at the end of the result heading indicates that this result's video is hosted by blinkx.tv and can be viewed on the blinkx.tv Direct Play Screen located to the right of the result.

- The eye and camera icon located at the end of the result heading indicates that this result's video can be viewed only from the original site.
- The blinkx podcast icon located at the end of the result heading indicates that this result is a podcast.

Blinkx is much better than other video search sites at weeding out videos that have been archived and are thus no longer available at the originating site, although sometimes you want to know that a video was once available even if you cannot access it now.

One other thing you need to know about Blinkx is how to use the "channel" buttons on the left side of the homepage. Here is the "news" channel; users can select all, some, or none of these channels to search by clicking on the buttons. In this case, the news channels with the green dots have been selected to be searched:



Blinkx employs keyword and Boolean search options. Blinkx also has a special "Self Casting" channel for videos uploaded by users. Most of these are probably of dubious value, but it's worth keeping an eye on them.

UNCLASSIFIED#FOR-OFFICIAL USE ONLY

Summary:

- > Sources of video data: multiple news, entertainment, radio, <u>podcasts</u> from across the web.
- Video viewed: videos hosted on Blinkx servers use "Direct Player"; videos from external sites use site's default media player.
- Search options: keyword, Boolean, & phrase searching; "conceptual search"; option to limit search by source.
- > Sort options: using slider, sort by date, relevance, or somewhere in between.
- Upload video? Yes.

CBS News Video Search

http://www.cbsnews.com/sections/i_video/main500251.shtml

CBS News has both an excellent video search, which appears to pull from the entire archive of CBS News videos. For example, I searched for ["hurricane isabel"] and found the very first CBS News video titled "Isabel Starting to Scare," September 12, 2003. Unfortunately, there is no way to force the search to look for a specific date, although there is the option to limit the search to specific topics (e.g., US news), to a specific CBS News show (e.g., "60 Minutes"), or to search in the Video Library by topic. However, there are almost no special search options, which makes it hard to find the nuggets in this treasure trove of videos. Despite this drawback, I believe the CBS News video search and archive is one of the great secret resources on the web. It is free to view the full videos, which open in the CBS News.com window using either RealPlayer or Windows Media Player (both free). All videos before November 20, 2003, play only in Real.

- Sources of video data: CBS News video only.
- Video viewed: RealPlayer or Windows Media Player.
- Search options: keyword searching; search by ANY, ALL, EXACT PHRASE; no date limit.
- Sort options: sort by relevance or date.
- Upload video? No.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

CNN Video Homepage

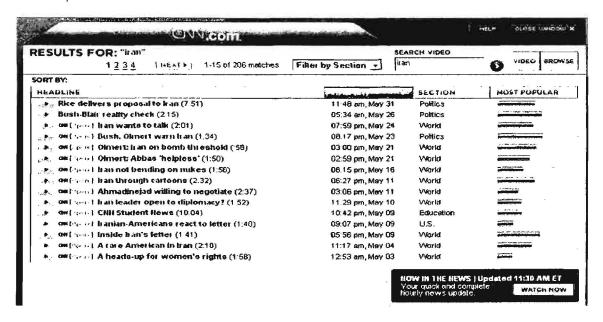
http://www.cnn.com/video/

CNN Video Almanac

http://www.cnn.com/resources/video.almanac/

CNN offers only seven days of free access to its vast video archive of news and features. After a week, users must sign up for <u>CNN Pipeline</u>, which requires both payment and downloading and installing a CNN video player.

Before CNN archives its video, you can search, browse, and view the free videos at the CNN website using Windows Media Player. All free videos are preceded by commercial advertisements. I find the CNN video set up very frustrating to use. If you go to the main CNN Video page http://www.cnn.com/video/, there is no way to limit your search to videos. However, if you go to the CNN homepage http://www.cnn.com/, scroll down to the middle of the page to watch video, and click on any video, the popup window includes a video only search box. Enter your search terms and CNN Video will return a list of video matches by date and time. You can then resort the results by Section or Most Popular. Note in this example that only the first two videos are still free. The rest have been archived to CNN Pipeline.



CNN has one other video feature you need to know about: its **Video Almanac**, a collection of the best video since the birth of CNN in 1980 through 1997. Users can select a year or topic, e.g., the video of the failed attempt to rescue the hostages in Iran in April 1980. The Video Almanac provides a very limited number of videos, most of which are very high profile events, though there are notable exceptions ("George Burns dies"). The videos play using QuickTime.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Summary:

- Sources of video data: CNN video only.
- Video viewed: Windows Media Player version 9 and above.
- Search options: search for one or more keywords or exact phrase in double quotes; filter by section; no date limit.
- Sort options: sort by date/time, section, or popularity.
- Upload video? No.

C-SPAN

http://www.c-span.org/

C-SPAN Store

http://www.c-spanstore.org/shop/

C-SPAN is the only one of the online video and video search sites that is a non-profit entity. In case anyone is unclear who and what C-SPAN is, here's what the network says about itself: "C-SPAN is a private, non-profit company, created in 1979 by the cable television industry as a public service. Our mission is to provide public access to the political process. C-SPAN receives no government funding; operations are funded by fees paid by cable and satellite affiliates who carry C-SPAN programming." http://www.c-span.org/about/index.asp?code=About

Until a video is archived, it is available for free at the C-SPAN site. C-SPAN offers both stored and streaming (live) media (video and audio, because there is also a C-SPAN radio). It is very important to remember that free C-SPAN video searches only index program titles, event descriptions, and air dates. "Searches are not based on text within the video files or closed-caption transcripts and do not include video available at the C-SPAN Archive http://www.c-spanstore.org/shop/, where you can buy almost any C-SPAN program." Because the C-SPAN search is not based upon text or closed-captioning, you may have to work harder to find a specific video.

Although the C-SPAN site says that "most events will remain in the archive for 15 days or less," I found many videos going back years that are still available, so it appears to be hit or miss as to what is or is not available for free. Keep in mind you can find and view many but not all Congressional hearings at C-SPAN. The site has an advanced search option that lets users limit a search by date ranges as well as by C-SPAN series and programs. The advanced search also offers the option to search by phrase and even a fuzzy word search ("spelled like" or "sounds like").

The C-SPAN video collection includes videos from these and other sources:

UNCLASSIFIED#FOR OFFICIAL USE ONLY

- Congressional hearings.
- ➤ US Government Executive branch (e.g., DoD, State Department) press conferences.
- Special commissions.
- Some State and Local videos.
- > C-SPAN TV series, such as "Washington Journal," "Booknotes," etc.

Summary:

- Sources of video data: C-SPAN video (and audio) only.
- ➤ Video viewed: Stored clips require RealPlayer; streaming media offers a choice between RealPlayer and Windows Media Player.
- > Search options: search for one or more keywords; advanced search offers date, program filters and fuzzy keyword search.
- Sort options: sort by relevance or date/time.
- ➤ Upload video? No.

Google Video

http://video.google.com/

Google Video underwent a huge metamorphosis during December 2005. Google Video moved from being primarily a video search engine to being a store front to preview and buy video. This time Google's target isn't Microsoft but (primarily) Apple's iTunes and to a lesser degree cable television companies. Google Video is more focused now on selling videos, including television shows, than on searching current videos. However, there are many free videos still available on pretty much any subject you can imagine. Very importantly, Google Video only searches for videos hosted at its own website. Google Video has a link to AOL Video on its homepage, but the link takes you to the AOL Video website and Google Video does not search on AOL's videos from its site.

In May 2006 Google Video (finally) wised up and let users upload their videos directly online without having to install any software and without the previous submission time lag of up to 24 hours. ("Google Inc. Tweaks Its Video Service," AP report in Forbes.com,

http://www.forbes.com/technology/ebusiness/feeds/ap/2006/05/17/ap2751923.html) YouTube had always permitted direct video uploads and was basically cleaning Google Video's clock.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Google indexes the closed captioning and text descriptions of all the videos in its archive to facilitate search. Google Video search supports keyword and phrase searches. While Google Video does not support search by source, it does offer the option of searching by title. The correct syntax is [title:keyword]. Note that phrase searching does not work with the title: syntax.

Google Video has several options for viewing or limiting results. Users can choose to view the results as a list or as a grid (the list provides more information while the grid shows a larger image of the video clip). There is also an option to limit the results by the length of the video—long (20+ minutes), medium (4-20 minutes), or short (up to 4 minutes). Results can be sorted relevance, date, or title.

Google explains how it ranks videos in terms of popularity: "We use algorithms to identify videos that are suddenly becoming popular, and then rank them based on how popular they are—and how suddenly they became popular. We've been using this list internally, and now it's ready to share with you, so check it out. Right now this feature highlights videos from close to 40 countries, including Argentina, Australia, Belgium, Brazil, Canada, Estonia, Finland, Greece, Hong Kong, India, Israel, Japan, South Korea, Mexico, and New Zealand, to name a few." (Jon Steinback, "Movers, Shakers, and Hoops on Video," Google Blog, June 9, 2006, http://googleblog.blogspot.com/2006/06/movers-shakers-and-hoops-on-video.html)

The software required to view streaming video at Google Video is Macromedia Flash Player 7.0+. To view purchased and downloaded videos, users still must install Google's own Google Video Player, which has frustrated many users and infuriated others. It will be interesting to watch Google Video's evolution as the site tries to move higher in online video popularity. closed captioning and text descriptions of all the videos in our archive for relevant results.

- > Sources of video data: only news, entertainment, radio, podcasts hosted on Google servers.
- Video viewed: free streaming media uses Flash Player; purchased videos require Google Video Player.
- > Search options: keyword & phrase searching; title: search; no search by source option.
- > Sort options: sort by relevance, date, or title; view as List or Grid; limit results by length of video.
- Upload video? Yes.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

IFILM http://www.ifilm.com/

IFILM "is a leading online video network, serving user-uploaded and professional content to over ten million viewers monthly. IFILM's extensive library includes movie clips, music videos, short films, TV clips, video game trailers, action sports and its popular 'viral videos' collection. IFILM is one of the leading streaming media networks on the internet." http://www.ifilm.com/about/> Viacom purchased IFILM and made it a part of its MTV Network, so most of the commercial videos are from some segment of that network. However, what distinguishes IFILM is its vast supply of user created and uploaded videos.

IFILM specializes in viral videos, that is, videos that gain widespread popularity across the Internet through blogs, email, IM, websites, and old-fashioned word of mouth. Not all viral videos are humorous. One example of a viral video from IFILM was "Mercenary Sniper in Iraq," described in Defense Review.com as an "interesting 'Viral Video' clip that puts the viewer inside a helicopter in Iraq (urban environment), and then puts you right there with a sniper team engaging hostile enemy targets (i.e. insurgents/terrorists) from a rooftop (again, inside Irag)."

http://www.defensereview.com/article826.html

Summary:

- > Sources of video data: multiple news, entertainment, and user-created videos.
- Video viewed: QuickTime, Windows Media Player, Macromedia Flash.
- > Search options: keyword & phrase searching; limit search to specific collections, such as "User Video."
- Sort options: no.
- Upload video? Yes.

MSN Video¹¹⁹

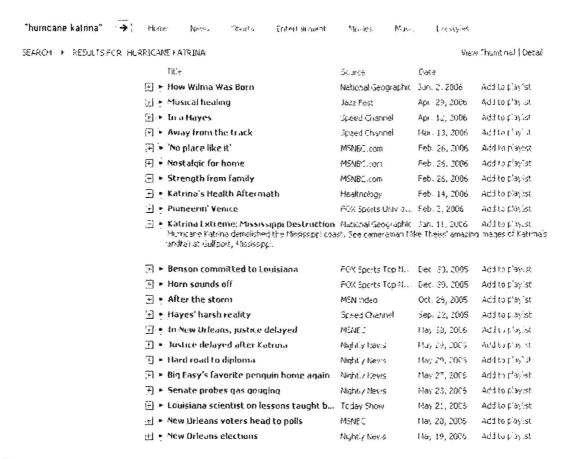
http://video.msn.com/

Don't bother trying to access this site from any browser other than Internet Explorer 6. Once you get past that extremely annoying requirement, MSN Video is not a bad site. Users have access to all NBC news, sports, and entertainment, as well as other sources such as the National Geographic Channel and the Discovery Networks. MSN Video also partners with IFILM, so you can access IFILM videos from this site.

¹¹⁹ As of this writing, MSN Video has not been renamed Live Video even though there is a video search incorporated into Live Search.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

I cannot determine exactly how long videos are kept; some videos date back years, but most of the news videos appear to be from within the past 30 days. What I like best about the site are the many options for handling results. Users can choose between thumbnails or details; view all results; and sort by title, source, or date. Here is a screenshot of the results for the query ["hurricane katrina"] in detail view sorted by date:



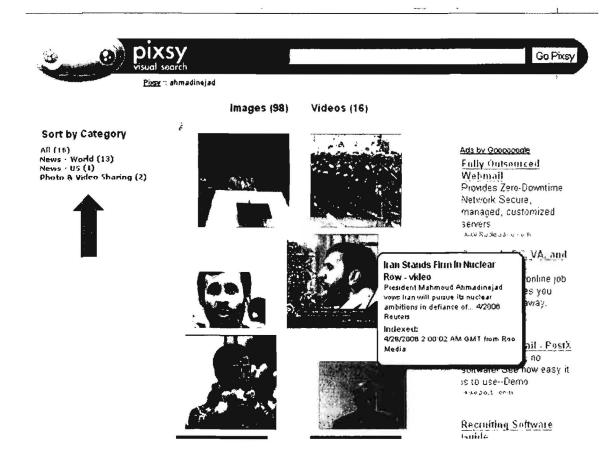
- Sources of video data: selected news, entertainment, etc., from NBC and MSN's partners as well as user-created videos from IFILM.
- ➤ Video viewed: Windows Media Player 7+, Macromedia Flash 7+ (site only works in Internet Explorer 6).
- Search options: keyword & phrase searching; limit search to specific collections, such as News.
- > Sort options: by title, source, or date.
- Upload video? No.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Pixsy http://pixsy.com/

Pixsy is a metasearch engine for both still images and video that searches a large numbers of video content providers. "As a meta-aggregator of image and video thumbnails images from RSS providers, Pixsy provides a visual search alternative for consumers and a source of free, high quality traffic for RSS syndicating image & video content providers. New RSS providers include YouTube, Revver, SmugMug, RollingStone, StumbleUpon, Defamer, People Magazine, Pictopia, Metacafe, TheOnion, Rotten Tomatoes, Buzznet, CNN, NPR, PBS, and many more." (Loren Baker, "Pixsy Adds New Video Search Content from YouTube, CNN, Defamer & Others," Searchenginejournal.com, http://www.searchenginejournal.com/?p=3500)

The thumbnails are small enough to review quickly but still of good quality, and a mouseover of each thumbnail image reveals the title of the video, a summary of its contents, the date of the video, its source, and the date and location where it was indexed. The only sort options at Pixsy are by category. The Pixsy homepage also links to the latest and to featured videos.



UNCLASSIFIED#FOR OFFICIAL USE ONLY

Summary:

- > Sources of video data: multiple news, entertainment, radio, podcasts from across the web.
- Video viewed: videos play using external site's default media player inside a Frame at Pixsy with option to remove Frame.
- Search options: keyword and phrase searching; no search by source option; no advanced search.
- Sort options: searches for both images and videos; tab to video; only sort option is by category.
- Upload video? No.

Reuters Video

http://today.reuters.com/tv

I have to mention Reuters because of its importance and reach, but the site is frustrating to use because there is no separate video search at the site. All you get is a list of available videos sorted by topic (news, entertainment, business, etc.). If you need Reuters' videos, I recommend using one of the video search sites—RocketInfo, SingingFish, RooTV, Blinkx, Yahoo Video Search, or Pixsy.

RocketInfo

http://www.rocketnews.com/ [select the VIDEO tab]

RocketInfo offers some of the most extensive sources for news and video on the web, drawing from over 16,000 news sources on a continuous basis. The downside is that RocketInfo only searches for the past four days at most, and users can limit a search to today's videos. The search has no advanced features, not even phrase searching. It looks for all the keywords, so be careful not to search on too many terms.

RocketInfo is a metasearch engine that searches many video sources, including BBC, RedOrbit videos, Reuters, CBS, CNN, MSNBC, local news, etc. For current news searches, RocketInfo has to be ranked as one of the best and most comprehensive.

- Sources of video data: news from over 16,000 sources across the web (not all or even most of these sources offer video)
- Video viewed: all videos play at originating site using that site's default video player
- Search options: keyword search only (searches for ALL terms)

UNCLASSIFIED//FOR OFFICIAL USE ONLY

> Sort options: by date or relevance

Upload video? No

<u>http://www.rootv.com/</u>

RooTV or Roo is a powerful video search tool, but I find it annoying because it automatically plays a video when you visit the homepage (and it really doesn't want to stop!). There is a simple keyword search but no way to limit the search to just news, for example. The videos play at the RooTV site, using Windows Media Player or RealPlayer. RooTV requires ActiveX; if you are going to use RooTV, I recommend opening the site in Internet Explorer (you will have to add it to your Trusted Zone) so you do not need to install an ActiveX plug-in in Firefox or Netscape, something I would discourage. Sources include Reuters and AP. Supported by advertising.

Summary:

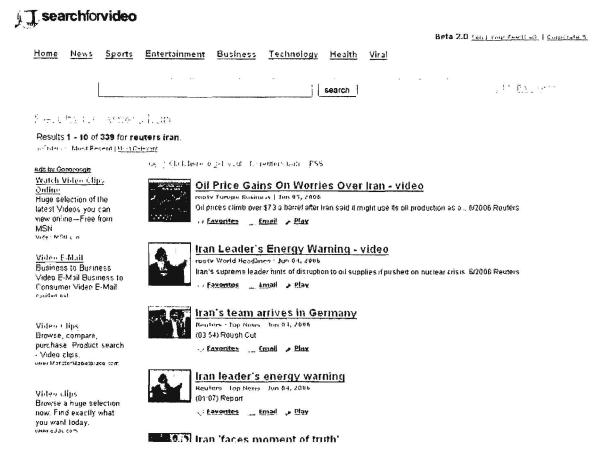
- Sources of video data: news from sources across the web, including Reuters and AP.
- Video viewed: all videos play at the RooTV site using either Windows Media Player or RealPlayer; requires ActiveX.
- Search options: keyword search only (searches for ALL terms).
- Sort options: none.
- Upload video? No.

<u>Searchforvideo</u>	http://www.searchforvideo.com/home/index.html
Searchforvideo IM Service	http://www.searchforvideo.com/misc/im.jsp
Searchforvideo Reel Time Lens	http://www.searchforvideo.com/misc/reel.jsp

Searchforvideo is one of the best video search engines available. Searchforvideo searches a vast number of sources, including BBC, ABC, CBS, CNN, AP via iVillage, Reuters via RooTV, RedOrbit, the Discovery Networks, and many others. The site also has separate pages by topic—news, sports, entertainment, business, technology, health, and viral—where users can browse the top videos of the day.

Interestingly, even though there is no search by source, I found that if you use the name of a source and a keyword, that seems to work very well because you are searching on all metadata associated with a video clip. For example, [reuters iran] returns only videos from Reuters sources containing the keyword "iran":

UNCLASSIFIED#FOR OFFICIAL USE ONLY



Searchforvideo also offers IM search that works with AOL, Yahoo, and MSN. There is also a very interesting option to view videos that are being added or watched in real time at the Searchforvideo Real Time Lens page.

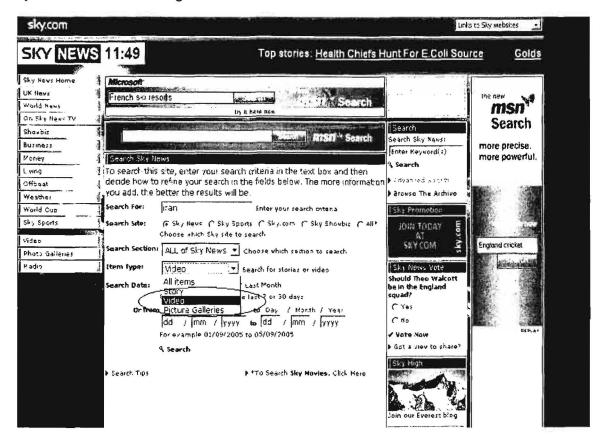
- > Sources of video data: searches wide variety of video sources across the web
- Video viewed: at originating site using that site's default video player, except...
- Offers video podcasts that can be downloaded to video podcast software at Searchforvideo Publishers http://www.searchforvideo.com/pub/>.
- Search options: keyword searching; add source keyword, e.g., [bbc], to limit by source.
- Sort options: sort by relevance or date.
- Upload video? No.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Sky News Video

http://www.sky.com/skynews/video

The UK's Sky News, billed as Europe's first 24-hour news channel, offers a "video channel" and advanced search to limit queries to videos only. Keep in mind that the simple search on the main video page does not restrict your search to videos: you must use the advanced search page. Here users have the option to limit the search by "item type," and Video is one of the types. Notice you can also limit your search by section and date range.



In the advanced search, the only characters recognized are letters, numbers, and hyphens. You can use a hyphen to create a phrase: [tony-blair]. Results are listed in date order with the most recent items listed first. Each result shows the relevance to your search. The videos play using Windows Media Player.

- Sources of video data: Sky News video only.
- Video viewed: at Sky News site using Windows Media Player.
- Search options: keyword searching; query terms limited to alphanumerics and hyphens; search by section (e.g., news), type (e.g., video), date range.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

- > Sort options: None; listed by relevance.
- Upload video? No.

TVEyes http://tveyes.com/

If you want to search the content of TV and radio news broadcasts, I recommend the TVEyes search engine. TVEyes captures, indexes, analyzes, archives, and distributes in real-time content captured from television and radio broadcasts. "Currently live content alerts and searchable archives are available from:

- > US
- > UK
- Canada
- Australia

and the Al Jazeera network. Coverage is being extended to include France, Mexico, Guatemala, South Africa with these countries on stream by Q4." http://www.tveyes.com/coverage/index.htm>

TVEyes uses voice recognition technology to create something the company calls its Spoken Word Index(TM), so that users can search every word in a broadcast. The search only seems to work if you use one keyword or a phrase because it interprets more than one word as a phrase. It works very well finding the right broadcast; unfortunately, once you go to the originating site of the news clip, you may well discover it is no longer available. However, TVEyes can help you locate a specific clip, especially if you did not know the source. You may have to pay for that video, but at least you know where to go to get it. TVEyes offers thumbnail versions of the clips, including some video no longer available at the originating site, which play from the TVEyes site. Full clips must be viewed at the originating site.

- Sources of video data: television and radio broadcasts in US, UK, Canada, Australia, Al Jazeera network (to be extended in 4th Q 2006 to many other international sources).
- ➤ Video viewed: thumbnail versions of the clips, which play from the TVEyes site. Full clips must be viewed at the originating site and may no longer be available.
- > Search options: keyword; more than one word is interpreted as a phrase.
- Sort options: sort by date or relevance.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

➤ Upload video? No, but TVEyes "can respond immediately to any request to add TV channel or Radio station coverage."

Yahoo Video Search

http://video.yahoo.com/

Yahoo News Video

http://news.yahoo.com/video

Yahoo video search is one of the most popular on the web and for good reason. It draws from many sources across the web and around the world, and it has a powerful advanced search. Yahoo video search is one of the few that has an explicit site delimiter. If, for example, I only want to search for videos from the BBC, on the advanced search page, I would select "only search in this domain/site:" and enter [bbc.co.uk]. The syntax for this search (with or without keywords) is [fromsite:bbc.co.uk] if you prefer to use the simple search interface. The other advanced search options are to limit your search by format, size, or duration. Yahoo video advanced and simple search recognize + (must include), - (must exclude), and double quotes for a phrase.

Yahoo video search has another excellent option in the form of channels. Channels are collections of videos created by a common source; clicking on a "channel" displays all video results from that source. Unfortunately, there does not appear to be any central list of all channels. The best way to find the channels is to do the following: on the Yahoo video homepage, click the "Categories" tab, then select "News." Now you will see the channels listed, e.g., ABC News, CBS News, CBS.ca, AP, Reuters News, etc. Selecting a specific channel will show you all the available videos from that channel. To save channels to your list of favorites, you must be a registered member of Yahoo

In late 2006, Yahoo and CBS "announced an exclusive video syndication agreement in which local news video from 16 of CBS's owned stations will be made available on Yahoo! to the Internet's largest news audience. The relationship...marks the first video agreement between a network-owned television station group and an Internet news provider." Yahoo's video focus has always been on commercial and not homemade video, the opposite of the YouTube/Google approach, even though Yahoo does have a video site that tries to compete with YouTube. Yahoo already had deals with CBS's "60 Minutes," as well as with ABC, Disney, and CNN. This deal is with local news CBS affiliates. While some are criticizing Yahoo's approach to offering video from commercial/professional sources, it is an approach that continues to work for them. When users want to find and view news videos, they do not think YouTube, they think Yahoo News Video. I use it heavily not only because of the number and quality of videos available but also because it is very easy to use,

¹²⁰ Michael Liedtkey, "Yahoo adds CBS news to video lineup," AP/San Francisco in Businessweek.com, 16 October 2006,

http://www.businessweek.com/ap/financialnews/D8KPGBUO1.htm (24 October 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

the videos open at the Yahoo site, and you do not need to register to find and view the videos.

