

Web search for local communities in the Highlands of Scotland: A self-tutoring guide

MODULE III

Alternatives to Google: some other search tools worth a try

Web search for local communities in the Highlands of Scotland: A self-tutoring guide

MODULE I

How to get the most out of Google Web search

MODULE II

A concise guide to Google products, services, applications, and other offerings

MODULE III

**Alternatives to Google:
some other search tools worth a try**

MODULE IV

The best of the Web: a guide to some of the most information-rich resources on the Internet

Introduction

Google is a marvellous Web search tool and is as good as they get at present, but it is certainly not the only one. Other top search engines include Ask.com (aka as Ask Jeeves), Bing (formerly called MSN Search), and Yahoo! (and see → **General purpose, product, and visual search engines** below). According to data published by Experian Hitwise <http://www.hitwise.com/us/datacenter/main/dashboard-23984.html> in June 2011, Google still heavily dominates the market with a share of about 68%, while the market share of Yahoo and Microsoft's Bing currently is something just under 14% for both; Ask.com is in fourth place with around 2.6%, and AOL Search in fifth place with about 1.4%. The picture is roughly the same if ranked by number of visits, although Bing does better than Yahoo in this category.

A whole new generation of much-hyped search engines and other search tools have appeared in recent years trying to compete with the search giants, many of them being touted as "next-generation" search engines and offered as a possible alternative to Google. Some were clever and inventive, while others used imaginative visual display interfaces to show search results. But very few have been successful competing with the likes of Google and many of the new ventures (Cuil, Daypop, Grokker, Kartoo, SearchCrystal, SearchMash, Teoma, Ujiko among them) were short-lived and never posed a real threat to Google; or if they did achieve a measure of success and popularity their operations were swiftly gobbled up by giant media corporations.

In addition to general purpose and product search engines there are now all sorts of speciality search engines, including meta search engines, visual search engines, question & answer search engines, a plethora of so-called vertical search engines, real-time search engines, search engines for social networking sites, or search tools for blogs, books, forums, song lyrics, maps, medical information, finding people on the Web, and more. A number of these are briefly described and appraised in the sections set out below. All websites in this module were accessed and verified during the course of May and June 2011.

Contents

General purpose, product, and visual search engines	5
Meta search engines	9
Real-time search engines	11
Question & answer search engines	14
Vertical and blog search engines	17
People search engines	24

General purpose, product, and visual search engines

This section briefly describes the major search engines other than Google, as well as reviewing some interesting newcomers to the Web search arena.

Also included are a number of visual search engines. Visual search engines search for information on the Web and display the search results not as a series of text descriptions accompanied by links but in a more visual format, using a variety of techniques and approaches. This can be useful for people who prefer images to text, or for situations where you need to display the terms in your query in a more visual manner, for example for travel or tourist information. There are now a number of quite exciting new search tools and some of these are described below.

A9.com <http://a9.com/>

An innovative product search engine that says it is “good at finding goods”. Search and browse for millions of products offered by the top online shopping destinations and the world’s leading e-commerce sites, including Amazon sites worldwide. It says it “processes billions of queries for hundreds of millions of products on Amazon sites in many parts of the world.” Powered by Google, A9 lets you not only undertake Web searches, but also looks for book results from the Amazon.com Search Inside the Book™, which opens in a window alongside the Web results. Its sophisticated technology aims to let users find the exact book they want, “even if all they can remember is a character name, a sentence they can't place, or an obscure phrase they can't forget.”



A9.com. An innovative product search engine

Already available for iPhone™ and Google’s Android mobile phone, an additional feature is A9.com’s visual search, using pictures to find products. With its visual product matching service that powers both the iPhone SnapTell Application <http://snaptell.com>, and Amazon Remembers (the Amazon App for iPhone and iPod Touch), customers can use an everyday camera phone to access product reviews, prices and features, easily and instantly. Using sophisticated image recognition technology it helps shoppers find information about the products they see and for which they would like to compare prices. The SnapTell image recognition translates a picture into searchable data points that are matched in a database of images to find and return information about the product. Snap a picture, or scan the barcode of nearly any book, CD, DVD, or video game sold in the US or UK, and within seconds get links and product reviews to Amazon, Google, YouTube, Wikipedia, and others.