Summary:

- > Sources of video data: searches wide variety of video sources across the web, including user-submitted videos.
- Video viewed: at originating site using that site's default video player.
- ➤ Search options: keyword searching; + (must include), (must exclude), and double quotes for a phrase; limit search to specific site [fromsite:domain]; search by format, size, or length of video.
- Sort options: None; listed by relevance.
- Upload video? Yes.

YouTube

http://www.youtube.com/

If you have heard of only one online video and video search site, it's probably YouTube, the most popular such site by far, garnering almost 43 percent of all visits to video websites in mid-2006. YouTube is currently serving 100 million videos per day, with more than 65,000 videos being uploaded daily. YouTube started as a personal video sharing site and opened to the public in February 2005. While other video sites try to mix commercial and homemade videos, individual users create almost everything at YouTube. However, because of YouTube's phenomenal success, some commercial enterprises have latched onto the site's popularity to promote their own products via videos. Most notably, YouTube now offers official movie trailers and promotion videos from established media companies.

It came as a surprise to no one when Google bought YouTube in late 2006 (when Video replaced Froogle on the Google homepage, that was a pretty clear indication of Google's direction). But this is a gamble, possibly a big gamble. Not everyone has forgotten what happened to Napster; Mark Cuban, chairman of HDNet (among other things) says in his blog, "It will be interesting to see what happens next and what happens in the copyright world. I still think Google Lawyers will be a busy, busy bunch. I dont think you can sue Google into oblivion, but as others have mentioned, if Google gets nailed one single time for copyright violation, there are going to be

¹²¹ "Hitwise Data Shows Overall Visits to Video Search Sites Up 164%," Hitwise Competitive Intelligence, 24 May 2006, < http://www.hitwise.com/press-center/hitwiseHS2004/videosearch.php (24 October 2006).

¹²² YouTube Fact Sheet, YouTube.com, http://www.youtube.com/t/fact_sheet (24 October 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY

more shareholder lawsuits than doans has pills to go with the pile on copyright suits that follow." 123

However, YouTube was in the process of cleaning up its copyright act before the acquisition, and Google is certain to ensure that happens. Both Google and YouTube have recently made deals with major video partners, so Google may dodge most of the lawsuits, especially if video producers realize there is serious money to be made with this partnership. Nonetheless, big, rich Google makes a much more tempting target for potential litigants than YouTube ever did. For now Google says that YouTube will continue "to operate independently to preserve its successful brand and passionate community." 124

Nonetheless, as of the first of this year, Google Video began to include results from YouTube. For now, when users click on the YouTube results, they are taken to the YouTube website. As of this writing, YouTube videos do not appear on the Google video homepage, only in search results where they are recognizable from their address.

With YouTube, people can:

- Upload, tag and share videos worldwide.
- Browse millions of original videos uploaded by community members.
- > Find, join and create video groups to connect with people who have similar interests.
- Customize the experience by subscribing to member videos, saving favorites, and creating playlists.
- ➤ Integrate YouTube videos on websites using video embeds or APIs.
- Make videos public or private—users can elect to broadcast their videos publicly or share them privately with friends and family upon upload.

YouTube is building a community that is highly motivated to watch and share videos. The service is free for everyone." http://www.youtube.com/t/about>

In May 2006 YouTube launched a new service that allows users to upload videos directly from their mobile phones and PDAs to the YouTube website. Clearly, two big concerns for a site such as YouTube are copyright infringement and pornography,

¹²³ Mark Cuban, "I Still Think Google is Crazy,": Blog Maverick: The Mark Cuban Weblog, 9 October 2006, http://www.blogmaverick.com/2006/10/09/i-still-think-google-is-crazy/ (24 October 2006).

[&]quot;Google To Acquire YouTube for \$1.65 Billion in Stock," Google Press Center, 9 October 2006, http://www.google.com/intl/en/press/pressrel/google_youtube.html (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

both of which they constantly try to avoid, in part by employing a "community policing" policy where users "turn in" offenders. When found, copyrighted material and pornography are removed.

Lest you think every video at YouTube is of the "Snakes on a Plane" variety, there is so much user-created video available that there is at least a little bit of everything as well as a lot of some things. Take a look at this snapshot of the News & Blogs category, which shows the most popular tags in this category and thumbnail images of the most popular videos. Video tags created and added to uploaded videos is crucial to the success of any video being found using search, which motivates video creators to properly and adequately tag their videos. However, the tags are solely the discretion of the video's creator (i.e., there is no standard taxonomy).



News, Biogs, Local Issues.

Popular Tags in News & Blogs

77 osama Cilish pil di l'Uneral asozial techno whigiteld 911 rave y/k ultrabelat be d'ance <u>waich firmy</u> america schwarzenegger american bigot ban brooke fun heute apht september corian melody born

terror Videos in News & Blogs 1-15 of 559







Kurds against Turkish TERRORISM Part 2 07:39 is ideal 1 day ago Fra Cristians PHK V 4 48 ** 3 ratinos



The Truth 00:59 / 15 to 1 day ago to interesting owner Sel 13



the real terrorists 02:49 .cd i 1 day ago गतम विशिष्ट्य 16 N 192 **** 1 retina



Hitnakut, singing girls NVei Dekalin (?) 05:22 An hal 1 day ego from lisvaschad 5 iv 22

We cannot afford to ignore sites such as YouTube because this is where users go every day to share and view their own homespun videos. It is an important new form of communication via the Internet.

- Sources of video data: only searches for user-submitted videos hosted at its own website.
- Video viewed: at YouTube using MacroMedia Flash.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

- > Search options: keyword searching for tags from uploaded videos; double quotes for a phrase will search for a video title.
- ➤ Sort options: Relevance, date added, title, view count, users' rating (using star system).
- Upload video? Yes.

Sites requiring registration, payment, and/or software downloads

AP Archive

The Associated Press video archive is different from the other video search services in several important ways. First, it only searches the AP archive, but that is hardly a small thing. With over half a million stories in the archive, the AP probably has what you want. There is a catch: this is not a free service. The process for ordering video clips is complicated, but if you really need a specific video, you should read the AP Archive's "How We Work" page to learn the particulars.

http://www.aparchive.com/APArchive/pages/admin/how-we-work.html

However, you do get some very useful data from the site for free. Look at the information provided about a May 2006 video entitled "Iran Nuclear":

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Tale: Iran Nuclear

Tape Number: EF08/0422

Duration: 00:01:53 latime: 10:51:09 Date: 2006-05-17 Source: IRINN SHOTLIST:

1. Zoom-out from banners to wide of rally

2. Close-up of people holding picture of Ahmadinejad

3. SOUNDBITE: (Farsi) Mahmoud Ahmadinejad, Iranian President.

"They say they want to give us incentives! Do you think you are dealing with a four-year old child to whom you can give some walnuts and chocolates and get gold from him?"

4. Cutaway crowd waving and chanting

5. SOUNDBITE: (Farsi) Mahmoud Ahmadinejad, Iranian President

"I tell you that we do not want anything beyond our legitimate rights. We want our rights within the NPT (Non Proliferation Treaty) and we will not accept one lotalless or more than our rights."

6. Pan of crowd

7. SOUNDBITE: (Farsi) Mahmoud Ahmadinejad, franian President:

"Don't force governments and nations to renounce their membership of the Nuclear Nonproliferation Treaty."

8. Crowd chanting and waving flags

9. SOUNDBITE: (Farsi) Mahmoud Ahmadinejad, Iranian President.

"The Iranian nation won't accept any suspension or end to its nuclear activities,"

10. More of crowd at rally

STORYLINE:

Iran's President Mahmoud Ahmadinejad on Wednesday rejected a European plan to offer his country incentives, including a light-water nuclear reactor, in return for giving up uranium enrichment.

"Do you think you are dealing with a four-year old child to whom you can give some walnuts and chocolates and get gold from him?" Ahmadinejad told thousands of people in central Iran, in a speech broadcast live on state television.

European nations have weighed up adding a light-water reactor to a package of incentives meant to persuade Tehran to permanently give up uranium enrichment - or face the threat of UN Security Council sanctions.

On Tuesday senior diplomats and EU government officials said tentative plans were being discussed by France, Britain and Germany as part of a possible package to be presented to senior representatives of the five permanent UN Security Council members.

Ahmadinejad also repeated his threat to pull out of nuclear Non Proliferation Treaty (NPT) if international pressure to give up uranium enrichment continued.

"Don't force governments and nations to renounce their membership of the Nuclear Nonproliferation Treaty," he said, asserting that Iran had the right to a civilian nuclear power programme.

"The Iranian nation wont accept any suspension or end to its uranium enrichment activities," Ahmadinejad said.

He also said from trusted the European Union in 2003 and suspended its nuclear activities as a gesture to boost negotiations over its nuclear programme - only to have the Europeans eventually demand from permanently half the programme.

The 2003 deal called for guarantees that Iran's nuclear programme wouldn't diverge from civilian ends toward producing weapons.

Iran agreed to the request, but negotiations collapsed in August 2005 when the Europeans said the best guarantee was for Iran to permanently give up its uranium enrichment programme.

Iran responded by resuming uranium reprocessing activities at its uranium conversion facility in Isfahan, central Iran.

Earlier this year it resumed research and uranium enrichment after the International Atomic Energy Agency, the UN nuclear watchdog, referred Tehran to the UN Security Council.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

As you can see, the "storyline" provides extensive information about this clip and is free at the AP Archive website. Users would be hard-pressed to find more or better news video and text than in the AP Video Archive.

AP Video Archive

http://www.aparchive.com/

CNN Pipeline

CNN offers only seven days of free access to its vast video archive of news and features. After a week, users must sign up for CNN Pipeline, which requires both payment and downloading and installing a CNN video player. Here is what CNN Pipeline offers:

"CNN Pipeline empowers you to watch up to four live news feeds at once, changing feeds at any time with a single click. Our free video player only allows you to view one video at a time. Additionally, while our free video contains commercial advertisements, Pipeline offers commercial free access to video content on-demand, including free video, with multiple features for ease in searching and browsing. Our extensive archive, not available through the Free Video Player, gives you unlimited access to search and browse CNN.com's online video library." http://www.cnn.com/help/pipeline/#28

CNN Pipeline

http://www.cnn.com/pipeline/index.html

Vanderbilt Television News Archive

If you must have a video of a news broadcast and you are willing to register and pay for it, the Vanderbilt Television News Archive is the place to go. "The Television News Archive collection at Vanderbilt University is the world's most extensive and complete archive of television news. The collection holds more than 30,000 individual network evening news broadcasts from the major U.S. national broadcast networks: ABC, CBS, NBC, and CNN, and more than 9,000 hours of special news-related programming including ABC's Nightline since 1989...The archive makes two kinds of video tape loans to clients: duplications and compilations. We charge fees for loans to offset the costs in providing this service." http://tvnews.vanderbilt.edu/

Vanderbilt Television News Archive

http://tvnews.vanderbilt.edu/

Conclusion

There seems to be no end to the number of sites that provide some sort of online video access, whether it is news or homemade videos, for-fee sites or free ones, archives or live feeds. Every time I look, I find more sites, but I had to decide where to draw the line and when to stop compiling sites. If you believe I missed a critically important source, please let me know (and tell me what features make it special).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Online Audio, Podcasts, and Audio Search

Last year I predicted that podcasting would be the "next big thing" for the Internet. At first, I thought I was wrong. While podcasting initially garnered a lot of attention, it did not take off until the second half of 2006. Fueled in large part by the immense popularity of digital audio devices such as iPod and the spread of broadband connections, podcasts and online audio have revolutionized the way people share and get news, entertainment, and information. While music downloads continue to dominate online audio, podcasting is coming on strong. And podcasting is not just for audio: video podcasting—know as **vodcasting** or **vidcasting**—is increasing in popularity, too. Podcasting is recording and broadcasting any non-musical information—be it news, radio shows, sporting events, audio tours, or personal opinions—for playback on a computer or a mobile device. "Though podcasters' web sites may also offer direct download or streaming of their content, a podcast is distinguished from other digital audio formats by its ability to be downloaded automatically using software capable of reading feed formats such as RSS or Atom."

Think of a podcast as analogous to a radio or television broadcast. The podcaster first makes a file available on the Internet, either an audio file (usually in MP3 format) or video file 126 (these can be in a number of formats). Then the podcaster announces the availability of the file using a feed (RSS or Atom) that lists the available podcasts (very much like individual radio or TV shows) with the title, date, and a short text description of each episode. Finally, the user either plays the file on a computer or downloads the file to play it on a portable device such as an iPod capable of playing MP3 and/or video files.

Podcasting has caught on because it is easy, inexpensive, mobile, flexible, and powerful. Many websites now serve as directories to help users find podcasts of every variety anywhere in the world. Podcast search engines generally index podcast metadata such as title, description, and length, which usually makes searching for podcasts fairly accurate. To date, audio podcasts remain more common and popular than video podcasts, but that is changing.

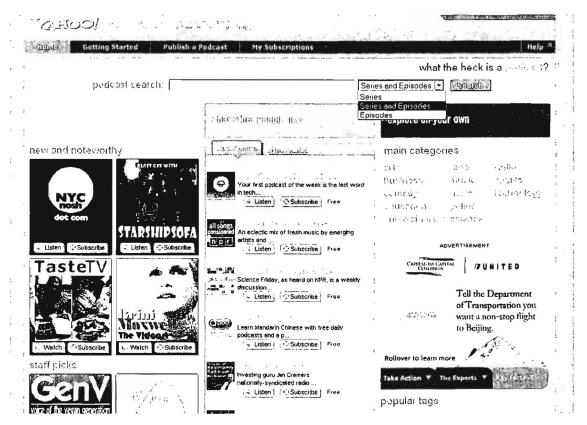
^{125 &}quot;Podcast," Wikipedia, <http://en.wikipedia.org/wiki/Podcast> (19 October 2006).

¹²⁶ While the terminology is changing, I would say the term "video podcast" subsumes <u>vlogs</u>, which are a special type of video podcast, i.e., a weblog containing video.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Yahoo Podcast Search

Yahoo got out in front of the podcasting trend with its new Podcasts Search (Beta) site after a study the search giant published with Ipsos Insight, which disclosed that most of the people who are using do so without even knowing it. Yahoo's audio search option also finds podcasts, but if you are looking specifically for podcasts, the Podcast Search site is better. Yahoo Podcast Search indexes metadata such as keywords, categories, or user-generated topic tags to match queries to podcasts. Users do not need to register at Yahoo to listen to podcasts, but registration is required to subscribe to podcasts. Yahoo Podcast Search includes its own player that launches at the site so users can listen to podcasts.



Yahoo Podcast Search (Beta)

http://podcasts.yahoo.com/

Podzinger

Podzinger takes a completely different approach to indexing podcasts. "Podzinger is a podcast search engine that lets you search the full audio of podcasts just like you search for any other information on the web...Podzinger takes search a step further

¹²⁷ Joshua Grossnickle, et al., Yahoo! and Ipsos Insight, "RSS: Crossing into the Mainstream," [PDF] October 2005, http://publisher.yahoo.com/rss/RSS whitePaper1004.pdf> (24 October 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY

by searching the spoken words inside the podcast in order to find more specific and relevant results. The text-based search results include snippets from the audio to help you figure out if the result is relevant. You can even click on the words to listen to the audio from that point." While this speech recognition technology works well, it only indexes English and Spanish podcasts at this time and the Podzinger index is much smaller than others. Podzinger searches audio, video, or both types of podcasts.

Podzinger also introduced a special tab on its homepage that allows user to "search inside" YouTube videos, that is, to search the spoken text of YouTube videos. A search of YouTube using Podzinger will return the results with the keywords highlighted within the transcribed text at the exact time at which they occurred:



Series: Series

AKBAR GANUES one of the Most Prominent Dissident Journalists in time. He Spent Over 6 years in Prison in time for Publishing a Series of "Investigative Aricles" Regarding the (additional episode information).

Tag:

amy goodman akbar ganji investigative journalist dissident H MI political assassination pacifica iranian government

- 0:00:12 ... arrested after it took part in the conference some political reform in the he was released six years later in march of this year. Since his arrival in the United States has been speaking out against human rights abuses in han he took part in three -- hunger strike outside the UN. Aimed at forcing the Iranian government to release political prisoners last week. He declined to personal invitation to the White House to meet with top US officials overseeing in dipolicy. He rejected the offer he says because he believed that current US policies could not help promote democracy in Iraq -- and she joins us now in the studio. The same can't come come Molfy also joins us to help with translation we welcome you both to democracy now it's good to have you with us. -- -- welcome can you begin by talking about. Your investigations as a journalist in han in the late 1980s -- ...
- 0:08:37 ... Intention is to bring together the anti-war forces within Enron from It are and from the rest of the world together -- -- -- In the US has threatened to attack an frain that is. The president President Bush has revolved around him part of the axis as evil and Patricia. What does this does
- 0:09:20 ... somebody you don't have compounds Younis first it is impossible. And maybe kars in the same manner that Iraq and that kind of stunned hurl and made it. Have to accept quoted him on the ...
- 0:11:31 ... into----- on how Canada pro democracy group finishers in trait and being empowered food should refuctant. Shouldn't be -advancing democracy education I don't believe we have a widespread democratic movement within ...

Requirements for Podzinger are, for Windows Systems

¹²⁸ Podzinger FAQ, Podzinger.com,<<u>http://www.podzinger.com/about.jsp?section=faq</u>> (19 October 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

- Internet Explorer 6.0 or higher and RealPlayer 9.0 or higher
- Internet Explorer 6.0 or higher and Windows Media Player 9.0 or higher
- Firefox 1.0 or higher and Quicktime 6.0 or higher

For Macintosh Systems

Safari 1.0 or higher and Quicktime 6.0 or higher

If your system does not meet these requirements, you can find podcasts and read the text transcript but you cannot play them. Video podcasts can be downloaded and played later or opened and played using a multimedia player such as Windows Media Player or Apple Quicktime. Podzinger supports the following video formats: mp4, mov, m4v, flv, mpg, or mpeg.

Podzinger

http://www.podzinger.com/

Podzinger Spanish Search

http://www.podzinger.com/index.jsp?il=es

Odeo

Odeo has been around since 2004 and now boasts over a million audio files, mostly podcasts from all over the world. It offers both a search and browse option. Users can either download the file or listen to it in a neat little player at the Odeo site. There is an option to create an account so that you can subscribe to feeds and save them at Odeo, but an account is not necessary. The search and categorization scheme is based upon tags, simple keywords or categories.

Odeo

Podcast.net

Podcast.net has more search options than most podcast search engines. Users can search by title & description, keyword, location, host, or episodes. I think the location search is the most useful because you can locate podcasts around the world.

Podcast.net

http://www.podcast.net/

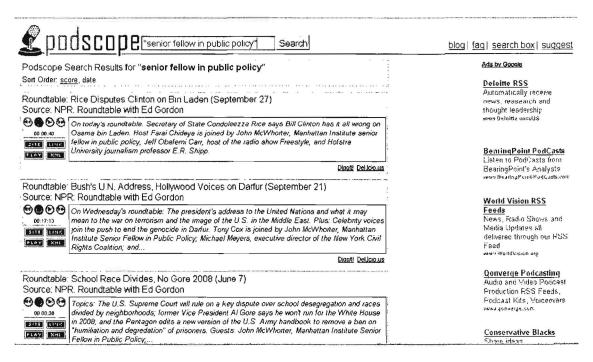
http://odeo.com/

<u>Podscope</u>

Like Podzinger, Podscope transcribes audio from podcasts into searchable text. This means you can search on words and phrases that occur in the podcast; however, I found that Podscope's search is not entirely reliable (the search misses some words that have been transcribed). Nonetheless, it is a valuable podcast search site.

Podcasts play at the site using Flash 8+: simply click on the green play button

UNCLASSIFIED#FOR OFFICIAL USE ONLY



Podscope

http://www.podscope.com/

Other Podcast Search Sites & Directories

A number of sites offer podcast search or function as directories that index podcasts usually by topic.

Podcast Directories

iPodder <u>http://www.ipodder.org/directory/4/podcasts</u>

Podcast Directory http://www.podcastdirectory.com/

Podfeed http://www.podfeed.net/

Podcasting Station http://www.podcasting-station.com/

Podcast Shuffle http://www.podcastshuffle.com/

Other Podcast Search Sites

Blinkx http://www.blinkx.tv/

Podcast Alley http://www.podcastalley.com/

Special Topics—News, Blogs, & Technology Search

Newsgroups, Forums, & Mailing Lists

News is one of the types of information most amenable to the Internet: both are fast moving, worldwide, and in high demand. Along with the proliferation of news sources on the Internet have come an even greater number of ways to share opinions on every subject, limited only by the scope of the human imagination. In this article, I am going to discuss several different ways of accessing both news and opinions on the Internet: newsgroups, message boards or forums, and mailing lists. I discuss weblogs in the next section because they require special software and thus demand a separate discussion.

Newsgroups in general and Usenet in particular have lost a great deal of their prominence in recent years as more outlets for sharing information and communicating on the Internet have appeared. Usenet is a system that allows individuals to post messages and have them read around the world within about three hours. One of the drawbacks of Usenet is that it requires special software (a newsreader) configured to send and receive data via a user's Internet Service Provider's news Microsoft embeds server. its newsreader into software such as Outlook and Outlook Express.

Usenet Warning

This is a place on the Internet I recommend exercising great caution and skepticism because anyone with Internet access can literally post anything to a newsgroup, and often do.

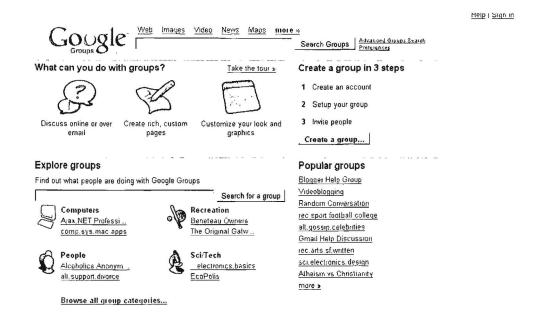
Usenet newsgroups are noted for being chaotic, notoriously unreliable (lots of gossip), and confusing. So why would anyone look at them? Because occasionally there is valuable and unique information posted to newsgroups.

By August 2000, there was only one Usenet search engine remaining: Deja (formerly Dejanews). Fortunately, Google acquired Deja in 2001 and made available the entire Deja archive consisting of over 700 million posts from 1981 to today.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Google Groups indexes more than 1 billion Usenet postings, and new postings normally appear in Google Groups within 10 minutes.

Google Groups put forward a new look in 2004 (out of Beta in early 2007) while retaining all the old postings in its index.



While Google Groups has many options and features that require an account, many of the most valuable features do not:

ACTIVITIES THAT DON'T REQUIRE A GOOGLE ACCOUNT:

- Reading posts in public groups
- · Searching for groups, posts, or authors
- Posting to groups via email if they are unrestricted or you're already a member
- Joining a public Google Group via email

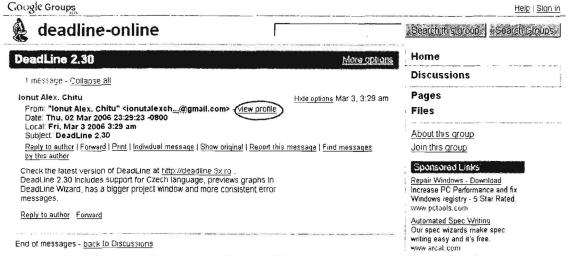
ACTIVITIES THAT DO REQUIRE A GOOGLE ACCOUNT:

- Creating and managing your own Google Group
- Posting to groups via our web interface
- Creating pages and uploading files
- Subscribing to a Usenet newsgroup and receiving posts via email
- · Joining a Google Group via our web interface
- Changing your subscription type (No Email, Abridged Email...)
- · Reading a restricted group's posts online

Google Groups has an **advanced search** that lets you limit your search by language, subject, forum, author, message ID, and/or date. It also has a feature known as "author profile." When you are viewing a posting, click on "More options" (next to the date); then select "View profile" next the author's From address. By clicking on a "Show options" next to an author's name, Google will present the option

UNCLASSIFIED//FOR OFFICIAL USE ONLY

to "Find messages by this author" and automatically search for every posting from a particular address. Google Groups no longer shows the author's entire email address for privacy reasons, but many addresses are guessable.



Google Groups' Option to "Find messages by this author"

Google Groups is also one of the most valuable sources of information on technical and computer-related problems. Someone has probably encountered that same complex and mysterious error message that has you flummoxed. Simply by copying and pasting the error message into Google Groups, you have a very good change of finding out what it means and, even better, how to fix the problem. **Despite what the site says, Google Groups no longer limits the query to 10 words, but to 32** as Google Search does. Also, the Google wildcard (*) does not work in Google Groups.

Google Groups

http://groups.google.com/

Yahoo Groups is also a huge source of newsgroups on virtually every topic. Yahoo Groups is very easy to use because it is arranged like the Yahoo general subject guide, i.e., it is browsable by topic and keyword searchable. Some groups require membership (registration), but many are publicly accessible.

Yahoo Groups does not offer a search by "author's profile," but it does link directly to its **Member Directory**. You may be able to view the member profile, which may provide the user's real name, location, age, occupation, and email address.

Yahoo Member Directory

http://members.yahoo.com/

However, the popularity of newsgroups has declined with the rise of forums, also known as message boards. One of the main advantages of forums as a means of communicating via the Internet is that they do not require software other than a web browser for users to read and post to them. Forums usually are focused on a specific topic, anything from computer games to politics. "In terms of countable

UNCLASSIFIED#FOR OFFICIAL USE ONLY

posts, Japan is far in the lead with over two million posts per day on their largest forum, 2ch. The United States does not have any one large forum, but instead several hundred thousand smaller forums...China, the Netherlands, and France are also home to hundreds of independent forums."¹²⁹

There are several major search engines specifically designed to index and search forums, message boards, and discussion groups. BoardTracker, Lycos Discussions and BoardReader make it possible for users to search many more online discussion communities than ever before.

BoardReader, which was offline for most of 2005, is back and improved. The site has several software partners, including Vivisimo, the clustering metasearch engine behind <u>Clusty</u>, and Thunderstone to search and group information from Internet forums and message boards. There is an option to limit searches by date, site, and language. The sources indexed by BoardReader still tend to be very different from those accessed by Lycos Discussion, Google Groups, or Yahoo Groups, making it a valuable addition to "hidden web" search. BoardReader does recognize non-Latin character sets. The biggest change to BoardReader is the addition of a Domain Profile option, which uses data BoardReader gathers by indexing forums to create a detailed picture of a domain, including number and sources of inbound links, the pages in a domain with the most active inbound links.

BoardReader

http://www.boardreader.com/

BoardTracker searches only for the content of message boards/forums and can be used as a traditional keyword search engine or a browsable directory of message boards by topic. BoardTracker has an "instant alerts system" whereby it "will notify you in a number of ways (email, Jabber, ICQ) as soon as a thread matching your search term is posted on any of the thousands of forums we track." While BroadTracker does not require registration to use, its alert service obviously must require registration, which in turn requires a valid email address. Nonetheless, BoardTracker's registration does not ask for your name or address, so users can employ a free email service such as Hotmail or Yahoo Mail for BoardTracker alerts.

Notice the search options BoardTracker offers, including the search for Threads or by Author, to sort by Relevancy or by Date, and to search in All Categories, your own selected threads, or in a specific category, which often has more subcategories within it, e.g., World & Regional includes specific continents, General, Middle East, and Warfare. I am also happy to report that **BoardTracker searches equally well in non-Latin character sets as it does in English**. I am also intrigued that it will search for results from the "past 6142 years."

¹²⁹ "Internet Forum," Wikipedia, < http://en.wikipedia.org/wiki/Internet_forum> (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

B	DardTracker			
Home	Search My Alerts Browse My Threads	Setup About Login Signup		
Search: Crypto	order by: Relevancy	In: All Categories All Categories "My Threads" Arts & Humanities Search st 6142 years [Safe Filter Is ON]		
Data Transfor Rapid, visual da XML. www.capeclear.co	ata transformation. Replace ETL and EAI with The l	Business & Economy Computers & Internet Culture & Society Beadin Education ons. Entertainment Corporate Cooking Government & People Health & Medicine Home & Family		
Board	Title	Ati News & Media Created at		
VBWire	Missing namespace: System Security Cryptograp	Fri Science & Technology 21:10 9-Apr-05		
VB Forums SignedXml class is supposed to reside in the system.security.cryptography.xml namespace, but I haven't got that namespace available. I have the system.security.cryptography namespacemore				
	Cryptography Research ideas	Jason Isom		
1	Can anyone suggest a good topic to research on Cryl a paper published by a popular website or journal suc	otography? In my information Security course, we have to get the as http://www.ebcvmore		
6.7	Research paper/Cryptography	JaWiB 😝 9 133 06:05 5-Sep-05		
Cprogramming	thinking about researching computer security-specifically cryptography . I don't really know much about cryptography , so I'm wondering how good an idea this would be for a research p <u>more</u>			

BoardTracker

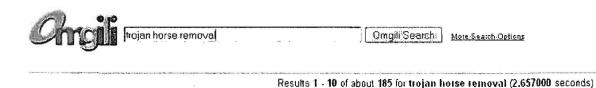
http://www.boardtracker.com/

OMGILI is a new vertical search engine that only searches the content of web forums. Why would anyone want to limit a search to discussion forums? "The information contained in online forums is typically presented in a 'question and answer' or debate style format. How is this significant? Many times you will have a question that has already been answered. Using Omgili, you can avoid posting already asked questions and quickly find your answer...Unlike ordinary search engines that prioritize articles and edited web pages, Omgili only indexes discussion forums. Using Omgili's advanced search capabilities you can choose to independently search titles, topics or just the replies of a discussion."

Think how many times you have needed to find the answer to a question—for example, about trojan horse removal—and had to struggle through typical search engine results that were littered with useless advertisements. Now look at the results for that query using Omgili:

¹³⁰ "About Omgili," Omgili.com, <<u>http://www.omgili.com/about.html</u>> (31 October 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY



trojan horse removal - tomcoyote forums 🥯

help to remove a trojan horse-virus? i have recently acquired the trojan horse virus &aud ...

exe c:\program files\trojanhunter 4.0\thguard.exe c:\program files\scanbutton 2.1\scanbutt...

http://faisms.tom.coyata.org/index.php?showtopic=23493 • Fage with topic and at least 11 replies.

trojan horse removal - swi forums 🗢

effect that the following trojan has been detected: "trojan house dropper.genericzh&..., / panda software anti-trojan: a free / personal / trojanhunter (30-day trial) / onlin... http://forums.spywareinfo.com/index.php?showtopic=56746 - Paga with topic and at least 2 repiles.

cexx org message boards view topic - trojan horse fig removal?
me that i had "trojan horse dialer 6g (maybe g6) and that it was in my temp int fil...
could you download hijackthis, extract it from the zip file into it's own directory like c...
http://boards.cexx.org/siewtopic.php?t=5044 - Page with topic and at least 1 replies.

trojan horse downloader agent problem - pc pitstop forums 🤝

shield picked up a trojan, according to avg it's a trojan house downloader agent q. every...

trojan.win32.agent.cs aka adware-virtumundo this may not be a simple repair.....syman...

http://popilstop.invisionzone.com/index.php?showtopic=105712 - Page with topic and at official U.S. Web site for all Norton and

Sponsored Links

Bed and Breakfast.com

The Trojan Horse Inn on Bed and Breakfast.com. Find information on over 27,000 properties on Bed and Breakfast.com. Guest reviews, maps, photos, gift-centificates and more.

Find Trojan Horse Virus with Business.com

Computers & Software - Business.com will help you find just what looking for. Browse through more than 65,000 categories. Find all the vendors and goods you need to grow your business.

Links To Antivirus Software

Detect viruses, Trojans, worms, scan and clean incoming and outgoing email Detect viruses in instant-message attachments.

Norton: Trojan Horse Virus Removal

2006 version available now. Stay protected with the trusted antivirus software at the official U.S. Web site for all Norton and

The Omgili crawler analyzes a forum differently from a webpage, recognizing and assigning different weights to the topic, title, and replies. One thing to keep in mind is that you will have to register at forums where registration is required, i.e., Omgili does not offer a cached option. I would recommend adding Omgili to your forum search bookmarks. Oh, in case you're interested, Omgili stands for "oh my God I love it!"

Omgili

http://www.omgili.com/

Finding News Groups and Mailing Lists

In contrast to searching newsgroup *content*, **Tile.net** will help you find newsgroups or mailing lists of interest. Tile.net offers an alphabetical listing of all Usenet newsgroups by description or newsgroup hierarchy. Don't worry if you don't understand the hierarchy because Tile.net has a search engine that will find the appropriate newsgroups to match your keywords. Tile.net also provides a searchable listing of mailing lists and discussion groups.

CataList is the official source to "browse any of the 58,638 public LISTSERV lists on the Internet, search for mailing lists of interest, and get information about LISTSERV host sites." LISTSERV is one of the most useful and now venerable Internet

UNCLASSIFIED//FOR OFFICIAL USE ONLY

programs (dating from 1986), scanning email messages for the words "subscribe" and "unsubscribe" to automatically update mailing list subscriptions. At the CataList site, users can get the following public List and Site information:

List information

- Search for a mailing list of interest
- View lists by host country
- View lists with 10,000 subscribers or more
- View lists with 1,000 subscribers or more

Site information

- Search for a LISTSERV site of interest
- View sites by country

Finally, both Google and Yahoo Groups allow users to search or browse groups by topic.

Google Groups

Yahoo Groups

http://groups.google.com/

Tile.net http://www.tile.net/

CataList http://www.lsoft.com/lists/listref.html

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Weblogs & RSS Feeds

Weblogs (more often known simply as **blogs**) are web pages that are usually defined as online journals or diaries but can be many things. "Blog posts are like instant messages to the web. Many blogs are personal, 'what's on my mind' type musings. Others are collaborative efforts based on a specific topic or area of mutual interest. Some blogs are for play. Some are for work. Some are both." Blogs can be wonderful sources of news and information, or they can be absolute schlock. I find blogs most useful as sources of rumors and opinions on particular topics or breaking news stories. Making blogs useful is tricky, but some of the best ways to access them are via specialty blog search tools because many search engines do not offer blog-specific searches. The first major blog search tool was <u>Daypop</u>, an excellent news search engine that lets users limit a search to blogs or RSS feeds. Google has a separate blog search engine, Yahoo subsumes blog searches under its news search, and Live has a "feed" search.

Many—in fact probably most—blogs make their content available in **RSS**. What is RSS? RSS stands either for Rich Site Summary or RDF Site Summary (there is some dispute about this). "RSS is an XML format for syndicating web content. A website that wants to allow other sites to publish some of its content creates an RSS document and registers the document with an RSS publisher. A user that can read RSS-distributed content can use the content on a different site. Syndicated content includes such data as news feeds, events listings…excerpts from discussion forums or even corporate information." Look for the RSS/Atom feed icon at websites to subscribe to a feed.

To add to the confusion, RSS is not the only format used for blogging and newsfeeds. The other major format is Atom. Here's a good explanation of the difference:

RSS/XML/Atom are technologies, but syndication is a process. RSS and Atom are two flavors of what is more or less the same thing: a 'feed' which is a wrapper for pieces of regularly and sequentially-updated content, be they news articles, weblog posts, a series of photographs, and more. For the purposes of this article, *consider the terms interchangeable*. XML is the base technology both are built on, but that's almost totally irrelevant; the

¹³¹ "About Blogger," Blogger.com, < http://www.blogger.com/about.pyra (14 November 2006).

¹³² "RSS." Webopedia, http://www.webopedia.com/TERM/R/RSS.html (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

orange buttons are mislabeled, and should read 'RSS' or 'Atom' instead. Strange, but true. 133

News aggregators (sometimes referred to as RSS aggregators) are programs designed to read XML formatted content, which is very popular in the blogging community. News aggregators retrieve RSS/Atom feeds and present these feeds in an easy to read format. Bloggers and many news websites use RSS/Atom feeds in XML format to publish information. Most news aggregators are downloadable programs that need to be installed on your computer, but some are implemented on websites.

There are now a number of RSS and blog search engines. Unfortunately, the relationship between RSS feeds and blogs is not as clean and clear as one might hope. While RSS search engines get their content from RSS feeds and not from crawling the web, any type of website can distribute content using RSS. This means RSS search engines are searching more than just blogs. On the flip side, not every blog uses RSS to distribute its content, so some blogs are not searched by RSS search engines.

So what exactly do blog search engines provide that traditional news search engines do not? What users often get from blogs are biased insights and opinions. For good or for ill, blogs are somewhere between newsgroups/chat rooms and true journalistic sites. Why use blog search engines? First, remember that traditional search engines are the least useful for news or date-sensitive information. News search tools are best for timely objective reporting. What blog search adds is diversified opinion (sometimes useful, sometimes not) on virtually every topic imaginable. Also, even good news search sites may index a limited number of sources whereas a good blog site may get news out faster and more efficiently. However, for a number of reasons, blog search engines' algorithms are not as good as general search engines' algorithms at weeding out spam, so you will probably have to wade through a lot of inappropriate and useless sites when using a blog search tool.

The list of blog search engines seems to be growing weekly, but I expect some will fall by the wayside as others become more popular. I recommend trying several of these tools because you will probably get very different results. In late 2006 Google Blogsearch overtook Technorati as the most popular blog search site. Technorati offers a somewhat different approach. The concept behind Technorati is that it "watches" over 63 million weblogs, analyzing who is linking to a blog, website, or news article. By entering the url of any blog, website, or news article, users can see how many blogs link to it, which bloggers are linking to that page, and what they are saying about it. In essence, Technorati is a very simple website analysis tool that

^{133 &}quot;What is RSS/XML/Atom Syndication?" Mezzoblue.com,

http://www.mezzoblue.com/archives/2004/05/19/what_is_rssx/ (14 November 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY

lets users see what others are saying about almost any site of interest. Technorati also offers a keyword search of all the weblogs it tracks.

Technorati and the Associated Press "initiated a service to connect bloggers to more than 440 AP member newspapers nationwide...Increasingly, what the blogosphere says about a news story becomes part of a more complete story, lending diverse perspectives and often expert commentary. The new service will bring blogger commentary about AP news stories to communities large and small throughout the USA, giving bloggers a voice in trusted local papers throughout the nation...When readers visit an AP member Web site that uses AP Hosted Custom News, they will see a module featuring the 'Top Five Most Blogged About' AP articles right next to the article text, dynamically powered by Technorati. Additionally, when readers click on an AP article, Technorati will deliver 'Who's Blogging About' that article."

At the participating sites you will see the "Blog Roundup" from Technorati:

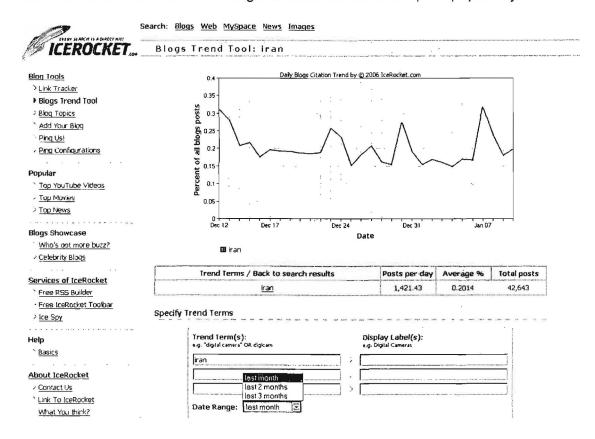
	Defense: Duke Accuser Changes	Blood Books diese
Politics	Story	Blog Roundup The most blogged
Washington	DURHAM, N.C. (AP) The accuser in	Associated Press articles.
Offbeat	the Duke lacrosse sexual assault case	
Podcasts	told prosecutors in December that one of	Congress Divided Over Bush War Plan
Blogs	the three players charged did not commit	25 new links
Weather	any sex act on her during the alleged	U.SLed Forces Detain 6
Raw Hews News SEARCH	attack, according to papers filed Thursday by the defense	Iranian Workers 18 new links
Search Go	Snow Latest Weather Woes for Seattle SEATTLE (AP) A cold snap swept	al-Qaida Suspects Still Alive in Somalia 11 new links
Archive Search SPECIAL SECTIONS Multimedia Gallery	through Seattle and surrounding areas	Cisco Sues Apple Over Use of iPhone Name 10 new links
Hews Summary (AUDIO) AP Video Network	more than 350,000 students, snarling traffic and causing at least one traffic-related death	Nicaraguan Revolutionary Back in Power 7 new links
Today in History	4 Hurt in Ind. Workplace Shooting	POWERED BY 🎥 Technorati
Photo Gallery	INDIANAPOLIS (AP) A man shot and	
PhotoWeek	wounded four co-workers Thursday	
SportsWeek	morning at a manufacturing business that employs disabled people through Easter	
U.S. Census Database	Seals, telling police he shot them over respect, notice said	

¹³⁴ Peter Hirshberg, "Technorati Teams With The Associated Press to Connect Bloggers to More Than 440 Newspapers Nationwide," Technorati Weblog, 23 May 2006, http://technorati.com/weblog/2006/05/107.html (31 October 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Clicking on the headline links you to the article; clicking on the "Links to this article" takes you to Technorati to read blog entries about the story. I think this development is further evidence that the definition of "journalism" is fuzzier than ever.

The Diarist Registry indexes more than 7100 blog sites, is fully searchable, and lists blogs by country. Bloogz permits searches in four European languages plus English and Blogwise offers an option to list blogs by country. IceRocket has a "trend tool" that shows keyword trends for the last one to three months as well as other information to allow users to get a broad sense of a topic's popularity.



Also, don't forget <u>Waypath</u>, which lets you see blogs that link to specific websites almost in real time. Waypath also searches the full text of blogs as well as by keyword.

Of course, we now have Google <u>Blogsearch</u>, which offers many search features and options but contains only fairly recent content. Also, keep in mind that **Google Blog Search only indexes the site feed, not the full content at the website** that originated the feed.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

All the blog search engines permit **sort by date** in addition to sort by relevance. Date sorting works perfectly well with weblogs, which (unlike webpages) have date/time tags.

http://www.blogdigger.com/ Blogdigger Blog Search Engine http://www.blogsearchengine.com/ Blogwise http://www.blogwise.com/ Bloogz http://www.bloogz.com/ Clusty Blog Metasearch http://blogs.clusty.com/ Daypop http://www.daypop.com/ Feedster http://www.feedster.com/ Google Blogsearch http://blogsearch.google.com/ **IceRocket** http://blogs.icerocket.com/ http://www.sphere.com/ Sphere The Diarist Registry http://www.diarist.net/registry/ Technorati http://www.technorati.com/

There are as many potential blogs as there are people willing to create and maintain them. Quite honestly, most are as dull as ditchwater. However, there is a type of blog—or perhaps it's more accurate to say, a type of blogger—that is inherently interesting: the insider. For example, there are several hundred Microsoft employees who are blogging, and a lot of what they are talking about is directly related to Microsoft products that are still in the planning stages. "Prolific Microsoft blogger Robert Scoble mentioned Windows XP Tablet PC Edition 2005 (and complained that the real name was far less interesting than the codename Lone Star)." A few days after his blog entry appeared, Microsoft officially announced that the service pack Tablet PC is tied to wouldn't be out as planned by June (the last month of the Microsoft fiscal year) but was being delayed until at the earliest July 2005, when Microsoft's 2005 fiscal year begins. Other Microsoft bloggers have discussed such things as technical details of work with the latest developer tools, Longhorn (which became Vista, the current Windows OS for desktop PCs) architecture, and how features are developed for Microsoft products like Word.

Microsoft has definitely taken a "if you can't beat 'em, join 'em" attitude towards blogs (so far). In 2004 the company went so far as to launch Channel 9, a website that combines blogs, discussion forums and other technology to improve communications with developers. Channel 9 refers to the audio channel many

^{135 &}quot;Inside Track," The Guardian, 20 May 2004,

http://technology.guardian.co.uk/online/story/0,3605,1220085,00.html (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

airlines use to let passengers listen in on crew conversations, and, according to the welcome message at the site, "We think developers need their own channel 9, a way to listen in to the cockpit at Microsoft, an opportunity to learn how we fly, a chance to get to know our pilots...Five of us in Redmond are crazy enough to think we just might learn something from getting to know each other...Join in, and have a look inside our cockpit and help us fly the plane."

Channel 9

http://channel9.msdn.com/

Companies walk a fine line between encouraging people to share information and protecting their proprietary "crown jewels." In fact, Microsoft fired a long-time temporary employee for posting photographs of Apple G5 Macs being delivered to the Microsoft campus. The employee originally posted the photo on his own site and intended it mainly for friends and family (he even cropped the photo so that the Microsoft campus wouldn't be visible). Nonetheless, the photo was quickly discovered and widely discussed across the web, leading Microsoft to dismiss him for "violating company security policy." Given the wide-open nature of the Internet, the ease of blogging, and the fact that people love to share information for whatever reason, blogging will undoubtedly become a bigger problem for companies and organizations. Employees may innocently or inadvertently post inside information that could be harmful to their organization. And it's not hard to imagine what damage a disgruntled employee could do.

General News Sources

Finding news sources around the globe has become fairly easy. All of the following are very good sources of different types of news sources: newspapers, magazines, wire services, etc. Many have links to online publications and news sources, and many of these publications are available free over the Internet.

ABYZ Newslinks http://www.abyznewslinks.com/ Guardian's World News Guide http://www.guardian.co.uk/worldnewsguide/ HeadlineSpot http://www.headlinespot.com/ http://www.esperanto.se/kiosk/engindex.html Kiosken Metagrid (newspapers & magazines) http://www.metagrid.com/ NewsCentral (online newspaper links) http://www.all-links.com/newscentral/ NewsDirectory http://newsdirectory.com/ Newslink http://newslink.org/ Online Newspapers http://www.onlinenewspapers.com//index.htm RefDesk (My Virtual Newspaper) http://www.refdesk.com/papmain.html

UNCLASSIFIED#FOR OFFICIAL USE ONLY

News Sites & Search Engines

In addition to the sites above that help you locate news sources, there are also a growing number of excellent search services dedicated to delivering and searching the news. In fact, so many people all over the world are relying on the Internet for their news that "old media" sources continue to lose revenue as they compete for readers and viewers. It turns out that some of the "new media" Internet news services such as Google and Yahoo now compensate old media sources such as AP on a pay-per-click basis. "Major Internet portals such as Yahoo and America Online have been paying for content since their creation in the mid-1990s... Earlier this year [2006], Google signed a deal with the Associated Press, one of more than 50 agreements AP Chief Executive Tom Curley has obtained from Internet players" since 2003.¹³⁶

Google News and Yahoo News are among the very best news search tools. Both have superb advanced search options that include the ability to limit searches to specific publications, dates, etc. Both let you sort the results by relevance or by date and time. News searching is one area in which it makes sense to sort results by date because news stories, unlike webpages, have meaningful date/time "stamps" or tags.

Because of partnerships between many news websites and news providers, readers often have access to subscription news stories without having to register (e.g., *The New York Times*) or give personal information at a news website. Occasionally, you may even discover you can access an article requiring payment because of a partnership arrangement. For example, this article is only available by paid subscription at its website, but by using Google News to find it, a reader can access the full article for free and without registering.

The Patent Office: Getting Wiki With It Law.com (subscription), CA - Jan 12, 2007 In August, when the Patent and Trademark Office acknowledged that it had taken Wikipedia off its list of acceptable research sources, the surprise was not ...

¹³⁶Elise Ackerman, "New media making deals with `old' news providers," *Mercury News*, 30 July 2006, http://www.mercurynews.com/mld/mercurynews/news/15157800.htm (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY



However, finding and understanding how to use these services can be confusing because most sites that permit unimpeded access to subscription news do not do a good job of advertising how to do it.

Some of the major news sources that require registration or personal information to access some or all of their news, but which are accessible via news search sites include:

The New York Times
The Washington Post
The LA Times
BusinessWeek
Forbes
The Guardian (UK)
The Times of London

I have included some ways to access subscription news without providing personal information in the discussion of news search engines below.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Google News

http://news.google.com/

Google News headlines are entirely generated using a computer algorithm that scours over 10,000 worldwide news sources. Google News also offers international editions for France, Germany, India, Italy, Spain, and several other countries.

Google News lets users:

- Sort results by date or relevance.
- Search on a title or headline using the syntax intitle: [intitle:iraq].
- > Use both the *intitle*: and *inurl*: or *source*: commands together to find a *specific topic in a specific publication*.
- Limit a search to a **specific U.S.** state or a **country** [location:ny] or [location:germany].
- Limit a search to a **specific news site** using the syntax *inurl*: or *source*: [inurl:washingtonpost.com].

Google also permits users to limit their search by news source in both simple and advanced news search. The **source**: command must be used with a keyword. The syntax in simple news search is **source**: and the query is:

[source:news_source_name keyword]

For example: [source:new york times iraq]

The **source**: syntax can be tricky to use. Here's a partial list I've come up with which works (as of now) in simple search:

source:bbc_news source:international_herald_tribune source:united_press_international source:guardian source:christian_science_monitor source:cbs_news source:abc_news

¹³⁷ Philipp Lenssen's research during 2006 indicated at least 10.584 unique sources for Google News. "Which sources does Google News index?" Google Blogoscoped, http://blog.outer-court.com/googlenews/> (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

source:washington_post source:reuters source:cnn source:msnbc source:forbes

Another option is to use the **Advanced Google News Search** that includes these options:

- Find results—all words, exact phrase, at least one word, without the word(s).
- Sort by relevance or date.
- Location by country or state.
- Occurrences of terms—anywhere or in headline, body, or url of article.
- ➤ Date—anytime, last hour, last day, past week, past month, or range of dates within past month.
- Source by simple name, e.g., CNN, New York Times.