About.com Websearch <http://websearch.about.com/>

Now owned by the *New York Times*, About.com aims to help users solve the large and small yet diverse needs of everyday life published in the form of a series of helpful guides (albeit with a fairly strong focus on the US). Its database contains 2.8 million handcrafted, original articles on 70,000 topics, written and updated by a network of nearly 800 expert guides.

AOL Search <http://search.aol.com/aol/webhome>

AOL's main Web search results are actually powered by Google and its audio and video search results <http://video.aol.com/> are powered by Singing Fish, one of the earliest search engines dedicated to multimedia content, acquired by AOL in 2003 and subsequently integrated into its search engine. AOL has some attractive ease-of-use features, and search results are presented in an accessible and user-friendly manner.

Ask.com aka Ask Jeeves <http://uk.ask.com/>

Founded as Ask Jeeves in 1996 and renamed Ask.com in February 2005, this is a popular question and answers search engine that uses technology designed to respond to questions and phrases (or single word searches), with questions posed in everyday, natural language (as well as traditional keyword searching). For example, questions such as "why is the sky blue", "how do i remove curry stains from clothing", "where can i find cheap flights to Australia", "why does swiss cheese have holes in it", "how do i repair a leaking toilet cistern", "when do midges bite", "is beetroot good for you", "where does snot come from", etc. As you begin to type your question you can see what other people have asked, and sometimes the results to a phrase search like this can be remarkably good and may help you to get the information you need quickly and easily. There is also an **Ask Jeeves for Kids version** at <http://www.askkids.com/>.

Bing <http://www.bing.com/> (formerly MSN Search)/

Bing Visual Search <http://www.bing.com/browse#toc=0>

Launched in June 2009 and formerly called MSN Search (and previous to that Live Search, or Windows Live Search) this is the Web search engine from Microsoft, "visually organized to help you make more informed decisions". Bing is now also fuelling → **Yahoo** search results in the US and Canada, and its market share in the US is about 12% or about 3.5% globally. A related product from the Microsoft stable, Bing Visual Search, offers visual search in a range of categories/sub-categories and through a number of themed galleries.

Duck Duck Go <http://duckduckgo.com/>

The quaintly-named Duck Duck Go claims to be a new search engine "with less garbage" and no advertisements, plus "zero-click info, more privacy, less spam, and more". It places a strong emphasis on privacy and safety, and if you go to settings at <https://duckduckgo.com/settings.html> you can select results and privacy settings, and as well as interface and "look and feel" settings. For Simpsons fans there are "goodies" such a complete A-Z list of all Simpsons characters <http://duckduckgo.com/?q=Simpsons+characters>, with a little profile for each, and with links to Web pages devoted to any of the characters. Overall Duck Duck Go is quite fun to use, but its search capabilities are no match for Google.

Exalead <http://www.exalead.com/search/>

Exalead is a European search engine and one of a number of search engines that offer you a preview of search results by displaying a thumbnail image of the Web sites found for each search, a kind of *visual* search engine. This can be quite fun, although the search results may be not as good as a Google search. Try it for "gairloch", "haggis", or "shinty" for example, and you can also do a search for images or videos. For Web searches, you can choose among three different result displays: text only, thumbnail preview images with short descriptions, or thumbnails with extended descriptions. A title and description is displayed for each result. Click the title to visit the page or view the file. You can also just click on a thumbnail image to preview a Web page or file without leaving the Exalead search engine.

Gigablast <http://www.gigablast.com/>

Touts itself as "The Green Search Engine" and has a fairly substantial index, in addition to providing a directory search within the human-edited Open Directory Project (but which has been in decline and is not very current nowadays). A useful feature is something called "GigaBits", enabling you to refine or broaden your search based upon related topics from search results, try it for "inverness" for example. The "GigaBits" appear on top of the search results; click on 'More' to see additional related topics or search terms.