The *source* option in advanced search is an improvement over having to use the *source*: syntax that required such entries as [source:united_press_international]. However, there are still two problems with the *source*: option. Google doesn't list its sources anywhere that I can find (it merely says there are about 4500 of them), and you still have to write out the full name of the source, so UPI or Fox will not work but United Press International or Fox News will. Google News now offers a Google Blogsearch link from the News homepage.

Google News Advanced Search http://news.google.com/advanced news search

Google News also offers **customization features** that let you add or remove news categories, decrease or increase the number of headlines you see, select from numerous regional editions of Google News from around the world, and more. For example, you could customize your Google News page to show only the World, Sci/Tech, Business, and U.S. pages, then add standard sections for German Business and another for France in general. Google News also permits adding a custom section in which you enter keywords (all must be in the story for it to appear in your section, so choose carefully). You can even use the "Advanced" option under custom section to limit your keyword search to a specific topic, e.g., "internet security" only in the Sci/Tech section. You can change the language of the stories in your custom section by clicking on "Advanced" and selecting a language in the pulldown menu. While these new customization options may sound a little complicated, they are fairly intuitive, and Google News is very forgiving. If you make a mess of your customization, there is an option to "Reset page to default" link to return your page to its default settings (i.e., the standard version without any

UNCLASSIFIED#FOR OFFICIAL USE ONLY

customizations). This will, however, obliterate any customization you ever made, so you might want to delete unwanted sections one at a time instead.

Google News now permits users to get any Google News page, section, search, or customized page via RSS or Atom feeds. Google News provides detailed information on how to set up and access these feeds.

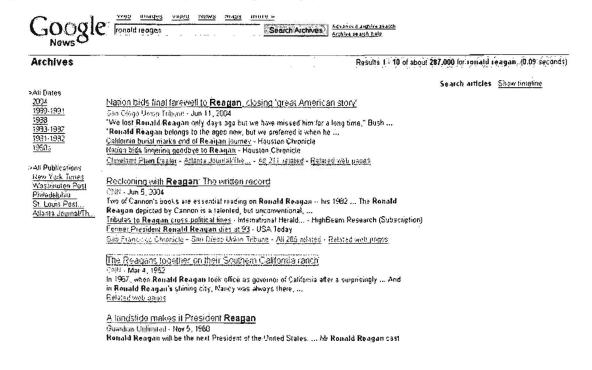
Google News Feeds http://news.google.com/intl/en_us/news_feed_terms.html

Google News keeps news stories in its database for about 30 days. This might be a bit misleading since much of the freely available news on the Internet is only available for less than 30 days (usually somewhere between 1 and 2 weeks). Unlike the Google search engine, Google News does not offer cached (stored) copies of these news stories. However, some news articles that cannot be accessed via the original site's archives because they require registration were indeed accessible using Google. So if you cannot get to a news article that is less than 30 days old via the original source, try searching for it in Google News and linking directly to the page.

Google News Archive

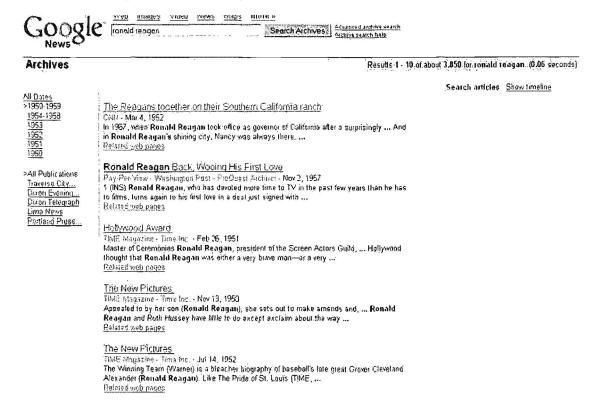
http://news.google.com/archivesearch

Google News unveiled a new service that, from the looks of it, will be very popular and useful—and potentially highly addictive. It's Google News Archive Search, which links from the Google News homepage. Here are the results for the search [ronald reagan]. On the right are the results listed by relevance, though they are a bit odd, considering that the third result is from CNN but dated 1952 (the article actually dates from Reagan's death in 2004 and is about his life in the 1950's). On the left you can limit the search by date range or by publication.



UNCLASSIFIED#FOR OFFICIAL USE ONLY

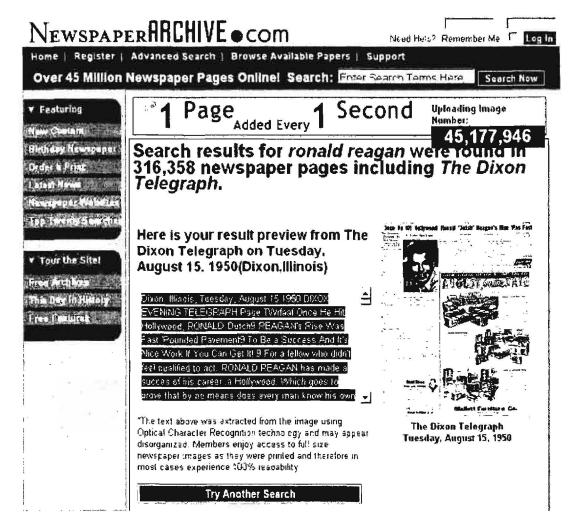
This is where the Google News Archive search is irresistible to researchers (not to mention to history and news buffs). Take a look at the results from the 1950s:



Ignore the first result, which we know isn't really from 1952. Most of the oldest news appears to come either from Time (if it's free) or from the Newspaper Archive. The Newspaper Archive is "the single largest historical newspaper database online, containing more newspaper pages from 1759 to present than any other service." It is subscription based, but free to public libraries and K-12 schools, so if you need a copy of the full sized newspaper image, you can get it at your public library.

The good news is that Google News Archive searches the Newspaper Archive and returns the results in a window. For example:

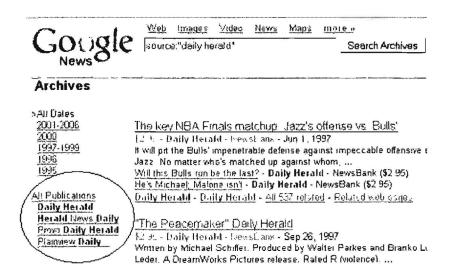
UNCLASSIFIED#FOR OFFICIAL USE ONLY



Although you cannot get the full size image using Google New Archive search, the complete article is here. You can scroll down, read, and even copy the entire newspaper article from the August 15, 1950, Dixon Telegraph using Google News Archive search, whereas even viewing results at the Newspaper Archive is unavailable to users unless they are paid members of the Newspaper Archive. You can limit your Google News Archive searches to the Newspaper Archive by adding [source:"newspaperarchive"]. Many of the results in the Google News Archive search come from subscription services, and these are clearly marked as such. The advanced search contains an option to "Return articles with the following price," including articles with "no price," meaning "free," not "priceless."

You can also limit your search to a specific news source, e.g., [source:"daily herald"], and Google News Archive will show the possible publications if more than one fits that query:

UNCLASSIFIED#FOR OFFICIAL USE ONLY



You may have noticed the option to "view timeline" above. When you select this, Google News Archive puts all the search results into a date ordered timeline that may make complex historical events and news stories easier to research:



Engineerities - this lungton buttle Engineers architer - Apr 20, 1976. *Unider today's conditions, Ronald Reagan would sweep delegates from the 13 southern and border states against Gerald R. Ford.* ...

<u>Uncomputional Windom</u> - TIME Magazina - Time Inc.

2 16 - stap - Earlie Leeb pages

1976 ** TME Magez he Arctive Article - How Ford Won and Reagan Lost ...

THE THINK A TO MAR 8, 1976
At Rohald Reagan's headquarters in Concords dingly New Hampshite Highway Hotel, confident aides had chilled several bottles of Almaden bland de blands ...

TIME Magazine Archive Article - Never underestima - TIME Magazine - Time Inc. A A 1976 P. Date with page.

The Rousing Numbers Game

UNCLASSIFIED#FOR OFFICIAL USE ONLY

The oldest newspaper articles I saw are from the Public Register or Freeman's Journal (dating back to 1763), Edinburgh Advertiser (to 1772), the Times of London (to 1788), and the Daily Universal Register of London (to 1786). You can read articles in British newspapers about the upstart American colonials and follow the war news up to a report on November 30, 1781, in the Edinburgh Advertiser that news "arrived off Cape Charles [Virginia] on the 24th [of October], when we had the mortification to hear that Lord Cornwallis had proposed terms of capitulation to the enemy on the 19th. This intelligence was brought us by the pilot of the Charon, and some other persons Who came off from the shore." If that doesn't make you salivate, nothing will.

Yahoo News

http://news.yahoo.com/

Yahoo Advanced News Search

http://news.search.yahoo.com/news/advanced

Yahoo News is one of the most impressive news sites on the Internet. In fact, it is my favorite news search site. Yahoo News crawls over **7,000** news sources in **35** languages. In each topical section, Yahoo News offers the option of viewing all articles from a specific source and access to the full-text articles. The news sources include major news sources that require registration or personal information to access some or all of their news, but which are accessible via Yahoo News, including but not limited to:

The New York Times
The Washington Post
The Los Angeles Times
Baltimore Sun
Chicago Tribune
Agence France-Presse (AFP)
The Guardian (UK)
Canadian Press (CP)
BusinessWeek
Forbes
Motley Fool
The Deal

News from the wire services, Reuters and the Associated Press, is stored for two weeks. News from other sources is usually stored for one month, although for some providers it could be as little as seven days.

Yahoo News search now includes results from <u>blogs</u> as well as from traditional news sources. When you search using Yahoo News you get results from <u>blogs</u>, the <u>Flickr</u> photograph sharing site, and links from <u>My Web 2.0 Beta</u>, a Yahoo "social search"

UNCLASSIFIED#FOR OFFICIAL USE ONLY

property that encourages grassroots "reporting." Traditional news source results appear on the left side of the screen and the new blog results on the right.



Yahoo Advanced News Search offers many options, including date, location, language, and source limits. To limit your search to a specific source using Advanced News Search, scroll down to Source and enter the source, e.g., [Los Angeles Times]. Users can also search by source" or site in the simple search screen as follows:

Limit a search to a specific news source using the syntax url: inurl:, or source:
[Index times as an an an arrange of the syntax url: inurl: times a search or force or an arrange of the syntax url: times a search or force or an arrange or a search or force or an arrange or a search or force o

[url:nytimes.com] or [inurl:nytimes.com] or [source:new_york_times]

Yahoo News also lets users:

- Sort results by date or relevance.
- Search on headline: or title word to see exact title (headline) matches [headline:iraq].

UNCLASSIFIED#FOR OFFICIAL USE ONLY

- Search on headline: before each of multiple keywords to search for titles that must contain all selected words. [headline:mideast headline:peace]
- Use parentheses to search for headlines that contain at least one of your selected words.
 [(headline:mideast headline:peace)]
- Enter phrases to search for titles that contain a specific terms. [headline:"mideast peace"]
- Limit a search to a specific U.S. state or a country: [location:ny] or [location:germany]

These queries can be used in combination as well as alone or with keywords.

Pandia Newsfinder

http://www.pandia.com/news/

Pandia's newsfeed is powered by **Moreover**, which means the number of sources queried is huge and includes the usual—*The New York Times*, AP, Reuters, BBC, the *Washington Post*, ABC, MSNBC—as well as the uncommon—*Financial Times*, Stratfor, Wikinews, and more. Use the *Pandia News Extended Search* on the pull-down menu for a greater number of news sources. Pandia Newsfinder is like everything else at this excellent Internet search site: easy to use, well organized, and high quality.

In addition to news search engines, there are some excellent sites that offer one interface to many news search sources. These are some of the best.

HavenWorks

http://havenworks.com/news/search/

Don't give up on this resource just because HavenWorks has one of the ugliest, most cluttered webpages you'll ever see. HavenWorks lets you use some of the best news search engines from one page: AltaVista and AlltheWeb News (both use Yahoo News sources), Google News, Daypop (news and blog search), and Rocketinfo (only five days worth). Unfortunately, Moreover no longer permits individual users to search from its home page, so the HavenWorks' Moreover search is no longer operational. Havenworks also has some excellent news search tips and specialized news searches if you can find your way through the maze of clashing colors and general untidiness.

JournalismNet

http://www.journalismnet.com/

As its name implies, JournalismNet is designed primarily as a tool for journalists.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

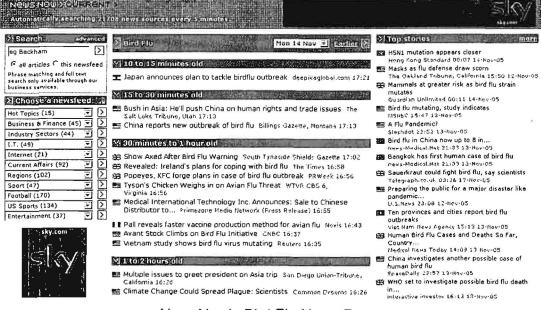
Journalism Net's international newsfeed is powered by Moreover, but it's not designed for searching. Jnet provides links by country and region—North America, Europe, and Africa—to the respective newspapers, news archives, television, radio, etc. There are also links to search tools, people finders, fact finders, newspapers, magazines, TV, radio, etc. The site is very heavily oriented towards tutorials—lots of "how to" pages, which are very useful.

NewsNow

http://www.newsnow.co.uk/

NewsNow offers a newsfeed that automatically updates itself every five minutes from news sources around the world. NewsNow is a UK resource that provides constant news monitoring of nearly 22000 worldwide sources on a large variety of topics. From the home page users can select a specific newsfeed from top-level subjects that include hot topics, business & finance, industry sectors, IT, Internet, current affairs, and regions. The newsfeeds are listed by "new in the last 5 minutes," "10 to 15 minutes old," "15 to 30 minutes old," and earlier articles. Although earlier articles from the same or previous days are available, NewsNow is primarily a current news source, not an archive. NewsNow also offers a free simple search that is quite good.

All sources are clearly indicated both by the name of the media outlet and by a national flag of the source's home country. Almost all the articles are in English, though I have seen a few non-English language articles on occasion. NewsNow is a great resource and is one of the best "breaking news" sources because of the automatic updates, the number and quality of news sources, the fact it does not require registration/payment, and because it has no ads. NewsNow will not display properly in older versions of Netscape.



NewsNow's Bird Flu News Page

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Topix.net

http://www.topix.net/

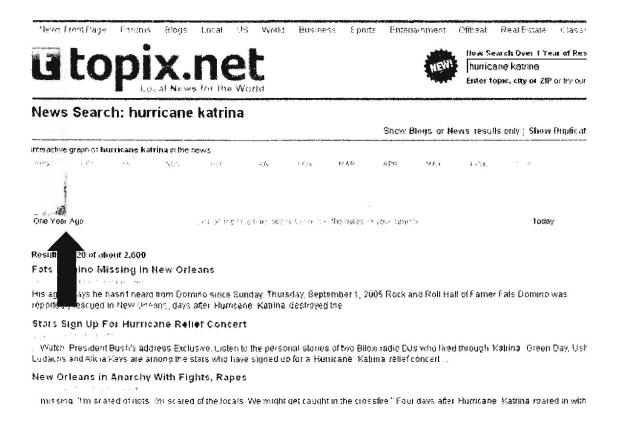
Topix gained a lot of attention and new users during the past year. Here are some of its features:

- It indexes more than 10,000 news sources (mostly US, but not all).
- ➤ It has more than 360,000 topically based pages (including world news pages on virtually every country).
- > It offers a "live" newsfeed of the news stories Topix is indexing at the moment.
- > It has a large searchable database.
- > It offers perhaps the best local US news aggregation and search available (I've not seen better).

Topix tweaked its news ranking algorithm during 2006, producing much better results for top stories than in the past. Topix also added two new features during 2006 that have elevated it at least in one regard. You are supposed to be able to search using capitalization, which has sadly become a rare search feature. The only problem is that Topix's new feature is not working properly or consistently. I tried a number of variations and the results were inconsistent. For example, a search on [ICon] contrasted with [icon] or [Icon] worked, but searches on [ZIP] [zip] [Zip] were unpredictable. This feature needs more work.

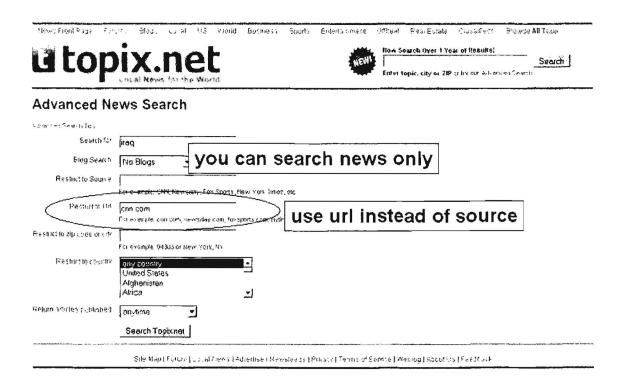
However, the **archive function** works just fine. Here is a search for Hurricane Katrina, and you can clearly see by the blue graph the peaks and valleys of news stories on Katrina. In order to see the news from earlier times, users simply need to click anywhere along the timeline:

UNCLASSIFIED#FOR OFFICIAL USE ONLY



The archive feature is very significant because finding and accessing news older than the past 30 days is usually both hard and/or costly (until <u>Google News Archive Search</u> showed up). Most news sources require at least registration and often payment to get their archived news. Topix will find news from up to one year ago, but what you can actually read for free and without registration is hit or miss. For example, I had very good luck with *Washington Post* stories, but not so much with the *New York Times*. CNN was perfect: I was able to find and read all the CNN stories I tried from last year, something you cannot do at CNN itself. However, I found that the "Restrict to Source" [cnn] option did not work well whereas the "Restrict to Url" [cnn.com] option worked perfectly. Also, I recommend selecting "No Blogs" if you only want to search news sources.

UNCLASSIFIED#FOR OFFICIAL USE ONLY



Despite the fact that the capitalization search is not (yet) working properly, the fact that the archive search is working makes these Topix upgrades important and valuable to researchers.

In short, Topix is a good entry into the news aggregation, archive search, and email alerts arena. For those who can use email news alerts, Topix makes it very easy. Just click the "email alert" link on any topical page to get related results delivered either daily or weekly. In case you're interested, Google News has an "as it happens" option for news delivery http://www.google.com/alerts>, something Topic needs to add.

Daypop

http://www.daypop.com/

Don't forget that Daypop is also a news search site and draws from a large number of international news sources. Daypop's Advanced Search lets you limit your search by time (from three hours to two weeks), by language (all or one of a dozen), and by country. Daypop is one of the few news search sites that caches news stories, so you have the option of viewing the stored copy.

Worldnews http://www.wn.com/

Worldnews is one of the best international news sources, offering the option of searching in English or one of 20+ other languages. Worldnews also presents separate home pages for many world regions, e.g., "Arab World News" as well as specialty news searches by topics such as business, politics, and science.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Technology News Sources

There are many fine telecommunication news sources on the web, so I will limit my discussion to a few of the best. *Users can now subscribe to the all of the following publications' technology feeds and blogs* without registering at the various websites. Look for the RSS/Atom feed icon at the news websites.

TechNews.com

http://www.washingtonpost.com/wp-dyn/technology/

The Washington Post absorbed Newsbytes to create TechNews, which offers current headlines, technology news by topic (biotech, telecom, software, Internet, etc.), and breaking technology news off the wire from Reuters. The site now requires free registration, but both Yahoo News and Google News let you read full-text TechNews articles without registering. For both Yahoo News and Google News, the search syntax is:

[source:washington_post keyword]

Because of the way the *Post* names its links, I have not been able to come up with a unique keyword to limit the query to TechNews articles only. The best way to find them is to use the author's name or a word from the title.

TechWeb

http://www.techweb.com/

TechWeb is part of the CMP Technology Network, a huge enterprise that publishes InformationWeek, *EE Times*, and *Windows Developer* magazines. TechWeb does not do original reporting; instead, it compiles the latest technology news from around the world and updates the website continuously. TechWeb is one of the single best sources of technology news on the web; it and all other parts of the CMP network are fully searchable.

Wired News

http://www.wired.com/

Wired News, part of the Terra Lycos family, has always been known as a slightly "edgy" site, and there is no doubt their original tech news features are among the most popular and sometimes controversial on the web. Wired doesn't only cover technology but brings in all aspects of our "wired" world: culture, business, and politics as well as technology. Wired also receives breaking news from AP and Reuters.

New York Times Technology News

http://www.nytimes.com/pages/technology/index.html

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Consistently a leader in technology news, *The New York Times*' feature writers include John Markoff and David Pogue. The Times also publishes the best tech stories from the wire services. I consider the Times technology section a "must read" to stay abreast of significant developments and to find in-depth articles on topics others deal with superficially.

The Register

http://theregister.co.uk/

If you like irascible journalism, check out The Register from the UK. Their slogan—"biting the hand that feeds IT"—says a great deal about the tone of the site. However, *The Register* is known for gutsy original reporting and is unafraid to take on controversial topics in a straightforward way. They are often cited and quoted by other sources. One of my favorites.

ZDNet News

http://news.zdnet.com/

One of *Wired's* major competitors is ZDNet News, part of the vast **CNET** network. In addition to original reporting and features, ZDNet News offers news by section: wired and wireless, security, IT management web technology, personal tech, etc.

Newsfactor Network

http://www.newsfactor.com/

The Newsfactor Network does both original reporting and compilations of technology news from many sources. Newsfactor has many special sections, such as a superb Enterprise Security section, data management, tech trends, Internet life, among others.

UNCLASSIFIED HOR OFFICIAL USE ONLY

Telecommunications on the Web

PTTs and Private Telecommunication Providers

Researching Public Telephone and Telegraph organizations (PTTs) and major private telecommunication providers, both national and international, is at the heart of much telecommunications research, and most providers, as well as the growing field of private telecom companies, are coming on line with websites. The hard part can be finding them. Fortunately, several fine websites have compiled lists of links to both public and private telecommunication companies.

Bandwidth Market Telecom Links

http://www.bandwidthmarket.com/component/option,com_weblinks/ltemid,4/

BWM's links pages offer an exceptionally fine collection of telecommunications-related information. The site includes a hyperlinked list of **telecom operators** alphabetically by country. This page makes it extremely easy to see which telecom companies operate in which country and go directly to the company pages. The site also has an impressive hyperlinked list of over 200 **telecom manufacturers** from around the world listed alphabetically. These are the categories and number of links at BWM's site:

- Analysts and Consultants (57)
- Associations and Organizations (32)
- R General (6)
- * Government Sites (36)
- Internet Exchanges (11)
- " Internet Telephony (15)
- Maps (165)
- Satellite Telecom (26)
- * Telecom Manufacturers (203)
- Telecom Operators (184)
- Telecom Publications (153)
- Telecom Standards (16)

ITU's Global Directory of Regulators (select Regulators for PTTs)

http://www.itu.int/cgi-bin/htsh/mm/scripts/mm.search

The ITU's Global Directory may be the most authoritative source of PTT's available. The site lists all members and can be searched by country, company, or organization. The site may load slowly.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

International Telecommunication Union Our Sites News Events Publications Site Map About Us	Français Print V
Home: Overview: Global Directory	
ITU Global Ofrectory	
TIES Main Page I FAQ I Email I GD Map Wed Jan 24 15:38:16 MET 2007	
Search the ITU Global Directory 🚧	
desult of your search: [Regulators] display totals?	
Name Albanie - Albania - Albania	ITU-R ITU-D Categorie
Interraction and the Submisses Televation , TIRANA	: ADMIN RELATED / REGULATOR
The electron last or feep paron of this , TIRANA	ADMIN RELATED / REGULATOR
Allemagne - Germany - Alemania	Control of the Contro
Foder of the Control of the Control of the Control of the Analysis of the Control	ADMIN RELATED / REGULATOR
Andorre - Andorra - Andorra	
Statistics - electrophymnistic (6-vi) van (5 tr) , ANDORRA	ADMIN/REGULATOR
Angola - Angola - Angola	- M
Secretarity to at the engine retrient (NALLEY) LUANDA	ADMIN RELATED / REGULATOR
Antiqua-et-Barbuda - Antigua and Barbuda - Antigua y Barbuda	
Call to set the anathrag Excepting track Programmed attack, Telecommunication Division , ST, JOHNS	ADMIN/REGULATOR
Antilles neerlandaries - Hetherlands Antilles - Antillas Neerlanderas	
Part ou of Felt communeration and Pert , CURAÇÃO	ADMIN RELATED / REGULATOR
Stable saoudite - Saudi Atabia - Atabia Saudita	2 1320
Color in faction from the book for the property of operation , RIYADH	ADMIN RELATED / REGULATOR
Argentine - Argentina - Argentina	
I and lander, must be a managed area to the BUENOS AIRES	ADMIN/REGULATOR
Almente - Almenia - Almenia	and the same of th
Public terior feed at ay Los Fither, YEREVAN	ADMIN RELATED / REGULATOR

World Wide Web Telecommunication Resource Center

http://home.planet.nl/~wvhwvh/teletop.htm

Country Index for Major PTTs, PTOs, and Major Service Providers

http://home.planet.nl/~wvhwvh/countidx.htm

The World Wide Web Telecommunication Resource Center has a separate listing of over 300 links to PTTs and major service providers in 120 countries. This site was last updated in 2004, but much of this information has not changed.

Goodman's International Telecom Companies

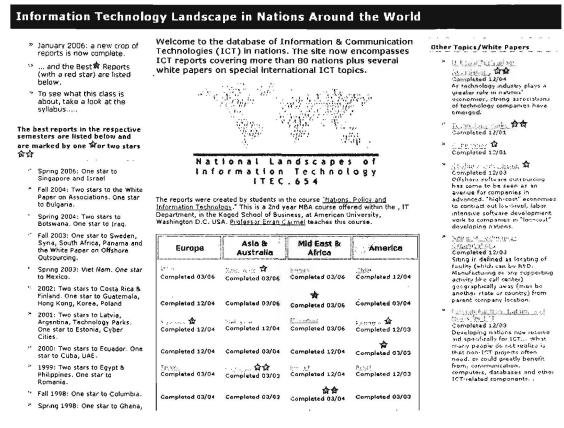
http://www.gbmarks.com/html/international.html

Another excellent source is Goodman's International Telecom Companies, which lists and links to telecom and telephone providers and authorities for each country alphabetically.

Finally, I suggest you stop by a website maintained by the American University's Management of Global Information Technology program. Each semester students prepare descriptive and analytical reports on all aspects of one country's information technology. Be aware the reports vary in quality, but every report is worth reading if only for the bibliography and links to other websites. The best reports (as judged by

UNCLASSIFIED#FOR OFFICIAL USE ONLY

the instructor) on IT Landscape in Nations Around the World are indicated with stars.



IT Landscape in Nations Around the World http://www.american.edu/academic.depts/ksb/mogit/country.html

Telecommunications Directories

Analysys Telecoms Virtual Library

http://www.analysys.com/vlib

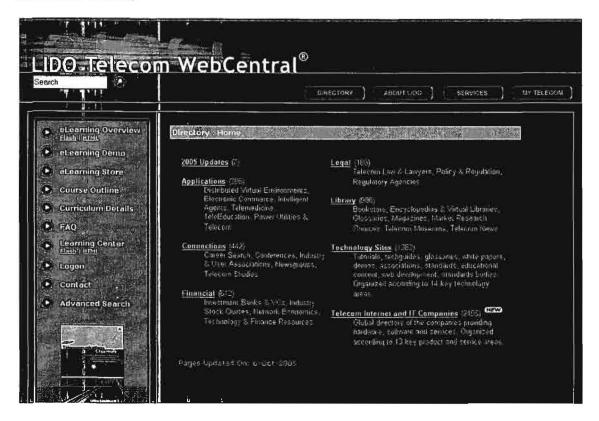
Analysys Telecom Virtual Library has over 8000 links to international telecom sites. The links can be browsed by topic or searched by keyword. This site is one of the best for finding international telecom providers of all types.

LIDO Telecom Web Central http://www.telecomwebcentral.com/secure/links/

One of the best directories of telecommunications information on the web, LIDO offers an easy to use directory or internal search for full range of telecom topics. - LIDO Telecom Web Central is a searchable web directory covering such telecommunications topics as applications, IT technology, companies, regulations, and

UNCLASSIFIED#FOR OFFICIAL USE ONLY

research resources. The impressive directory is kept fairly current and contains thousands of links.



Computer and Communication Entry Page

http://www.cmpcmm.com/cc

The Computer and Communications Entry Page has great links to companies, mostly in the US but some international. It's also an excellent source for standards as well as for programs and projects. The site is internally searchable and once again appears to be up to date.

Goodman's Bookmarks

http://www.gbmarks.com/

Goodman's Bookmarks have also been around for years and has undergone a major makeover. Goodman covers such topics as wireless, IP telephony, vendors, and international telecommunications providers. In fact, his list of international telecom operators is one of the best on the web, providing links to the websites of almost all international telephone and telecom companies and organizations.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Radio, Television, and Satellite Broadcast Information

Other useful sources of telecommunications information include sites devoted to broadcast radio and TV stations and to satellites. Many radio stations around the world now have audio feeds to permit users to listen to them over the Internet, and television is also making more and more of its content accessible over the Internet. There are a number of excellent sites that will help you locate these stations and find out which ones have Internet feeds. Communications satellites are often the primary source in many parts of the world that lack the infrastructure for terrestrial communications.

I have included a link to the Union of Concerned Scientists' satellite database, which includes information not only on communications but also other types of satellites: "the UCS Satellite Database is a listing of operational satellites currently in orbit around Earth. It...contains only the official name of the satellite in the case of government and military satellites, and the most commonly used name in the case of commercial and civil satellites."

Radio Locator

Radio, TV, and Satellite Links

Live Radio

Radio Station World

Mike's Radio World

http://www.nadio-locator.com/
http://www.live-radio.net/info.shtml
http://www.live-radio.net/info.shtml
http://radiostationworld.com/default.asp

vTuner http://www.vtuner.com/

USC Satellite Database

http://www.ucsusa.org/global_security/space_weapons/satellite_database.html

Heaven's Above Satellite Database http://www.heavens-above.com/selectsat.asp

SatcoDX Satellite Chart http://www.satcodx.com/eng/

NASA's J-Track Satellite Tracking http://science.nasa.gov/RealTime/JTrack/

Small Satellites Home Page http://centaur.sstl.co.uk/SSHP/

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Research How-Tos

Finding People

While searching for people remains one of the most popular and important Internet research projects, it also continues to be one of the most frustrating. It is also one of the most lucrative for businesses—legitimate, shady, or simply incompetent. Sites that claim "you can find anyone, anywhere!" are all over the web. Try running a search for the phrase "search for people" and you are likely to find these categories of results, each of which has its drawbacks and most of which are so limited in some way as to be useless except under special circumstances:

- pay services provided by individuals or companies (everything from sleazy to legitimate PI services)
- for-fee software ("net detective"-type software)
- free software (this probably is a case of "you get what you pay for")
- links to a gazillion "people finder" websites

It all boils down to this: there are three basic types of information on the Internet useful in searching for people:

- 1. information you must pay for
- 2. information you must register to get
- 3. information that is free

While it is true you may have to pay for help and/or information either to find a person or to retrieve information about a person, it isn't always a case of "you get what you pay for" with people searches. People are actually selling books with bloated claims and largely useless tips on looking for people. Most of this so-called "expert" advice is worthless. If you have searched for people on the Internet, you've probably encountered some or all of the suggestions, claims, offers, and outright scams from people claiming to help you find someone using the Internet. Many sites will gladly take your money to search for people. Some of these are legitimate companies with access to huge databases containing a great deal of personal information. Some of these for-profit companies give you nothing more than a list of web links that in turn may only lead you to a 404 error. In many, many cases, you end up paying for information that was freely available elsewhere on the web!

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Internet researchers agree: searching for information about a person is one of the hardest, most frustrating types of research to perform, especially when you're limited to using free resources that do not require registration. The major reasons this is so difficult are:

- The decentralized nature of the Internet. Despite what some people may fear, there is no giant database "out there" where everyone's names and addresses are stored. There are many, many databases all independent of each other spread around the world. Also, there are as yet no reliable publicly available mobile telephone directories.
- 2. <u>How quickly the Internet changes</u>. Web pages come and go, individuals frequently change email addresses and/or service providers, networks go up and down, etc.
- 3. <u>Data about people is often found in the deep web</u>. In most cases, the information is in databases that search engine spiders cannot crawl and index. This means that, in addition to traditional search engines, we have to use other resources to search for people.
- 4. People have become much more conscious of and careful about their personal information. In the early days of the Internet, people thought nothing of putting their contact information on their webpages, using de-obfuscated email addresses in newsgroups, the finger utility was still in use, etc. Today, people are much more likely to hide personal information to protect their privacy and their computers from malicious hackers.

The Usual Suspects Round-up: Debunking People Finder Lists

I get really irritated at sites and people who proclaim how fast and easy it is to find people using the Internet. Because you are going to encounter lots of bogus claims about how simple it is to find people via the Internet, I thought it would be a good idea to go down the list of the usual types of sites recommended for people searches and discuss their advantages (few) and drawbacks (many). What follows is an overview of the usual sources cited as ways to find people and the problems with and benefits of each.

Sites that claim "you can find anyone, anywhere!"

These are all over the web. The bigger the hype, the smaller the likelihood they are worth anything. Approach with care and skepticism or, better yet, avoid altogether.

Big Directories

Number one on virtually all "search for people" lists are large directories (telephone and/or email). They are usually disappointing. What's the

UNCLASSIFIED#FOR OFFICIAL USE ONLY

problem? The big residential telephone directories on the web often come from printed directories that have been scanned in by companies who then make them searchable via the Internet. Email directories are almost entirely made up of addressed culled from USENET newsgroups, voluntary registration, ICQ users, and (sometimes) web pages. Other limitations to big phone and email directories are:

- they generally focus on the US (maybe Canada, too)
- their sources of information are very limited (people registering themselves or from trolling newsgroups)
- they require you already know details about the person you're seeking
- o they are woefully out of date, and/or
- their data is inaccurate

Public Databases

Public databases are similarly frustrating. Almost invariably, the public databases listed on "people finder" sites are US-based and many of the best ones require at least registration and even a justification of use if not payment. I recommend you look for public databases in the country you're researching, but expect to encounter similar problems. In general, public databases suffer from many of the same problems as big directories as well as having drawbacks of their own. Public databases are wonderful sources of information if you are seeking such public records as real estate transactions, lawsuits, liens, judgments, bankruptcies, fire arm transactions, sex offenders, motor vehicle registrations, corporate records, etc. However, this kind of research simply is beyond the purview of this book. Other drawbacks of public databases are:

- they generally focus on the US (maybe Canada, too)
- o they often require registration and/or payment
- they tend to be restricted to specific types of data (e.g., sex offenders or property records in a specific state)
- o they require you already know details about the person you're seeking
- o they are out of date, and/or
- o their data is inaccurate

Private or Internal Directories

These are not inherently bad, but they are by their very nature extremely limited. Many private directories belong to universities, government agencies, corporations, and organizations. If, however, you are looking for someone whom you know works for an organization or attends (or is even an alumnus of) a specific university, these directories can be gold! See below for details.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Genealogy Sites

Genealogy sites always pop up on "search for people" lists—nice for generating your family tree, but not usually pertinent for current research.

Biographical Directories

The big biographical directories are fine for one thing: finding information about well-known and/or successful people living and dead. Otherwise, they are useless. People search lists always include links to "celebrity searches." However, a few <u>specialized biographical directories</u> might prove useful.

The US Census

Census data for a number of different years is now online and searchable. Again, a great genealogy tool, but probably not very useful for your research.

> Fee for Service Sites

There are a myriad for-profit people finder websites on the Internet. They may or may not be good, but I am only discussing free products, services, and sites.

> Finger

Finger is a UNIX utility that uses an email address to return information about the owner of the address. Even today finger always seems to appear on people finder webpages, but the fact is that the vast majority of sites disabled the finger command years ago for privacy and security reasons, so it's basically useless.

> The MIT interface to the USENET Address Database

USENET is still accessible via the old MIT interface. I mention the MIT interface (which contains over 4 million email addresses collected from USENET newsgroup postings between July 1991 and February 1996 only) because it continues to show up on "people finder" lists. However, it's almost a decade old, which may as well be a century in Internet time. Instead of wasting time on the MIT USENET interface, why not go directly to Google Groups? One reason: MIT does not obfuscate email addresses; Google now does, so you might find an older but full email address in the MIT USENET database that is obscured in Google Groups.

MIT USENET Address Database (1991-1996)

http://usenet-addresses.mit.edu/

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Recommended People Search Techniques

Now that you have a pretty good idea of sites to avoid or at least not waste a lot of time on, what kinds of resources are available to help you find people? How do you find information about a person? Here are some recommended steps to use in searching for people, keeping in mind that there is no one way to do this and certainly no guarantee of success.

- 1. Start a search for a person by name, address, email address, phone number or any other uniquely identifying information using not one but several good search engines (Google, Yahoo, and several foreign search engines as appropriate). Common names will return too many hits to be useful, so you must find some limiting query term to narrow down the search.
- 2. Try to find the person in Google and Yahoo Groups, weblogs, and news stories using several high-quality news search engines.
- 3. Try telephone, email, address, and other lookups at a variety of online directories. The quality of these directories varies greatly. Be sure to look for directories specific to a city or country.
- 4. If you know the person's profession, you might find additional information about him in a database that contains such things as licensing information. The US is very good about licensing all types of professions; check other countries for similar information.
- 5. Property ownership and transactions are carefully recorded in the US and many such records are publicly accessible. This may also be true in other countries. Look for such public databases of these records and transactions.
- 6. If a person has other interests, such as hobbies, you may find more information about him at a site devoted to that sport or hobby, especially if it's an unusual one.
- 7. If you know where a person works, that organization (be it government, academic, or corporate) may have a publicly accessible directory of its employees, faculty, students, alumni, or members.
- 8. Check the ICQ directory. Instant Messaging (IM) is popular worldwide.
- Whois databases contain information about thousands of people associated with the Internet. It's worth a look. The Whois databases maintained by ARIN, APNIC, AfriNIC, LACNIC, and RIPE are all searchable by name using their advanced search forms.
- 10. Famous dead people are not the only ones listed in biographical directories. If the person you're seeking is at all well known, he might have an entry.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Recommended People Search Resources

There is no Internet equivalent to the Rosetta Stone for finding people. Searchers need a good set of bookmarks that include at least the following sites.

ICQ User Directory

http://people.icg.com/whitepages/

Unlike AOL's AIM instant messaging system directory, ICQ's phone and email directory is publicly available, listing more than 120 million users in over 245 countries. The ICQ User Directory's Advanced People Search is impressive. You can search on as few or as many fields—including first/last name, nickname, email address, language, phone number, country, occupation, etc.—as you wish. While there are several interfaces to ICQ search (depending on whether you want to search by email address or phone number, for example), I recommend starting with the main People Search page and then going to the Advanced People Search for more options.

Google Groups

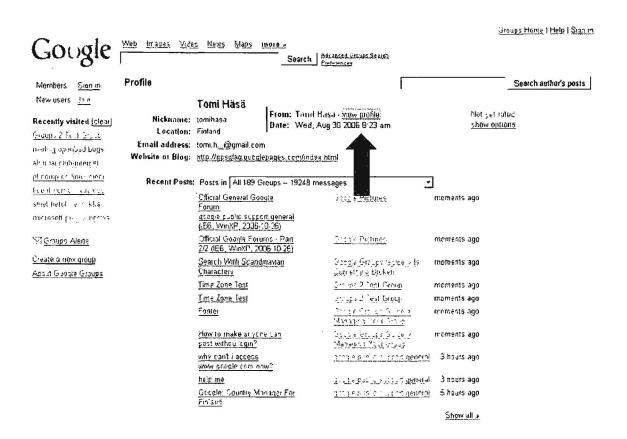
http://groups.google.com/

Not only is Google Groups much more user friendly than the MIT USENET interface, it is also current, including postings in almost real time. Google Groups can be searched in many ways, but for the purpose of locating people, the *Author* search is exceptionally valuable. However, keep in mind that anyone can make up an alias, change addresses, and use other people's email, so USENET postings are of real but limited value. Also remember that USENET postings are one of the places where a sort by date works and makes sense because all newsgroup postings are date/time stamped. There are three basic ways to run an *Author* search (important note: *Google Groups now obscures the userid in email addresses by inserting an ellipsis (...) for privacy reasons*; however, many obscured addresses can be easily guessed). However, *if you know the author's full email address*, you can use search syntax in the main Groups & Groups Beta query screen—[author:email address]. This will return all the postings associated with that email address.

Even if you do not know the full email address, you can still see all the posting associated with one email address (these instructions apply to Google Groups Beta):

- While viewing a specific posting, click on "More options" (next to the date);
- Select "View profile" next the author's From address;
- Google will automatically search for and display every posting from a particular address.

UNCLASSIFIED#FOR OFFICIAL USE ONLY



Yahoo Member Directory

http://members.yahoo.com/

Similar to Google Groups, Yahoo's Member Directory lets you view information about the registered members of its groups. Some people provide a lot of personal data, including their real name, location, age, occupation, and email address (and sometimes even a photograph). I am skeptical that you will find anyone of great interest unwise enough to provide this kind of data, but it is worth a look.

Specialized Biographical Directories

While the big biographical directories are virtually useless, there are a few specialty directories that include the less than famous that you might want to know more about. The Forbes People Lists include: 400 Richest Americans, Best Paid CEOs, World's Richest People, and the Midas List of the top investors in information technology and life sciences. The Biography Center lists nearly 25,000 biographies total, with almost 11,000 in English, 6,000 in French, 7,000 in German, and 2,000 in Spanish. The search form requires you to search in one of these languages; there is no way to search in all at one time and the results are different for each. There is also a Chinese Biographical Database containing 3500 biographies, all in English. Wolfram's Science World Biography includes famous and not so famous scientists living and dead.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Forbes People Lists http://www.forbes.com/lists/

Biography Center http://www.biography-center.com/

Biography Reference Center from MacGill University

http://www.library.mcgill.ca/refshelf/biograph.htm

Chinese Biographical Database http://www.lcsc.edu/cbiouser/

ISI Highly Cited Researchers http://www.isihighlycited.com/

Wolfram's Science World Biography http://scienceworld.wolfram.com/biography/

Edgar Online Database

There are several free search tools of SEC filings that allow you to search by a person's name or by a specific company name. As of mid-November 2006, the users can "search the contents of the disclosure documents filed electronically with the SEC using a new full-text search tool on the Commission's website. The newly searchable information includes registration statements, annual and quarterly reports, and other filings by companies and mutual funds filed during the past four years on the Commission's EDGAR database." This means users can search on names of people as well as companies.

SurfWax offers a wonderful tool using its "LookAhead" (AJAX) technology to search the SEC EDGAR filings for 2004-2006. This is a very flexible and powerful search tool, and it permits searches on the names of individuals.

The Price Waterhouse Coopers' website now offers free full-text search of EDGAR filings, which means it can search on names. Even though others have said this query tool is slow, my trial searches ran fairly quickly. My impression is this is a remarkably powerful tool, especially given it is free and does not require registration. Make sure you use the link below to access the advanced search of the free service.

Search the SEC's Edgar Database

http://searchwww.sec.gov/EDGARFSClient/jsp/EDGAR MainAccess.jsp

SurfWax SEC Search

http://lookahead.surfwax.com/edgar/

EdgarScan Advanced Search

http://edgarscan.pwcglobal.com/servlets/advancedsearch

¹³⁸ "SEC Enhances Online Search Capabilities for Investors," SEC Press Release, US Security and Exchange Commission, 14 November 2006, http://www.sec.gov/news/press/2006/2006-190.htm (20 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

University and other organizational phonebook directories

Such directories are usually very good sources of current information about personnel at a specific university, in an entire university system, employed by a company, or working for a government agency. Many universities use the PH system or WebPH, which is a web interface PH. PH is a UNIX-based Internet utility that lets you search for someone's email address if their email provider has a Ph server program. The problem with these PH systems is that there is no reliable way to locate them even if you know the domain name of the institution. However, as with most things on the Internet there are lists of publicly accessible PH servers available in several places. There are about 330 PH servers around the world currently on the PH server list, which is accessible at various locations such as:

Queen's University (Canada) http://alpha9.drake.edu/cgi-bin/WebPh?other ph servers
University of Medicine and Dentistry, New Jersey

http://www2.umdnj.edu/~webph/cgi-bin/WebPh?do=show others

Your best bet for using some sort of internal directory successfully is to identify the university, company, or other organization (including government agencies) and locate the directory at that website. Directory services are generally very popular, so most sites will have a link to theirs prominently displayed on their home page. If you don't see a link on the home page, try browsing the site map.

Whois Lookups

These should never be overlooked as a source of information about people. All five of the Regional Internet Registries (ARIN, APNIC, AfriNIC, LACNIC, and RIPE) have Whois access that permits users to search on any of the database fields, including the person object, i.e., the technical or administrative contact responsible for the registration. However, keep in mind people lie all the time when registering in Whois databases, so all information obtained from Whois entries must be validated using other sources. To learn how to search the three identically formatted Whois databases, see the instructions (RIPE's documentation is the most thorough, APNIC's the easiest to follow).

APNIC Whois Database Query Option

http://www.apnicnet/db/search/all-options.html

ARIN Whois Database Search Help

http://www.arin.net/tools/whois help.html

AfriNIC Database User Manual

http://www.afrinic.net/docs/db/afsup-dbgs200501.htm

RIPE Database Reference Manual

http://wwwripe.net/ripe/docs/databaseref-manual.html

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Weblogs

Blogs may indeed mention people that traditional sources miss. Blogs are better than traditional search engines for finding people whose names just surfaced in the news. See the blogsearch section for the best blog search engines.

News Search Engines

These are great for timely searching, which means they may be the best source of information on someone who has recently gained his/her fifteen minutes of fame. As with USENET groups, news searching is amenable to date/time order because news stories, unlike web pages, have dates and times associated with them. They are not generally good for historical research, i.e., more than one to three months old. There are numerous excellent news search engines detailed in the News Search Section.

Online Telephone Directories

Telephone directories may be useful or merely frustrating. Despite the drawbacks of many telephone directories, some are quite good. More and more cities and even whole countries are putting their white and yellow pages on line, often with flexible lookups, including reverse lookup. Check a country's search engines and directories for listings of telephone directories. Information on finding and using telephone directories on the web is available in the Telephone and Fax Directories section.

Email Directories

While these remain spotty and often unreliable, do not overlook this potential source of information that is growing and improving all the time. Information on finding and using email directories on the web is available in the <u>Email Lookups</u> section.

Recommended People Finder Websites

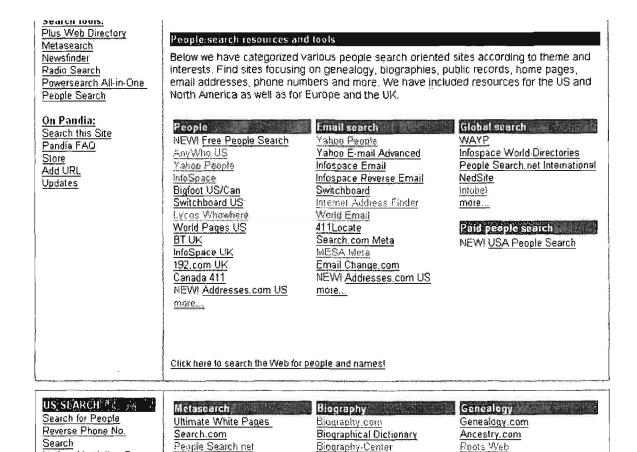
I have tried to give you some idea of the types of information and links you will encounter at people finder websites so you won't be surprised or become discouraged if they do not turn out to be all they claim. Still, you will have to sift through a lot of silt and sand to find the nuggets, but they are there. I think these are the best general sites for getting started looking for people.

Pandia People Search

http://www.pandia.com/people/

I would start here. One of the best web search sites for locating people, Pandia Search has a terrific section on people search that includes both metasearch interfaces as well as many links to the best people finder resources.

UNCLASSIFIED#FOR OFFICIAL USE ONLY



The Virtual Chase People Finder Guides

http://www.virtualchase.com/topics/people_finder_index.shtml

whee People Tra

Another excellent web search site, the Virtual Chase is devoted to legal research and, as such, its people finder guide focuses on US public records but includes reverse lookups, white pages, email directories, and non-US phone books. The first link will help you find online sources of public records and other people search tools. The second link is to an "annotated research guide to find sources of information about people" as well as articles on people research.

Searchbug People Finder

Nationwide Yellow Pages

http://www.searchbug.com/peoplefinder/

Family Soc

Searchbug is a metasearch site with links to many resources, but as with most such resources, the most potentially lucrative are not free or at least require registration. You are probably going to be disappointed when you discover that most of the links are to for-fee services. Also, the focus is distinctly US.

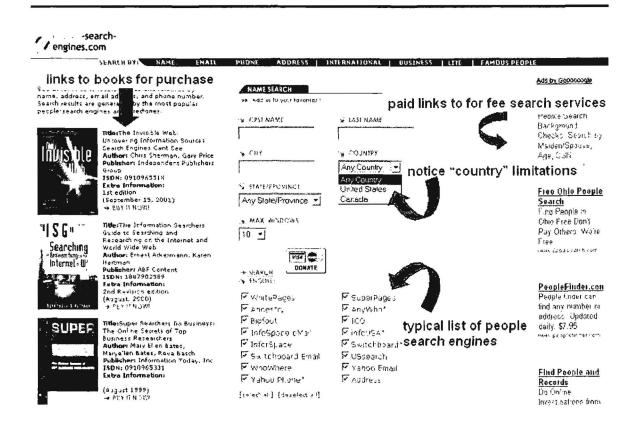
People Search Engines

http://www.people-search-engines.com/

This is a good metasearch site that searches some or all of multiple sources by name, email, phone, or address. This site typifies a respectable people metasearch

UNCLASSIFIED#FOR OFFICIAL USE ONLY

site. Unlike most people search sites, this one also has an "international" page where you can query resources that are not limited to the US and Canada:



PeopleSearch.net http://peoplesearch.net/peoplesearch/peoplesearch deluxe.html

This site is almost identical in the search engines it queries to People Search Engines, but without the international option.

ZoomInfo

http://www.zoominfo.com/

ZoomInfo is a "web summarization" site that contains over 27 million summaries of people on the web. According to its FAQ, "ZoomInfo automatically and continuously grows its base of Web Summaries from corporate and personal websites, government filings, press releases and other public sources. All information found by ZoomInfo and used to create your Web Summary comes from public sources and can be found by anyone by using most major search engines like Yahoo and Google." ZoomInfo does summarizes that data and presents it in easy to handle small packages. ZoomInfo is integrated into A9 search; see that section for details.

The Virtual Gumshoe

http://www.virtualgumshoe.com/

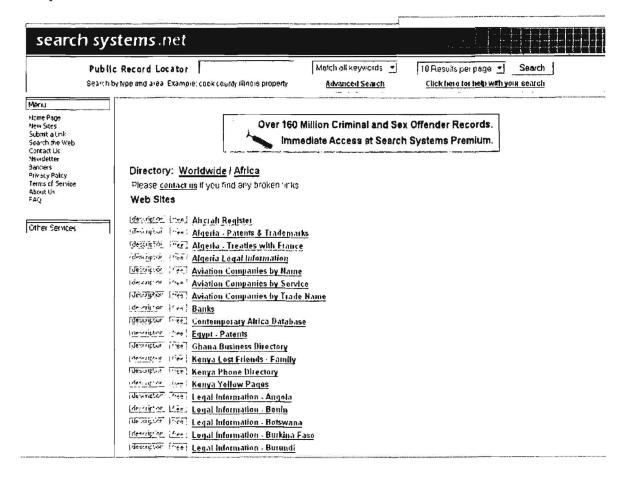
The Virtual Gumshoe provides lots of links to all sorts of investigative resources. The site contains many interesting categories, including gangs, terrorists, cults, criminal

UNCLASSIFIED#FOR OFFICIAL USE ONLY

histories, etc. According to the site, "We provide links to every information source known to us that can be obtained on the Internet for free. Some categories may list a professional company as a source for information, which is not free. This is merely an option when there is no legitimate source online (known to us) that provides the information for free." And there are plenty of categories of information you are not going to get (legitimately) for free.

Search Systems Free Public Records Database http://www.searchsystems.net

With links to nearly 30,000 mostly free databases, this is one of the most impressive of the people finder megasites. What makes this site especially valuable is the large Worldwide section, which is further divided into continents and regions. Here is a snapshot of the Africa page and, as you can see, the vast majority of resources really are free:



Langenberg.com Person Finder

http://person.langenberg.com/

This is an easy to use megasearch site that lets you query individual search tools from one page. Resources include the Lycos' WhoWhere phone book, Yahoo

UNCLASSIFIED#FOR OFFICIAL USE ONLY

members search, Google Groups search, ICQ, and Live Spaces (still labeled as MSN Member Directory), among others.

Public Record Finder Outside the US

http://www.publicrecordfinder.com/outside_usa.html

Notable for its unusual nature, this site lists a moderate number of international public records' resources.

Power Reporting People Finders

http://powerreporting.com/category/People finders

A site devoted to helping journalists, this is a no-nonsense guide to realistic people searching. No ads!

People Search Sites http://www.nettrace.com.au/resource/search/people.html

This Australian site will try to sell you people search products but has an impressive list of free resources as well. There are also many more non-US resources at this site for obvious reasons.

Deadline Online's People Finders

http://www.deadlineonline.com/peoplefinders.html

Another good megasite with no ads! A very clean interface and many useful people finder tools neatly organized by topic.

People Search Links

http://www.peoplesearchlinks.com/

This site is actually a directory with categories that include international search, criminal searches, phone & address search, public records searches, vocations, etc.

By this time you are going to start seeing the same resources again and again, which should tell you that you're exhausting the free resources.

Miscellaneous People Finder Websites

These are sources for finding people that do not easily fit into any category.

The Virtual Chase Criminal Records

http://www.virtualchase.com/topics/criminal_records.shtml

This site has many links to Federal, State, and international online criminal resources, including US government lists of blocked, debarred, or denied persons. The Federal Bureau of Prisons inmate locator, which contains records about federal inmates dating back to 1982, is now online and free. Many state prison systems now have online searches for current and former inmates.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

CrimeNet

http://www.crimenet.com.au/

This Australian site provides information on criminals, missing persons, wanted persons, and unsolved crimes in Australia.

The Black Book Online

http://www.blackbookonline.info/

Designed for use in legal investigations, this resource has links to some potentially lucrative US databases. From Crimetime.com, the Black Book includes unusual links such as a mail drop search form (to find out if a legitimate looking suite number is in fact a PO box). Many of the resources listed require payment.

Landings Certified Pilots Database

http://www.landings.com/ landings/pages/search/certs-pilot.html

A database of about 600,000 pilots registered with the FAA who have a current medical certificate. Updated monthly.

NameBase

http://www.namebase.org/

A review describes NameBase as "an index of people influential in politics, the military, intelligence, crime, business, and the media since WWII. It started in the late 1960's when New Left activist Daniel Brandt began clipping magazine and newspaper articles and collecting investigative books about the power structure...NameBase includes close to 100,000 names from approximately 260,000 citations. The names are drawn from over 700 books and serials...most books and articles come from a leftist perspective." Many of the resources require payment, but a basic search does not. The site is worth a look because of the extensive cross-referencing of names from various sources and the "Social Network Diagram" available for each name mentioned.

Final Thoughts on Finding People

Do not overlook search engines. They may still be your best way of finding people on the Internet because they crawl the web almost constantly, index vast amounts of data, index everything on a website (including names, addresses, phone numbers, email addresses, etc.), and they are fast and easy to use. Also keep in mind there is a remarkable lack of overlap among search engine results, so use a number of different ones or several metasearch engines in your search.

UNCLASSIFIEDHFOR OFFICIAL USE ONLY

Web Tip Look for information in unconventional sources.

Sometimes calling codes and dialing information can come from unusual sources. For example, one of the best explanations of the arcane dialing rules and list of up-to-date calling codes for the CIS countries is located, of all places, at the "Russian Bride" website. And why not? After all, these folks are running a business and must provide accurate information about how to contact their clients.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Using the Internet to Research Companies

One of the main issues researchers face when looking for information about companies is there either seems to be too much or too little available data. How is it possible to have both a superfluity and a paucity of information at the same time? When it comes to businesses, there is too much data that is not useful (advertisements, sale's pitches, etc.) and too little of the kinds of data most researchers need (details about ownership, personnel, products, contracts, etc.). However, there are some good starting points for researching companies on the Internet, and I have put together a set of tips and techniques to address this challenging topic.

But what about using free Internet resources to research companies? Yes, it can be done with success and sometimes there is information on the Internet that is not available anywhere else. What follows are some of the steps I recommend for researching companies on the Internet using only free and non-subscription resources. Those that require payment and/or registration are not discussed here.

Company Research Guides & Tutorials

Several sites have excellent online tutorials and guides for how to research businesses.

Researching Businesses and Non-Profits on the Web

http://www.ojr.org/ojr/technology/1028068074.php

From the University of Southern California's Online Journalism Review; heavily US-oriented.

Researching Companies Online

http://www.learnwebskills.com/company/

Excellent tutorial that includes separate section on identifying international business resources.

Virtual Business Information Center

http://www.vbic.umd.edu/

From the University of Maryland School of Business; offers a segment devoted to international business and specific countries.

Virtual International Business and Economic Sources

http://library.uncc.edu/display/?dept=reference&format=open&page=68

From the University of North Carolina, Charlotte; the site's emphasis is entirely international.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

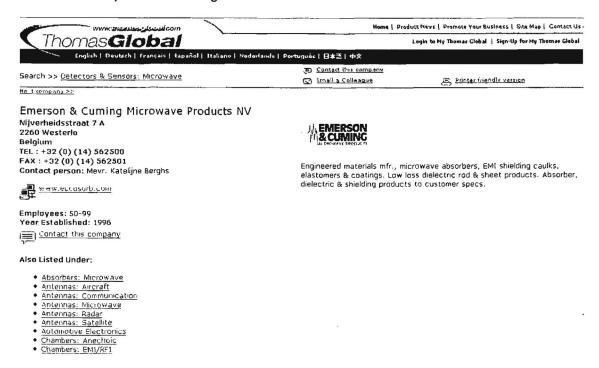
Free Company Research Websites

Two well-known company research tools have been available on the Internet since 2004, but both have severe limitations. In the past the excellent Thomas Register's search tool ThomasNet was restricted to US and Canadian companies. Thankfully, in 2006 Thomas Register combined its search products into ThomasGlobal, which still does not require either registration or payment to use. ThomasGlobal is a search tool for international products/services or company names and includes an option to limit a search by location and to browse by product/service category.

ThomasGlobal

http://www.thomasqlobal.com/

ThomasGlobal provides a brief company profile (including links to any website), contact and location information, other headings under which a company is listed, and a link to a product catalog where available.



This is a very well designed tool and its global reach makes it extremely powerful and, in fact, invaluable for company research.

Virtual Chase Company Information Guide

http://www.virtualchase.com/topics/company information index.shtml

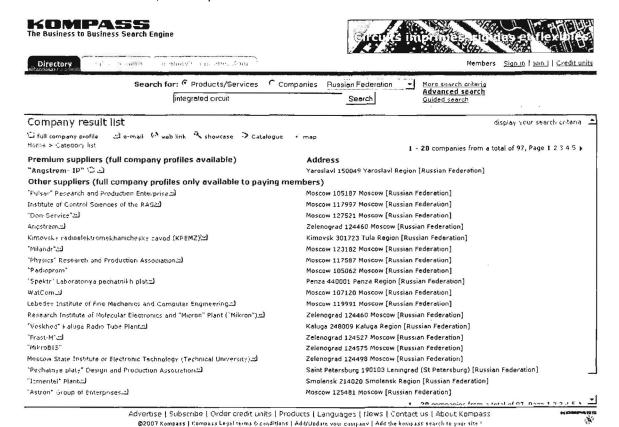
Virtual Chase provides an annotated research guide to sources of information about companies, executives, court records, SEC filings, news and even public opinion about companies. The site also includes a number of articles on how to research companies.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Kompass

http://www.kompass.com/

Kompass business-to-business search is among the best international company search sites. Users can search the huge Kompass database of over 2 million companies in 70 countries by products/services or company name, and limit the search to any combination of countries or world regions. While users must pay for full company profiles, searches are free and provide full company names, addresses, telephone/FAX numbers, website addresses, and a full list of companies' products and services. Here are the results of a search for "integrated circuit" in the Russian Federation, which produced 97 hits:



The Scannery

http://www.thescannery.com/

The second tool is much harder to use, or at least that is my experience. Called "The Scannery," it is a database of more than 12,000 public companies around the world organized by country and sector. The database contains many types of documents, including HTML, PDF, PowerPoint, Word, and more. The Scannery also offers many search options, including boolean operators, wildcards, "fuzzy" matches, synonym searching, etc. Sounds great, doesn't it? The only problem is, it's royally confusing and the results are often completely puzzling. For example, an unlimited search on IBM brings up as the top three hits, QLOGIC, Cable & Wireless, and Synopsys, but

UNCLASSIFIED#FOR OFFICIAL USE ONLY

not IBM. Why? Because according to Joseph Pols, the developer of The Scannery, "not all companies allow their websites to be indexed...Therefore, the only results would be those documents on other websites which contain the search phrase."

However, even given this limitation, The Scannery is a very useful tool. Most companies do permit indexing, so chances are the ones you are researching will be accessible via The Scannery. The problem usually isn't too little information but too much. For example, an unrestricted search on [alcatel] brings back more than 3100 hits! When I limit my Alcatel search to France, I get a more manageable 74 hits. Keep in mind that what you are seeing is any mention of your query term in any document indexed by The Scannery, so my search for Alcatel will return not only documents from the Alcatel sites but also documents at other company websites that mention Alcatel.

The Scannery has a related site with the oh-so-clever name TimBuckOne. TimBuckOne will permit users to choose a country, then choose a specific company in that country, and then search that company's website. However, TimBuckOne requires free registration in order to use.

SEC's Edgar Database

Search the SEC's Edgar Database

http://searchwww.sec.gov/EDGARFSClient/jsp/EDGAR MainAccess.jsp

SurfWax SEC Search

http://lookahead.surfwax.com/edgar/

EdgarScan Advanced Search

http://edgarscan.pwcglobal.com/servlets/advancedsearch

There are several free search tools of SEC filings that allow you to search by a person's name or by a specific company name. As of mid-November 2006, the users can "search the contents of the disclosure documents filed electronically with the SEC using a new full-text search tool on the Commission's website. The newly searchable information includes registration statements, annual and quarterly reports, and other filings by companies and mutual funds filed during the past four years on the Commission's EDGAR database." This means users can search on names of people and companies.

SurfWax now offers a wonderful tool using its "LookAhead" (AJAX) technology to search the SEC EDGAR filings for 2004-2006. This is a very flexible and powerful search tool, and it permits searches on the names of individuals.

¹³⁹ "SEC Enhances Online Search Capabilities for Investors," SEC Press Release, US Security and Exchange Commission, 14 November 2006, http://www.sec.gov/news/press/2006/2006-190.htm (20 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

The Price Waterhouse Coopers' website now offers free full-text search of EDGAR filings. Even though others have said this query tool is slow, my trial searches ran fairly quickly. My impression is this is a remarkably powerful tool, especially given it is free and does not require registration. Make sure you use the link above to access the advanced search of the free service.

More Company Research Sites

Keep in mind that even those sites that do require payment and/or registration for full access often provide some information for non-subscribers. Try these sites:

Annual Reports from Report Gallery

http://www.reportgallery.com/

Free annual reports searchable by company name, stock ticker, or sector. Reports may be in PDF or HTML formats.

Arab Data Net

http://www.arabdatanet.com/

Large amounts of information about business in the Arab world; be sure to check the Directory for data on companies.

Business.com

http://www.business.com/

Excellent business-related search engine/directory; not limited to US companies

Business Information on the Internet

http://www.rba.co.uk/sources/index.htm

Great UK site with huge number of resources for researching businesses not only in the UK but around the world. Working through this site is like taking a research tutorial.

Corporate Information

http://www.corporateinformation.com/

Reports by Wright Investors' Service; full access requires registration; the available free information is good.

Free Reports for Top 20 European Companies

http://amadeus.bvdep.com/amadeus/top20/ top20.htm

Only 20 companies, but if you are researching of them, you have hit the jackpot with this site. These are full financial reports for free.

Global Edge International Business Research (Michigan State University)

http://globaledge.msu.edu/ibrd/ibrd.asp

The Center for International Business Education and Research at Michigan State University maintains Global Edge, a site of international country studies and business information with links to over 5000 resources. The site includes current information on the business climate, economic landscape, and relevant statistical data for 197 countries. "Powerful features such as comparing countries using

UNCLASSIFIED#FOR OFFICIAL USE ONLY

multiple statistical indicators and ranking countries based on a selected statistical indicator are available. A rich collection of country-specific international business links adds to the vast collection of information."

Hoovers http://www.hoovers.com/

Hoovers is one of the best and most comprehensive company, industry, and market research resources. However, full access requires subscription. Even so, the information at Hoovers' free site is fairly extensive.

Irasia Investor Relations Asia

http://www.irasia.com/listco/

Asian Company Reports; excellent data on Asian companies by name or sector.

MacRae's Blue Book

http://www.macraesbluebook.com/

MacRae's EuroPages Search

http://www.europages.net/co branding/macraesbluebook/home-en.html

MacRae's Blue Book is a guide to more than a half a million US and Canadian industries and over 2 million product listings; search by product or company. MacRae's now offers a EuroPages search of European companies at a separate link.

Market Access and Compliance

http://www.mac.doc.gov/

US Department of Commerce site that is packed with information about trade and development around the world.

MSN Money's Key Developments

http://news.moneycentral.msn.com/ticker/sigdev.asp

Major milestones for publicly traded US companies; search by name or ticker symbol and see all developments, earnings, or product announcements.

PRNewswire

http://www.prnewswire.com/

A great source for press releases; search by company, keyword, or limit your search by country.

SEDAR

http://www.sedar.com/

Canadian securities-related information.

Yahoo Finance Press Releases

http://biz.yahoo.com/prnews/

Easily searchable source of PRNewswire press releases.

Company Websites

Look at a company's own websites, but don't limit your search to the official home page. Be sure to check out country-specific sites (most international companies

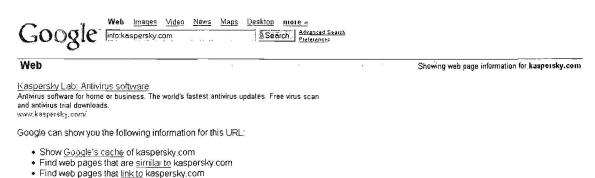
UNCLASSIFIED#FOR OFFICIAL USE ONLY

have sites specific to a country or region) and alternative domain names. Many companies have more than one domain name associated with them. For example, the Russian firm Kaspersky Labs has sites in multiple countries. In addition to the .com and .net domains, I counted 20 country top-level domains associated with the domain name *kaspersky*. I explored these names using Yahoo's <u>Site Explorer</u> and Google's *info*: command.

Google's Info: Command

This is a valuable tool for getting a fast overview of a website. Simply by inserting an address—kaspersky.com—into the Google search box, you will see the following options:

- view Google's cache of kaspersky.com
- find other pages similar kaspersky.com
- find external pages linking to kaspersky.com
- find pages containing kaspersky.com



Whois Lookups

Find web pages from the site kaspersky.com
 Find web pages that contain the term "kaspersky.com"

While people can and do lie about the information they enter into domain registration forms, the Whois databases remain an excellent source of information about companies and people (addresses, phone/fax numbers, email addresses, ISPs, mail servers, etc.). However, I strongly caution you to double-check these data with other sources to validate them.

ARIN (North America and some Caribbean)
RIPE (Europe, Middle East, North Africa)
APNIC (Asian-Pacific)
AfriNIC (Africa)
LACNIC (Latin America and some Caribbean)

http://www.arin.net/whois/index.html
http://www.ripe.net/perl/whois
http://www.apnic.net/apnic-bin/whois.pl
http://www.afrinic.net/cgi-bin/whois
http://lacnic.net/cgi-bin/lacnic/whois

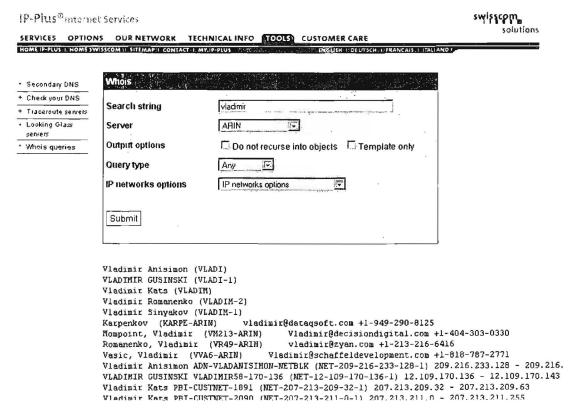
UNCLASSIFIED#FOR OFFICIAL USE ONLY

Look for Additional Domains in Whois Database

As an example, Kaspersky Labs owns and operates websites under the domain names *kaspersky-labs* and *avp* (for Anti-Virus Protection) for many country top-level domains as well as the .com and .net domains. How did I find these alternate domain names and websites? By investigating Whois data on Kaspersky Labs and noting email addresses @avp.ru and other server names (*kaspersky.com, kasperskylabs.net, kaspersky-labs.com*).

Correlate Whois Data

For example, when you find a person's name (the *person* object is the technical or administrative contact for a Whois entry), I suggest you search on that name in all the major Whois databases to see if his/her name shows up anywhere else, thus providing possible leads to relationships between seemingly unrelated companies or organizations. To see how this works, search on the name [vladimir] in the RIPE, ARIN, and APNIC databases at IP-Plus:



IP-Plus http://www.ip-plus.ch/tools/whois_set.en.html

Also, search on the name using search engines, phone directories, email lookup tools, etc.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Wildcard Whois Search Tools

One of the best ways to find other domains registered to a company is to use a wildcard Whois lookup tool. There are a number of these available on the Internet and they are detailed in a <u>later section</u> of this book. Wildcard searches work best with unusual or distinctive domain names.

Domain Surfer

Domain Tools Whois Source

Namedroppers.com

Netcraft

Whois.net

Whoix?

Advanced Search

http://www.domainsurfer.com/
http://www.domainsurfer.com/
http://www.namedroppers.com/
http://www.namedroppers.com/
http://www.namedroppers.com/
http://www.netcraft.com/
http://www.whois.net/
http://www.whoix.com/

Explore the Site

There are any number of things you can learn from looking at the various pages at a company website. Fortunately, it is easy to display all the pages at a website (at least all the pages indexed by a specific search engine); unfortunately, for a large website, finding useful information within all these can be a daunting task. The best tools for site exploration are Yahoo's Site Explorer and the various site: and link: commands available at major search engines.

In addition to Yahoo's Site Explorer, here is a summary of the various "site exploring" syntax available in Yahoo, Google, and Live Search:

Yahoo Search:

➤ site/domain: restricts results to a specific website or domain, including a specific top-level domain. May be used with or without keywords. [site:fr] returns all the pages Yahoo has indexed in the French top-level domain

[site:who.int sars] finds all the pages at the WHO website that contain the keyword SARS; will find pages at wpro.who.int as well as www.who.int.

Google Search:

site: restricts results to websites in a given domain. May be used with or without keywords.

[site:in] returns all the pages Google has indexed in the Indian top-level domain

UNCLASSIFIED#FOR-OFFICIAL USE ONLY

[site:who.int "avian flu"] finds all the pages at the WHO website that contain the phrase "avian flu"

[site:who.int "avian flu" -inurl:speeches] finds all the pages at the WHO website that contain the phrase "avian flu" and exclude any page with a url containing the keyword "speeches."

[cirrus -site:mastercard.com] finds pages containing the keyword cirrus that are not on the Mastercard website

Live Search:

site/domain: restricts results to a specific website or domain, including a specific top-level domain. May be used with or without keywords. [site:gov.it] returns all the pages Live Search has indexed in the Italian Government's domain [site:gov.it language:en] finds all the pages in the Italian Government's domain that are in English [site:gov.it -filetype:html] finds all the pages in the Italian Government's domain that are not in HTML format

Link: Command at Major Search Engines

It is always worthwhile to see who is linking to a site. Best bets for running the link command are the biggest search engines because, you guessed it, they have the biggest databases! If you get too many hits, try limiting the search by restricting the results to a specific domain or adding a keyword when permitted (Google does not permit the use of keywords with its link: command). All these search engines use the same syntax for link searching, i.e., link:, but remember, Yahoo has two link commands.

Also, the <u>Yahoo Site Explorer</u> is extremely good at showing "inlinks" to an entire domain or just a specific url. Unlike Google, Yahoo Site Explorer provides all the links to a site (which Yahoo calls "inlinks") instead of a limited subset of links.

Yahoo Search:

▶ link: restricts results to pages containing links to a specific url. Yahoo requires users to enter the full url including the http:// for the link: command. May be used with or without additional keywords and syntax.

[link:http://www.who.int] finds all the incoming links to this specific page

[link:http://www.who.int "avian flu"] finds all the incoming links to this specific page that also contain the phrase "avian flu"

[link:http://www.who.int "avian flu" inurl:h5n1] finds all the incoming links to

UNCLASSIFIED#FOR OFFICIAL USE ONLY

this specific page that also contain the phrase "avian flu" and have H5N1 anywhere in the linking page's url.

linkdomain: restricts results to all incoming links to an *entire domain*, not a specific page or url.

[linkdomain:who.int] finds all the incoming links to this entire domain

[linkdomain:who.int "avian flu"] finds all the incoming links to this entire domain that contain the phrase "avian flu"

[linkdomain:who.int "avian flu" inurl:h5n1] finds all the incoming links to *this* entire domain that contain the phrase "avian flu" and have H5N1 anywhere in the linking page's url.

Google Search:

link: restricts results to documents that link to a specific url, but excludes many inlinks as a Google policy to try to reduce webpage or keyword spamming. No other syntax or keywords can be used with the Google link: command

Live Search:

➤ **link:** restricts results to documents that link to a specific url. May be used with or without additional keywords and/or additional syntax.

[link:who.int loc:de geflügelpest] finds all the incoming links to this domain that are in the German (de) domain and contain the keyword Geflügelpest.

Link Harvester: http://www.linkhounds.com/link-harvester/

➤ Link search on steroids. This incredibly powerful tool shows where all the links come from in groups and by common Class C IP Address blocks. Link Harvester also links to the Internet Archives, two types of Whols Source data, and several variations of Google cache. Data can be exported as a CSV file and imported into a spreadsheet or database for additional analysis.

Google

http://www.google.com/

Live Search (beta)

http://www.live.com/

Yahoo

http://search.yahoo.com/

Also remember that you can use <u>Fagan Finder's URLinfo</u> page to run the link queries from all these search engines one at a time.

Fagan Finder's URLinfo

http://www.faganfinder.com/urlinfo/

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Newspapers, Magazines, Journals, Newsgroups, Blogs

There is a lot of hard news, opinion, and old-fashioned gossip on the Internet about companies and their goings-on...new products, new partnerships, contracts, lawsuits, you name it. Disgruntled employees may post to newsgroups or publish blogs. In fact, people have been fired for leaking company secrets in blogs. Obviously, the validity of this information ranges from very high to scandalously poor, but in many cases researchers are piecing together a monochrome jigsaw puzzle, so any bit of information may help.

Press Releases

These can be a great source of cutting edge information about companies. They want to make money and, if they are public, sell stock, so they love to tell the world of their successes. The two best ways to find press releases is using PRNewswire or going to the horse's mouth (the company website—look for the magic words *press releases* on the home page). PRNewswire has improved its search and I especially recommend using the advanced search option. The PRNewswire site also lets you select a specific region or country, so it's easier to find international information here than at Yahoo.

PRNewswire Yahoo Press Releases via PR Newswire

http://www.prnewswire.com/ http://biz.yahoo.com/prnews/

How to Research a Specific Country

In order to perform in-depth research on a particular country, it is a good idea to have numerous sites bookmarked as general starting places for country-specific Internet research. Some countries have a large web presence while others, for all intents and purposes, are absent or have a very limited presence. For the latter category (countries such as North Korea or Burma) you will have to rely largely on "third party" sites, that is, webpages *about* the country. Be aware that expatriates, exiles, or dissidents with large axes to grind created many of these. Third party country pages are usually most useful as sources for links to other webpages, and some are impressive in their size and complexity.

My best advice in beginning to research a specific country is to try to find the most comprehensive country metasites, and this is not always easy. Good starting places are *university area studies departments*. Many of these have country or regional

UNCLASSIFIED#FOR OFFICIAL USE ONLY

webpages with links to other metasites. For example, the University of Texas's Russia and Eastern Europe Network Information Center (REENIC) has compiled an impressive set of links for this part of the world and makes a good starting place, as do Bucknell University's Russian Studies Department, Columbia University's Middle East and Jewish Studies.

Russia and Eastern Europe Network Information Center (REENIC)

http://reenic.utexas.edu/

Bucknell University's Russian Studies http://www.departments.bucknell.edu/russian/ Middle East and Jewish Studies

http://www.columbia.edu/cu/lweb/indiv/mideast/cuvlm/

Occasionally *for-profit enterprises* will provide excellent free resources, such as *Aardvark: Asian Resources for Librarians*. Aardvark has two sections: Literature and Recommended resources and sites. I especially like the Asian Databases section, which currently lists over 650 databases. Thompson Gale published an excellent overview of Aardvark by Professor Péter Jacsó. 140

Aardvark: Asian Resources for Librarians

http://www.aardvarknet.info/user/aardvarkwelcome/

Also, do not overlook **non-profit organizations** as a resource. Friends and Partners is probably the most famous non-profit organization on the Internet devoted to improving Russian-American relations and, as a result has become a tremendous source of information about Russia—everything from current news to telecommunications to downloading Cyrillic character sets.

Friends and Partners

http://www.friends-partners.org/friends/

Try to find webpages that claim to index "all" the sites in a country. Even though that claim is undoubtedly exaggerated, these sites are still good starting places. Another type of excellent country resource is a commercially operated country website that attempts to create an overarching subject directory of a particular country or region. Every world region and most countries have such guides or directories, but the quality can range from poor to excellent. A good test is to look at categories such as "governments" or "technology." If these are sparsely populated, chances are the directory is oriented more to entertainment, travel, shopping, etc., than to research.

Péter Jacsó, "Aardvark," Péter's Digital Reference Shelf, Thompson Gale, October 2006, http://www.galegroup.com/reference/peter/aardvark.htm (26 October 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY



Al Bawaba: The Middle East Gateway

http://www.albawaba.com/

Yahoo Countries

http://dir.yahoo.com/regional/countries/index.html

If I am looking for information about a country with whose Internet presence I am unfamiliar, I often start with Yahoo's Country index. A simple alphabetic list of the nations of the world leads you to individual pages within Yahoo for each country, ranging in quality and quantity from the vast (the UK with nearly 222,000 entries) to the tiny (Kiribati with 20). Most Yahoo country pages, however, are fairly robust, with subject headings such as "business and finance," "computers and Internet," "government," "education," "news and media," etc. However, never assume the Yahoo country lists and links are thorough or even adequate representations of a country's Internet presence because they are not. Yahoo is nonetheless a very fine starting place for researching a country on the web.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Library of Congress Country Studies http://lcweb2.loc.gov/frd/cs/cshome.html

The Library of Congress website "contains online versions of books previously published in hard copy by the Federal Research Division of the Library of Congress under the Country Studies/Area Handbook Program sponsored by the US Department of the Army. Because the original intent of the Series' sponsor was to focus primarily on lesser known areas of the world or regions in which US forces might be deployed, the series is not all-inclusive. At present, 101 countries and regions are covered. Notable omissions include Canada, France, the United Kingdom, and other Western nations, as well as a number of African nations. The date of information for each country appears on the title page of each country and at the end of each section of text." Each country study presents historical, political, economic, social, and military analyses and descriptions. Users may search across all countries or any combination of countries or browse an alphabetical list of countries.

The Economist Country Briefings

http://www.economist.com/countries/

The Economist is a well-respected publication that has made a huge contribution to the Internet with its website. Specifically, the Economist Intelligence Unit has a resource known as "Country Briefings" that looks at about 60 countries and provides superb profiles, forecasts, statistics, and more about each. All the content at the site is original with the Economist and not merely a compilation of information from other sources. Some of the material is "premium content," meaning users must pay to view it. However, there is more than enough free information to make the site very valuable.

BBC Country Profiles

http://news.bbc.co.uk/1/hi/country profiles/

Similar to its counterparts at the Library of Congress and The Economist, the BBC offers Country Profiles at its website. The profiles not only include information on the history, politics, and economy of most of the world's nations, they also offer audio and video clips from the BBC's extensive archives. The main Country Profiles' page presents six drop-down menus for world regions with countries in each region listed alphabetically and one menu for International Organizations. Once users select a specific country to profile, they are offered an overview, a set of basic facts, information on current leadership, and details about the country's media. The BBC profiles are not quite as detailed as those from the Library of Congress or The Economist, but the BBC does provide a very good and user-friendly overview of most nations.

Northwestern University Library Foreign Governments

http://www.library.northwestern.edu/govpub/resource/internat/foreign.html

Northwestern University Library International Governmental Organizations

http://www.library.northwestern.edu/govpub/resource/internat/igo.html

Unrepresented Nations and Peoples Organization (UNPO) http://www.unpo.org/

UNCLASSIFIED//FOR OFFICIAL USE ONLY

The Northwestern University Library has a superb collection of documents related to governments, international organizations, and unrepresented nations and peoples. The governments' page has links to websites of countries from Afghanistan to Zambia. Even better, the page is kept up to date. Obviously, every website for every government is not listed, but the resource focuses on the major sites, such as president/prime minister, MFA, central bank, parliament, etc.

The international governmental organizations page lists intergovernmental organizations (IGOs) that have webpages. Most are in English and generally the list links only to the main page, though there are exceptions. Organizations cover a wide range of topics, from the African Development Bank to the Chemical Weapons Convention to the International Criminal Court.

Finally, the Unrepresented Nations and Peoples Organization (UNPO) page, which resides at a separate website, lists both current and former members. UNPO members comprise nations and peoples inadequately represented at the UN. Among the UNPO current members are Australian Aboriginals, Assyrians, Chechens, Iraqi Turkoman, etc. For each group, the site links to a UNPO page containing information about the people and their homeland, including geographical, economic, and historical information as well as links to UNPO documents relating to the nation/people.

Academic Info

http://www.academicinfo.net/

I am somewhat hesitant to recommend this site only because I can't find out much about who is running it (Mark Madin and a group of "volunteer subject matter experts"). However, many reputable sources link to and recommend Academic Info, including the superb University of California, Berkeley, library, which describes it as a "Rich selection of about 25,000 pages, selected as 'college and research level Internet resources' aimed at "at the undergraduate level or above." And it does contain a lot of useful information. The information appears to be fairly high-level and most of it is collected elsewhere. However, this is a good general starting place and has the advantage of having been updated frequently. Because of the currency, quality, and amount of information throughout the website, I would recommend bookmarking and using the site as a good starting place for research on a number of topics.

Google Directory Country Index

http://directory.google.com/Top/Regional/Countries/

Google Directory uses the Open Directory Project listings as a basis but adds to them and includes more links.

Admi.net http://admi.net/world/

Admi.net is a French website that attempts to be "the cyber-documentation center delivering general information about government authorities and public services." Admi.net's scope, therefore, is more limited than some others, but may thereby be

UNCLASSIFIED#FOR OFFICIAL USE ONLY

more useful under some circumstances. The site is very well organized: *General Information/Regions/*Countries. Within each subheading are the same subject headings, from "political organizations" to "law" to "companies" to "telecom" and much more in between. Obviously, much more information is indexed about some countries than others and not every heading has an entry for each country.

BUBL Country List

http://bubl.ac.uk/link/world/index.html

Carefully researched and annotated web directory by information professionals.

WWW Virtual Library Regional Studies

http://vlib.org/Regional

Links to African, Asian, Indigenous, Latin American, Middle Eastern, Russian, Eastern European, Western European, and Pacific studies. The Virtual Library was established in 1993 and is still kept up to date.

The Organization for Economic Co-operation and Development (OECD)

http://www.oecd.org/

The Organization for Economic Co-operation and Development (OECD) has launched webpages for each member country. The main features for each country website are:

- ➤ What's new: News, recent and forthcoming events or conferences, latest publications.
- Statistics: OECD statistical profile, OECD free access databases by country, and more.
- > Country and the OECD: ratification of the convention, role and activities of the delegation, financial contribution to the budget, etc.
- > Publications & documents: documentation listed in chronological order, by topic or document categories.
- Information by topic: every topic from the ageing society to transportation.
- > Don't Miss: Country profile including over 100 indicators, how to obtain OECD publications, list of translations, useful links and contact us.

What countries are included in the OECD project? All thirty member nations have country webpages and most of those country's also have a separate statistical profile page. OECD also has active relationships with 70 non-member countries. The OECD site includes numerous publications and documents, such as an Economic Survey of Russia (July 2004), Main Economic Indicators: Non-Member Countries, April 2005 (PDF), as well as information about co-operation operation between OECD member and non-member nations by country. The OECD site also permits

UNCLASSIFIED#FOR OFFICIAL USE ONLY

users to browse the site by topics, including biotechnology, corruption, energy, finance, money laundering—32 topics in all.

NationMaster

http://www.nationmaster.com/

NationMaster takes the data from the CIA World Factbook, the UN, and OECD and mines it for statistics, then makes all that information easily accessible. Also, users can easily generate maps and graphs on all kinds of statistics. One of NationMaster's nicest features is the easy ability to view profiles of individual countries, which includes many details about a country beyond just statistical data. There is no original data at NationMaster, but it is the single best source for pulling together, organizing, searching, and displaying data about every country on earth.

Finally, look for *directories and country-specific search engines within the country itself*. In my experience, the quality of such sites varies greatly, so do not give up if the first few you look at are not impressive. Once you have identified the major ISPs in a country, look to see if they have subject directories; many ISPs are creating both national and local guides for Internet users. The best will provide internal search engines and maybe even a translation feature. For example, one of Greece's major telecommunications and Internet companies, FORTHNet, offers a directory of more than 32,000 Greek websites.

FORTHNet Directory

http://dir.forthnet.gr/index-0-en.html

How do you find these sites? There is no one way of detecting the best websites in or about a country but try this: once you find a good website in or about a country, *run a "link" command* in Yahoo, Google, Live Search, and Gigablast to see webpages with links to the page of interest. You can quickly scan these for interesting-looking new websites.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

How To Research a Country

There is no sure-fire methodology for researching a specific country on the Internet. Try these sources:

- 1. University area studies departments
- 2. Non-profit organizations or foundations
- 3. Country-specific subject directories
- 4. In-country guides, portals, directories (check big ISPs)
- 5. Search engines' "link" command

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Finding Political Sites on the Web

If a country has a parliamentary form of government (and most do by the broad definition used by this organization), be sure to check out the **InterParliamentary Union**'s web page. Established in 1889, the IPU is an organization of parliaments of sovereign states. The IPE site has its own superb PARLINE database of country parliaments, which includes all relevant details about the form and structure of governments. In addition, the IPU provides links to official parliament web pages. Of the more than 200 countries listed, not all have official parliamentary web pages, but a surprising number do.

InterParliamentary Union

http://www.ipu.org/english/home.htm

Northwestern University's Foreign Governments

http://www.library.northwestern.edu/govpub/resource/internat/foreign.html

Northwestern University's International Governmental Organizations

http://www.library.northwestern.edu/govpub/resource/internat/igo.html

Also be sure to look at Northwestern University's Foreign Governments list as well as its International Governmental Organizations. Currently up to date.

Global Edge

http://globaledge.msu.edu/

The Center for International Business Education and Research at Michigan State University maintains Global Edge, a site of international country studies and business information with links to over 5000 resources. Much of the information comes from the CIA's Country Studies and the Department of State's Country Background Notes, but the site adds a tremendous amount of value to this data.

Foreign Government Resources on the Web

http://www.lib.umich.edu/govdocs/foreign.html

The University of Michigan also maintains Foreign Government Resources on the Web, a site with links to many types of foreign government sites, including official sites, embassies, constitutions, laws, etc. Last updated in July 2006.

Political Resources on the Net

http://www.politicalresources.net/

Political Resources on the Net: Unrepresented People

http://www.politicalresources.net/int6.htm

Political Database of the Americas

http://www.georgetown.edu/pdba

UNCLASSIFIED#FOR OFFICIAL USE ONLY

One of the best political websites is Political Resources on the Net. This is a superb site for locating websites for governments, political parties (mainstream and fringe), NGOs, institutes, and "unrepresented" peoples or areas, such as Kosovo, East Timor, Hmong, etc. The topics are arranged from global to regional to country-specific. For *anything* political, this is a very good first stop. However, the internal search does not work. For political information on the Americas, try Georgetown University's Political Database of the Americas; the site contains both original information as well as links to reference sources.

European Countries

http://europa.eu/abc/european_countries/index_en.htm

At the EU's Europa site, each of the European Union member states has a link from this page. The newly redesigned site has far less information about member nations' government site and now only provides a link to each country's official website. Europa still provides information about and links to EU candidate countries and to other European nations.

Council of the Baltic Sea States

http://www.cbss.st/

The Council of Baltic Sea States also offers a website with information including news, history, the structure of the Council, lots of documents, contact lists, etc. The site is completely in English, though external links to member states are usually in the appropriate national language. The homepage includes a map with links to the foreign ministries of each country.

East & Southeast Asia: An Annotated Directory of Internet Resources

http://newton.uor.edu/Departments&Programs/AsianStudiesDept/index.html

The East & Southeast Asia annotated directory of Internet resources from the University of Redlands (in California) is an excellent portal for information specific to Asian nations as well as Asian resources in general as well as "hot topics" such as the avian flu, the North Korean nuclear crisis, and the tsunami of 2004.

Current Rulers Worldwide

http://www.terra.es/personal2/monolith/

Rulers of the World

http://rulers.org/

For information on world leaders, be sure to see Zarate's Political Collections, which includes Current Rulers Worldwide. The site is updated regularly. Rulers of the World is also kept current but has an accent on history: "This site contains lists of heads of state and heads of government (and, in certain cases, de facto leaders not occupying either of those formal positions) of all countries and territories, going back to about 1700 in most cases." Current rulers are also included.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Finding Ministries of Foreign Affairs

To track down ministries of foreign affairs, you can try directories such as Google's, Yahoo's, and the Open Directory Project, or you can go directly to one of these resources that provide links to MFAs. Please keep in mind that no resource has everything. Just because the MFA you seek isn't listed by any of these sources does not mean it's not on the Web. Sometimes url guessing will work, e.g., www.mfa.gov.yu for the former Yugoslavia; sometimes it won't, e.g., www.ud.se for Sweden.

US Institute of Peace Library Foreign Affairs Ministries on the Web

http://www.usip.org/library/formin.html

The site has about 140 links to most MFAs and a few other international agencies. Current information.

Library of Congress: Portals to the World

http://www.loc.gov/rr/international/portals.html

First-rate alphabetical list of countries, with links to government sites. For most, finding the MFA is a simple click to the Government, Politics, and Law link.

Stefano Baldi's Ministries of Foreign Affairs Online

http://hostings.diplomacy.edu/baldi/mofa.htm

Hosted by Diplo Directory site, this is a superb resource, with 110 links listed.

Ministries of Foreign Affairs from Lawresearch

http://www.lawresearch.com/v10/global/ciministries.htm

Links to about 100 countries' MFA websites.

Finding Embassies

One task the Internet has certainly made much easier is locating information about embassies, including their addresses, phone/fax numbers, names of individuals associated with them, and even email addresses and, sometimes, websites. Even more fortunately, there are several websites that specialize in providing information about and links to foreign ministries on the web.

Also keep in mind that ministries of foreign affairs (MFAs) are usually in the business of providing information rather than protecting it, so they are often great resources for learning a wide range of details about a country. I strongly recommend that you always check out a country's MFA Web site for the most current information about its missions. For example, a country such as Afghanistan has

UNCLASSIFIED#FOR OFFICIAL USE ONLY

gone through many changes in the past few years both in terms of representation from and within the country. Never assume any site is absolutely current; it's best to go directly to the source when possible. Here is the Afghani MFA's list of its embassies and consulates from its website.



The following are sites helpful in finding embassies:

Embassy World

http://www.embassyworld.com/

Embassy World is a commercial site that has improved dramatically over the last few years. Today, it is probably the single best resource for finding embassies and consulates. The home page offers users the option to search for a specific country's mission in any or a single location. You can also use the alphabetical list by country to find a country's missions in other countries or other countries' missions in a specific country. I recommend searching both ways because the results are not always identical. Embassy World also includes UN Missions and an international telephone directory that varies in quality depending on the country of interest (poor for Palau, great for France...you get the idea).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Library of Congress: Portals to the World

http://www.loc.gov/rr/international/portals.html

First-class alphabetical list of countries, with links to government sites. Virtually all include a link to Embassies.

Embassies & Consulates Worldwide

http://www.mypage.bluewin.ch/caccia/

Search by country in which embassy or consulate is located. This site is current and contains other useful information about countries, generally geared toward visitors.

Yahoo Embassies and Consulates

http://dir.yahoo.com/Government/Embassies and Consulates/

An always-reliable if incomplete source; countries listed alphabetically with both embassies of and embassies in each country.

Tagish Worldwide Embassies

http://www2.tagish.co.uk/Links/embassy1b.nsf/

Tagish has hosted the European Union's Ethos site for some years and has a separate page devoted to diplomatic missions derived from the Ethos data, which now appears to be a couple of years out of date. Specifically, the Tagish site also includes a separate *list of embassy Web sites and embassy email addresses* (both sorted by either host country or country of origin).

Embassy.org

http://www.embassy.org/

Embassy.org is the source for information about any embassy in Washington, D.C.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

Research Round-up: The Best Research Tips & Techniques

Tip 1: Use the Right Tool

I am repeating Rule #1 because it is so important. The single biggest mistake researchers make is using the wrong search tool. For example, search engines are generally not useful for finding current news (use a specialized news search service). Wikis, custom search engines, and directories are generally better when researching a broad topic because they have a select group of sites (but watch out for paid placement; for a directory that contains no paid placement, use Open Directory). There are also many specialized or vertical search services that cover a huge variety of topics, everything from chemistry to message boards.

Tip 2: Search for the Most Obscure Term

I needed to find the url for an article at CNET's *News.com*. Sounds easy. I knew the date and title, but the internal CNET search tool wouldn't let me search by date and the title words returned nothing. Solution? I picked an obscure word from the article—in this case, the last name of the Microsoft executive interviewed for the article, ("wallent") and searched on that term. Bingo. Only one hit and it was the article I wanted.

Tip 3: Put the Most Important Search Term First

While it's not always true, search engines usually give more weight to the first term you list because the search software assumes it's the most important term (otherwise, why would you list it first)? Try these two queries in Google one after the other: [gardening roses] then [roses gardening]. The results are similar but not identical.

Tip 4: Search on the Singular Form First

While it is not always the case that search engines automatically search for plural forms of search terms, many (including Yahoo and Google) do. The converse, however, is not true, i.e., a search on [rose] will find *roses*, but a search on [roses] will not find *rose*. Therefore, it makes sense to search first on the singular form of a term.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

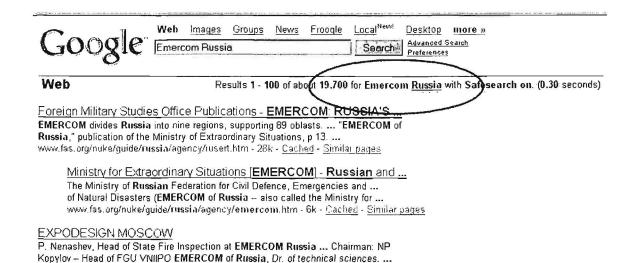
Tip 5: Use Regional Search Services, Directories, and Databases

It's so easy to fall into the trap of using US-based search engines (and, it's true, their databases are the biggest in the world). But I believe it is critical to use search engines, directories, and databases that are country- and region-specific, or language-specific. These international sites focus on collecting data from a particular part of the world and also offer more language options than US-based services. Don't just think in terms of search engines—local phone and email directories can be invaluable.

Tip 6: Search in the Native Language

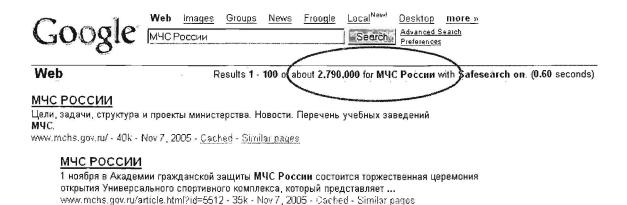
If you really want to exploit the power of an international search engine, directory, or database, you need to search in the native language. Doing so will vastly improve your chances for finding what you seek. Also, remember that many if not most international databases—for example, phone directories—only list information in the native language, so searching the online Moscow phone book for a name in English will produce zero results because all the listings are in Russian Cyrillic.

Look at the difference between searching using Google in English and in Russian [Emercom Russia] with 32K hits vs. [MЧС России] with 1.25M hits and, more importantly, the Emercom site as the top result:



www.expo-design.ru/2005/FIRE_XXI/infX_progr_eng.htm - 8k - Cached - Similar pages

UNCLASSIFIED#FOR OFFICIAL USE ONLY



<u>Инфо</u>рмационная газета "Спасатель" **МЧС России**

Аэромобильный госпиталь **МЧС России** завершил свою работу в городе ... Директор Департамента развития инфраструктуры **МЧС России** Надежда Герасимова стала ... www.spasatel.ru/ - 39k - Nov 7, 2005 - Cached - Similar pages

Tip 7: Follow Those Links

[More results from www.mchs.gov.iu]

Whenever you find a good website, always check its links. While in theory links at a web page that is indexed by a search engine should also have been indexed, the reality is often different. "Links" pages are often a gold mine of sites with similar information.

<u>Tip 8: Learn Two Words in Any Non-English Language in Which</u> You are Searching

Those two words are *search* and *links*. You need to be able to push the *search* or *find* button on a non-English web page, and you need to be able to find the *links* page.

Tip 9: Search on the LINK Field

Every time you find a good website, go to a search engine, such as Yahoo¹⁴¹, Gigablast, the new Live Search, or Google¹⁴² that permits <u>link: command searches</u> and find the pages that link to your newly discovered web page. *Any page that is really interesting to you is also interesting to others interested in the same subject.* Check the search syntax of each site: some search engines, such as Yahoo, require a full address, including the http://.

¹⁴¹ Yahoo and Live Search have two different link commands (<u>link:</u> and <u>linkdomain:</u>) that serve different purposes. Linkdomain: is a broader, more comprehensive search option.

¹⁴² The Google *link:* command is not showing all links as it once did because of Google's efforts to try to limit webspam.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Here's an interesting twist on link searching, that is, finding sites that link to a specific address. This search, which works with Yahoo and (to a lesser extent) Windows Live Search, *finds pages that link to a specific domain or domains but not to another specific domain or domains*. An example would help. Let's say I start by finding the sites that link to the Iranian Ministry of Defense. Here is the query I would use:

[linkdomain:mod.ir]

In Yahoo this query returns 545 hits. Now, suppose I want to see which sites link to both the Iranian MOD and the Iranian Electronics Industries. I can do that easily with this query:

[linkdomain:mod.ir linkdomain:ieimil.com]

However, I see lots of sites that also link to the ever-present CIA World Factbook, which, while a wonderful resource, isn't want I want. I would really like to see the sites that link to both the Iranian MOD and IEI sites but not to the CIA Factbook. Can I do this? Sure:

[linkdomain:mod.ir linkdomain:ieimil.com -linkdomain:cia.gov]

While this technique has obvious applicability for search engine optimization ("who is linking to my competitors but not linking to me?"), I think it is worth knowing about because you may come up with some creative ways to use it. Just as an interesting example, try these two queries in both Live Search and Yahoo. It's interesting to see what drops from the results' list on the second query.

[linkdomain:cia.gov linkdomain:nsa.gov]

[linkdomain:cia.gov linkdomain:nsa.gov -linkdomain:fbi.gov]

I believe you will consistently find that Yahoo! provides more results than Live Search for the linkdomain: searches. However, the results will vary, so it's worth using both search engines. Google does not offer a linkdomain: search, and its link: search has been hobbled.

Yahoo Search
Windows Live Search

http://search.yahoo.com/

http://www.live.com/

Tip 10: Look Beyond Search Engines and the Web

Search engines and directories index only a tiny portion of the Internet. With some notable exceptions, they are basically designed to index web pages. A vast amount of data is stored, for example, in online databases, many of which are free and open to the public. They often contain information useful to researchers—phone

UNCLASSIFIED#FOR OFFICIAL USE ONLY

directories, Whois databases, NIH PubMed, SEC <u>Edgar</u>, <u>Amazon's "Search Inside the Book" feature</u>, digital library collections. The <u>Domain Name System/Service (DNS)</u> itself is the largest distributed database ever created and freely accessible to any user via a simple query (<u>NSLookup</u>). Also, don't overlook mailing lists, newsgroups, and other non-web segments of the Internet.

Tip 11: Configure and Use Two Browsers

If you spend much time on the Internet, especially viewing non-US sites, you will probably encounter certain webpages that will not display at all, will not display properly, and/or will not print in the browser you're using. However, if you open the page in the other browser, it may be fine. So don't despair if a page isn't displaying or printing as it should. Chances are there are simply problems with the way the page was created and it will look fine in the other browser.

Tip 12: Try URL Guessing

It works more frequently than you would imagine. For example, I found the Iranian Ministry of Foreign Affairs by guessing www.mfa.gov.ir. And guess what the address for the Russian Ministry of Internal Affairs (MVD) is? Yes, it's www.mvd.ru. No search engine indexed either of these sites at the time I first found them.

Tip 13: Change URLs to Find "Hidden" Webpages

Sometimes a simple change inside a long url will disclose interesting pages deep within a website. For example, look at these two pages from the Federal Trade Commission:

http://www.ftc.gov/opa/2006/02/

http://www.ftc.gov/opa/2007/02/

Simply by changing the portion of the url that indicates year and month, you can view the FTC News Releases for a specific date. This is a simple example of a technique that can be used to uncover "hidden" webpages. It's especially useful on sites that update on a regular schedule, e.g., sites for press releases or news.

Tip 14: Be on the Lookout for URL Errors

Not surprisingly, many urls listed on webpages are incorrect. Among the most common mistakes are misspellings, putting a backslash (\) where a slash (/) should be, including or excluding the L in HTML, e.g.:

http://www.examlpe.com/pathname\bigmistake.html

UNCLASSIFIED//FOR OFFICIAL USE ONLY



Correct the spelling or use the tried and true method of "backward hacking" (backspace deleting) file and path names and retrying the link (you'll be amazed how often this works!).

Tip 15: Take a Look at the "Site Map"

Most large, well-organized websites have maps of what is on their web server. These site maps can be extremely useful and revealing, often showing pathways to more topics than were revealed on the home page. Also, internal search tools are notoriously unreliable; a site map may be much more useful at digging down inside a website.

Tip 16: Try Using the "Mouseover"

For non-English sites where you don't know the language, try the "mouseover" trick, i.e., move your mouse over hyperlinks. Often, the link information is in English or, if it isn't, quite often the url that appears in the toolbar at the bottom of the browser is revealing because it is likely to be written in English.

UNCLASSIFIED//FOR OFFICIAL USE ONLY

<u>Tip 17: Try Alternative Spellings, Especially of Non-English Names or Terms, or Even Purposeful Misspellings</u>

This hint especially applies to searching for non-English websites. For example, the Arabic word *khalifah* (usually written as *Caliph* in English) generally refers to the person who is the successor to Muhammed for leadership of the Islamic community. This term is transliterated as *kalifa*, *kalifah*, or some other variant. However, there is a related term *khilafah*, which is the office itself, i.e., *Caliphate* in English. This term is often associated with some of the most radical Islamic websites. I found I had to search on many variations of these terms or I would have missed a number of important sites. Even though many of the websites are in Arabic, the addresses for the sites are still written in the Latin alphabet, and searching for these key words using syntax as *inurl*:, *site*:, *domain*: revealed many useful links.

The same is true for commonly misspelled words, e.g., *http referrer* is usually misspelled *http referer* and *genealogy* is often misspelled *geneology*. A wildcard makes it possible to find both correctly and incorrectly spelled terms.

As you can see from the following example, a search on ["ministry of foregin affairs islamic republic of iran"] elicits 918 hits while a search with *foreign* spelled correctly returned only 38 results.



Did you mean: "ministry of foreign affairs of islamic republic of iran"

Ministry of Foregin Affairs of Islamic Republic of Iran

Documents Archive ...

www.mfa.gov.jr/output/english/ documents%5Contm84_200410.htm - 9k - Cached - Similar pages

Ministry of Foregin Affairs of Islamic Republic of Iran

LATEST NEWS: Thursday, 2. Dec, 2004 7:27 Local Time. ECO Secretary General Meets Dr. Kharrazi. Nov 29,2004 November 28, 2004 Visiting ... www.mfa.gov.ir/output/english/main.asp - 17k - <u>Cached</u> - <u>Similar pages</u>

Ministry of Foregin Affairs of Islamic Republic of Iran

Privious Page, Iraqi Deputy Prime Minister Meets Dr. Kharrazi Aug 29,2004. Iranian foreign minister, Dr Kamal Kharrazi on Sunday emphasized ... www.mfa.gov.iv/output/English/documents/doc4605.htm - 9k - <u>Cached</u> - <u>Similar pages</u>

Ministry of Foregin Affairs of Islamic Republic of Iran

Privious Page, Dr. Kharrazi's Statements at the Ministerial Conference of Iraq's Neighbors in Cairo Jul 21,2004. Iranian foreign minister ...

Tip 18: Always Look at a Website's Native Language Version

Usually, the native language version of a website will differ from the English version, sometimes a little, sometimes a great deal. To illustrate this point, check out the differences between the Al Jazeera Arabic and English home pages. You will see a different set of stories with different emphases on each.

UNCLASSIFIED#FOR OFFICIAL USE ONLY



Notice the completely different set of articles and even the way the pages are organized:



UNCLASSIFIED#FOR OFFICIAL USE ONLY

Tip 19: Use Wildcards to Maximize Effectiveness

I love to use wildcards in searching. Unfortunately, few search engines permit wildcard searches, which means the researcher must enter the term in a variety of forms for a thorough search. Here is how the major search engines handle wildcards.

Google: one limited wildcard (*); can only replace any single term with white space on either side (e.g., ["what a * web"] will find "what a tangled web" and "what a coiled web"). Cannot be used within or at the end of a search term (for example, to pluralize a term).

<u>Yahoo</u>: Yahoo does not support wildcard searching. The old cheat of using a "small" word, such as *a*, no longer works in Yahoo.

<u>Live Search</u>: no truncation, no wildcard. A search for [cat] finds *cat*, not *cats*.

<u>Exalead</u>: a search on [child*] return pages with *children* highlighted as a search result. The wildcard also can be <u>used inside a search term</u>, e.g., [kazak*stan]. This search accurately finds *kazakhstan*, *kazakh*, and *kazak* as well as *kazakstan*.

Tip 20: Examine Page Source Code

In addition to often revealing the webpage's language encoding, page source can provide other helpful details, including names, dates, email addresses, type of software used to create the page, etc. However, experienced webpage designers have learned by now that putting these sorts of details into the source code is an open invitation to spammers to harvest them, so finding useful bits of information in source code is less likely now than in the past. Still, many people are not experienced Internet users and have not yet learned to keep this information out of the source code, so it is worth a look. Below is a very good example of how analyzing page source code helped one company track down the person responsible for a fabricated television interview that spread potentially libelous information about the company. The page source contained an email address that ultimately led to the person responsible for the false information.

Daniel Janal, "How to Deal with Lies About Your Company (and You) on the Internet," Scambusters.org http://www.scambusters.org/Scambusters29.html

Tip 21: Ask for Help

One of the hardest questions to answer is, "how do you know when you aren't finding something because you're not searching correctly or because it's not there to find?" The best rule of thumb I can think of is to ask a more experienced Internet researcher for advice and assistance if you are hitting a brick wall. Experienced researchers generally have developed a pretty good sense of where to look for different kinds of information and, most importantly, the types of information you are not likely to find on the Internet.

432

Researching & Understanding the Internet

A Plain English Guide to Internetworking

The Internet, were it a person, would be something of a narcissist. It probably has more information about itself than any other single topic, which means the best place to research the Internet is the Internet. However, understanding how the Internet actually works can be frustrating and most books, websites, and other resources that purport to explain the inner working of the Internet exacerbate rather than enhance knowledge because they make assumptions about their readers that are often wildly inaccurate.

There are generally two types of guides on how the Internet works: one is the technical guide for network engineers and system administrators; the other type is a users' guide, a how-to book for people who want to use the Internet but don't really need to know anything about what's under the hood. The first type of guide is fine if you are an engineer, but these manuals or websites are written by and for technical people, meaning you already have to know their jargon in order to understand what they are saying. The second type of guide merely avoids the problem all together by ignoring technical explanations. After all, you don't need to know how an internal combustion engine works to drive a car.

I have attempted to create a little niche in between the extremes by offering a non-technical explanation of some of the basic concepts behind how the Internet gets traffic from one point to another and does it with remarkable speed, accuracy, and efficiency. For many of us who are not technically inclined, the answer has often seemed to be "magic happens here." But more and more often in our jobs we find we need to know a little bit about the mysterious inner workings of the Internet. I have no pretensions that this overview is complete. I knowingly gloss over some subtleties that I think are simply too arcane for a high-level look at Internetworking.

The Internet

The Internet isn't a single network; it is a network of networks. However, all these many networks are not directly connected to each other and, in fact, do not "talk" or exchange data directly with each other. Rather, the Internet employs an elaborate system of rules (known as protocols), numbering schemes (e.g., autonomous systems, IP addresses), hardware (e.g., routers and servers), and arrangements (e.g., peering agreements) to ensure that your request to view a website in Pakistan or send an email to Seoul is fulfilled almost in the blink of an eye.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Packets

At the (almost) simplest level of the Internet are things called packets. For our purposes, a packet is the fundamental unit of data that is sent between an origin and a destination on the Internet. How does a packet get from Point A to Point B? First, it needs two addresses: one for its origin and one for its destination. The origin address is known as the source IP address and the end address is known as the destination IP address.

IP Addresses

Every device attached to the Internet must have an address. In IPv4, ¹⁴³ these addresses are written as "dotted quads," i.e., four sets of numbers separated by periods, as in 204.180.95.2. ¹⁴⁴ Today, IP addresses are much more likely to identify the public end point of a path than an individual computer. The last IP address you see is likely the point at which the packet enters a private network, and you can't normally see beyond that point.

IP addresses are similar to telephone numbers or house addresses, but that analogy only goes so far because many IP addresses are temporary. Temporary IP addresses are called dynamic IP addresses; the IP address your cable or DSL provider assigns to you is probably dynamic. On the other hand, certain types of devices on the Internet are usually assigned permanent or static IP addresses. One of these types of devices is a router.

This brings us back to our traveling packet. Once we know that the packet has its own address and a destination address, we then need to know how the packet gets from one place to another. It gets there because of something called a router.

Routers

A router is a device that forwards packets from one network to another. Routers used to be called gateways and you may still encounter that term. So, how does routing work?

Routing

Think of routing as the process in which your computer (host) tries to find a destination computer (host). Basically, your source IP address host with data to send asks your ISP's router two very important questions:

¹⁴³ IPv4 is Internet Protocol version 4, the current Internet Protocol or set of rules for exchanging data across the Internet. It is in the process of being replaced by IP version 6. IPv4 uses 32-bit addresses (allowing 4.3 billion addresses); IPv6 uses 128 bits (allowing 3.4x1038 addresses).

¹⁴⁴ This type of IP address is an IPv4 address; the next generation Internet will use IPv6.

UNCLASSIFIED#FOR OFFICIAL USE ONLY

- A) how do I get to my desired destination address? and
- B) would you please take me there?

On the Internet, it is highly likely that there is more than one path from computer A to computer B. So the router has the job not only of knowing where to send the packet but deciding which path or route the packet should take. But how do all these routers on the Internet know which path to use to send packets so they get to their destinations?

Routers use routing tables to determine where a packet is going and how to send it. The Internet in part consists of a huge grid of routers, each of which only knows a piece of the whole picture, not the entire Internet (it's too big for that!).

Look at this image. You can see that in this simple example, Router A does not know Router B's connections. It literally cannot see Router B. The hub or switch at 10.0.0.0 knows both Router A and Router B's destination paths and can, therefore, route packets from 20.0.0.0 to 40.0.0.0 and back. All Router A needs to know is to send those packets via its Port 1 to 10.0.0.0 and, from there, Router A has no more to do with the packets. The reverse is true of Router B. 145

Router A **Routing Table** hub or To go to Route via network: port #: 10.0.0.0 10.0.0.0 20.0.0.0 30.0.0.0 40.0.0.0 Port 1 Port 1 Router A Router B Port 2 Port 3 Port 2

From Computer Desktop Encyclopedia @ 1998 The Computer Language Co. Inc.

20.0.0.0

30.0.0.0

40.0.0.0

¹⁴⁵ "Router." Answers.com. Computer Desktop Encyclopedia, Computer Language Company Inc., 2005. http://www.answers.com/topic/router (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

There is a further layer to the addressing and routing of Internet traffic that we need to know about. As the Internet grew, it quickly became apparent that it would be impossible for every network to know every other networks' "gateways," as they used to be known. This works only a small scale, and the Internet is anything but small. Thus, the idea of autonomous systems was introduced, by which only those devices within a specific AS would know all of that AS's routes.

Autonomous System

An autonomous system (AS) is a network or group of IP networks run by a single entity using common routing policies or protocols. Another way of looking at an AS is to think of it as a lot of networks being handled on the Internet as one logical domain. There are two types of Autonomous System numbers: Public AS numbers and Private AS numbers. A public AS has a unique number (ASN) associated with it; this number is used in both the exchange of exterior routing information (between neighboring Autonomous Systems), and as an identifier of the AS itself. The introduction of autonomous systems also meant that AS's could run their own internal network any way they pleased, which made sysadmins happy. We are only interested in Public ASNs.

In order to keep track of these new AS's, each AS was assigned an AS number (a 16 bit integer) by IANA, the Internet Authority for Assigned Names and Addresses. IANA assigns both IP addresses and ASNs to Regional Internet Registries (the three biggest are ARIN, APNIC, and RIPE), who then assign both types of numbers to customers. 146

The rules for how traffic is handed off among and between AS's on the Internet are called protocols, and there are a number of them currently in use. The core external routing protocol for the Internet is the Border Gateway Protocol (BGP), which we'll look at next.

Border Gateway Protocol (BGP)

BGP works by maintaining a table of IP networks or 'prefixes' that designate network accessibility between autonomous systems (AS).

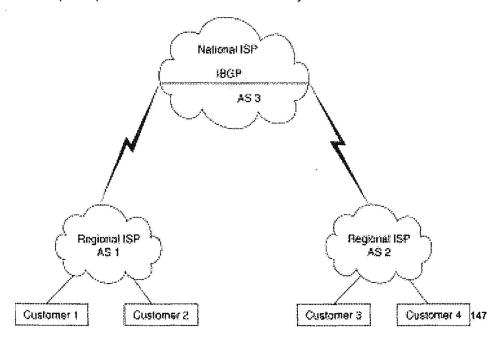
In this example, imagine Customer 1 in Baltimore wants to reach a website maintained by an ISP in Los Angeles, represented here by Customer 4 at Regional ISP (AS2). Customer 1's Regional ISP (AS1) routes the request through National

¹⁴⁶ Elquapo's Guide to Routing—Part 3, BPG, 20 August 2003,

http://www.kuro5hin.org/story/2003/8/19/18378/4228 (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

ISP (AS3) in Chicago because there is no direct connection between Regional ISP (AS1) in Baltimore and Regional ISP (AS2) in Los Angeles and going through multiple third party networks would be both slow and costly. Does this look familiar? It is the same principle that was at work in the way routers handle IP addresses.



Think of the routing of Internet packets as analogous to airline travel. Imagine you need to fly from Dallas to Damascus, Syria. There is no direct flight, so you must first fly to Washington Dulles Airport, change planes and carriers, then fly to Nicosia, Cyprus. Once there, you must again change planes and carriers to get to Damascus.

Internet traffic moves in much the same way. The airline is analogous to an Autonomous System. Each AS has a limited number of other AS's with which they exchange data, just as American Airlines does not have a direct connection with Syrian Airlines. Similarly, Dulles Airport is analogous to MAE-East, a major Internet Exchange Point along the East Coast of the US.

Internet Exchange Points

Internet Exchange Points are physical locations where different networks exchange traffic, much like large airport hubs such as Washington's Dulles or London's

^{147 &}quot;Border Gateway Protocol," Cisco Documentation, Cisco.com,

 (14 November 2006).

UNCLASSIFIED#FOR OFFICIAL USE ONLY

Heathrow airports. Internet Exchanges (IX) are also known less frequently as Network Access Points or NAPs.

The IX system is extremely important because it greatly facilitates the movement of Internet traffic both within a country and around the world. IXs permit traffic to be handed off directly between two different Autonomous Systems without incurring a cost instead of having to travel through third party networks, which is usually slower and more expensive. Over the past decade, there has been a marked increase in the number of IXs in the world. It used to be the case that in order for Internet traffic to travel within a country, it had to be routed to an Internet Exchange on another continent because there was no way for a network in one city to get its traffic directly to a network in another city in the same country, at least without paying a rather steep price to use a third party network.

Internet Exchanges principally consist of a large number of **switches**, network devices that select a path for sending data to its next stop. In general, a switch is simpler than a router because it does not require as much information about the network and routing policies.

Some IXs are relatively small while others, such as MAE-East and MAE-West in the US, may have many providers as members and many different locations. At the heart of this system, in which traffic is voluntarily handed off or exchanged among Autonomous Systems, is something called peering.

Peering

"Peering is the practice of voluntarily interconnecting distinctly separate data networks on the Internet, for the purposes of exchanging traffic between the customers of the peered networks...Peering usually indicates that neither party pays the other for the traffic being exchanged. There are, of course, examples of paid peering.

The act of peering typically involves the following elements:

- > The physical interconnection of the networks involved.
- > The exchange of routing information, through the BGP routing protocol.
- Commercial and contractual peering contracts or agreements". 148

Why is peering so important? One of the original concepts behind the creation of the Internet was the concept of global reachability, that is, every point on the Internet is

^{148 &}quot;Peering," Wikipedia,

http://en.wikipedia.org/wiki/Internet exchange#Exchange Points and Colocation Facilities (14 November 2006).

UNCLASSIFIED//FOR OFFICIAL USE ONLY

in theory reachable by and from any other point on the Internet. But at the same time, every network most certainly is not connected directly to every other network on the Internet, so there had to be accommodations made to ensure that traffic could flow freely around the globe. This mutually beneficial arrangement to exchange Internet data freely between two networks is a peering agreement. I'll accept and forward your traffic if you'll accept and forward mine.

CIDR

One final concept needs to be mentioned. About a decade ago people realized that something had to be done about Internet addressing. There were real concerns that we would run out of IP addresses and run out of capacity in global routing tables, with new networks being connected to the Internet at the rate of about one every thirty minutes. The solution was something called CIDR (pronounced "cider").

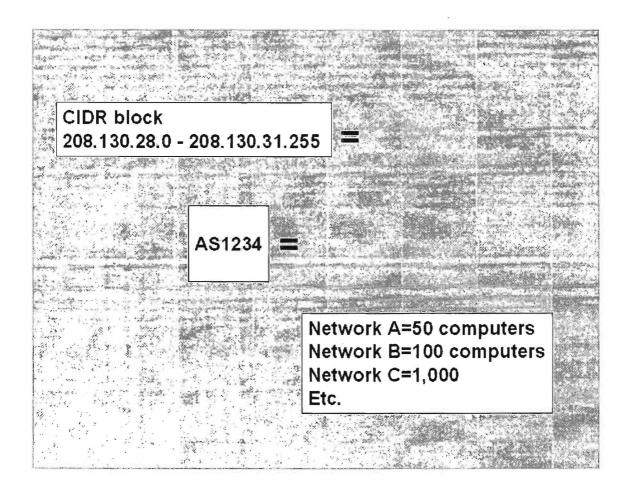
CIDR "CIDR stands for Classless Inter-Domain Routing...CIDR is an effective method to stem the tide of IP address allocation as well as routing table overflow. Without CIDR having been implemented in 1994 & 1995, the Internet would not be functioning today. Basically, CIDR eliminates the concept of class A, B, and C networks and replaces this with a generalized 'IP prefix'." CIDR permits blocks of addresses to be assigned to networks as small as 32 hosts or as large as those with over a half a million hosts. This type of address assignment more accurately reflects an organization's actual size and specific network requirements.

What is the relationship between CIDR and the Autonomous System? Perhaps the easiest way to understand their relationship is to think of "an Autonomous System as a collection of CIDR IP address prefixes under common technical management. For example, the CIDR block from 208.130.28.0 to 208.130.31.255, as well as network 201.64.75.0 might be AS2934." In turn, AS2934 may represent one network made up of 50 computers or fifty networks made up of thousands of computers.

¹⁴⁹ The CIDR FAQ, version 7.0, May 2004, http://www.interall.co.il/cidr.html (14 November 2006).

¹⁵⁰ "Autonomous System," Connected: An Internet Encyclopedia, Third Edition, April 1997,

UNCLASSIFIED#FOR OFFICIAL USE ONLY



UNCLASSIFIED#FOR OFFICIAL USE ONLY

Researching Internet Statistics

Typical research questions about the Internet concern its size, numbers of websites, numbers of users, where they are, who they are, what ISPs serve a particular city, etc. This information is available on the Internet and knowing a few of the best basic sources of Internet-related data will give you a head start in finding these facts fast.

Network Wizards Domain Survey

http://www.isc.org/ds

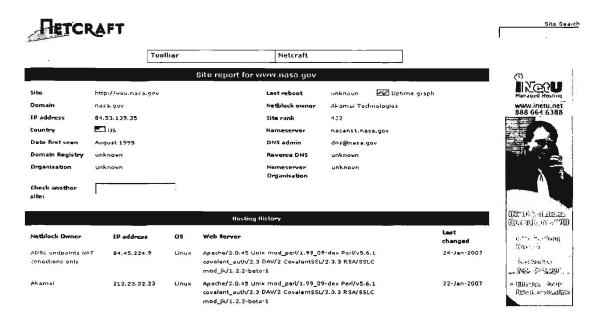
One of the best sites that tracks the growth of the Internet is the Network Wizards Internet Domain Survey. Network Wizards has been surveying the Internet since 1987 and provides very straightforward data about numbers of hosts on the Internet and their distribution by top-level domain. Thus, researchers can quickly find out the .INFO domain went from 2128 hosts in Jan 2002 to 76692 in the July 2006 survey. Network Wizards also provides a very good explanation of how they conduct their surveys. They make the important point that their numbers are "fairly good estimates of the *minimum* size of the Internet." They have no way of telling how many hosts or domains they did not locate.

Netcraft

http://news.netcraft.com/

Another Internet statistics site of long-standing is Netcraft, best known for its "what's that site running" data. Netcraft keeps track of a wide range of data about the Internet, including "market share of web servers, operating systems, hosting providers, ISPs, encrypted transactions, electronic commerce, scripting languages and content technologies on the Internet." Netcraft is probably the top site for information about Internet technology usage.

UNCLASSIFIED#FOR OFFICIAL USE ONLY



ClickZ Stats

http://www.clickz.com/stats/

Jupiter Media bought NUA, the Network Users Association, a European Internet consultancy that tracked many types of Internet trends and surveys around the world, and combined it with its former competitor. The ClickZ Stats sites provides detailed information about Internet usage in many sectors and topics: demographics, geographics, wireless, government, finance, hardware, etc. The statistics tend to be consumer-oriented (e.g., DVR usage) but there are a number of technology surveys as well, e.g., the global VoIP market.

More good sources of statistical and usage information

Global Reach's Global Internet Statistics by Language

http://www.glreach.com/globstats/

Internet Traffic Report

http://www.internettrafficreport.com/main.htm

Zooknic Internet Statistics

http://www.zooknic.com/