ISseek <http://www.iseek.com/iseek/home.page>

This is an interesting clustering engine using technology that enables it to dynamically and instantly target results that it believes are most relevant. It analyzes your search results each time you do a query, identifying themes, topics, authors, and more, and organizes the results for you. These are also

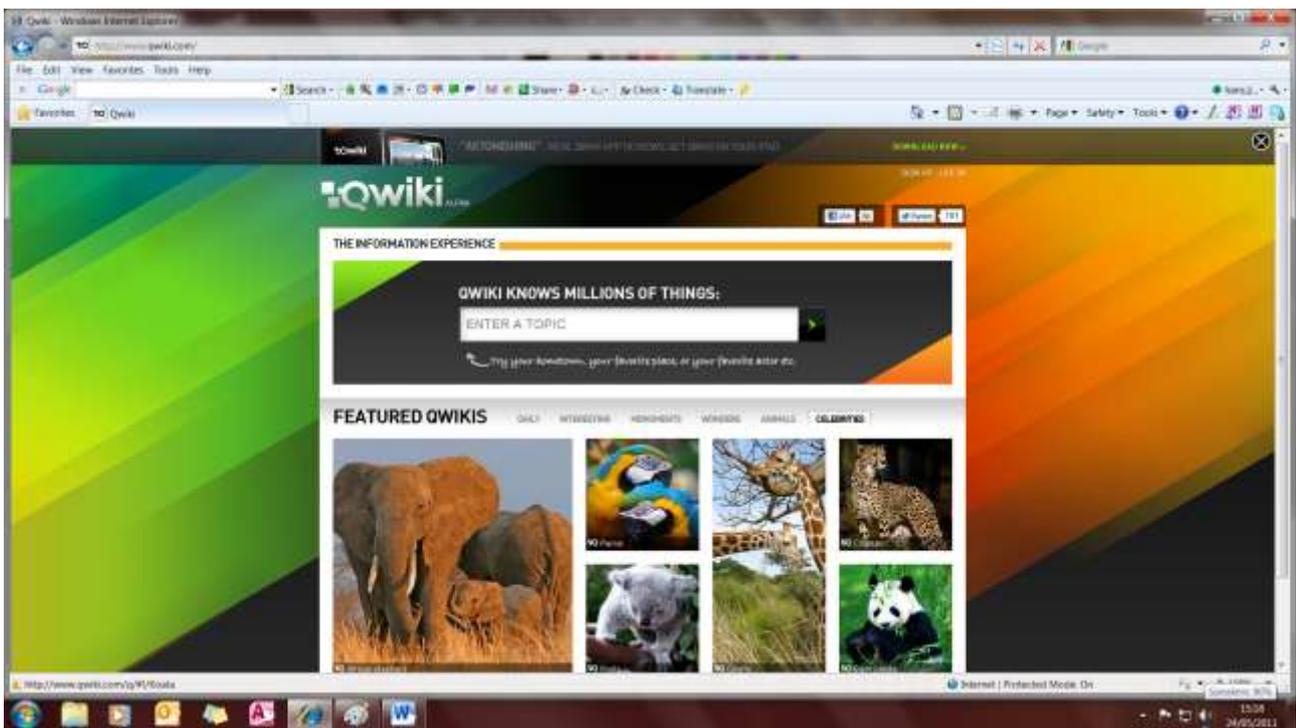
displayed as cluster results by topics on the left hand side of your results screen. Instead of relying on keyword pattern matching, you can use natural language to do your searches, and ISeek says it “knows the meaning and concepts embedded in the questions you ask, so it understands what you are asking and retrieves relevant information.” A separate section, ISeek Medical Resources, gathers results from peer-reviewed sources, government organizations, and other American Academy of Continuing Medical Education (AACME) accredited content providers. Not a serious competitor to Google, but worth a try.

oScope <http://www.oskope.com/>

Set up by a media company in Switzerland oScope calls itself a “visual search assistant” and lets you graphically browse and organize items from Amazon, Ebay, Flickr, Fotolia, Yahoo!Image Search and YouTube. You first select one of these services on the menu, and you can then fine-tune your search by selecting from a range of categories and sub-categories, or you can search for specific items. Results will appear as small thumbnail images and more information will appear when hovering over the image with your mouse. Click on any image to play a video (where available) or enlarge it, and which will also reveal a link to the original page. Selected items can be saved by dragging them into a special folder on your screen. You can switch between views and display modes: grid, stack, pile, list, and graph. You will need to register to use the service, but it is free. oScope offers quite an inventive approach for visual search, although I found the search facilities somewhat flawed.

Qwiki <http://www.qwiki.com/>

Among the latest crop of visual search engines and still in Alpha, Qwiki is really quite impressive, attractively presented with high-quality photographs, and is interesting in as far as it also provides an *audio* commentary – quite reasonably informed and up-to-date for the most part (and apparently based largely on Wikipedia entries) – for search results, together with visual links to related topics, or locations, etc. as well as links to Wikipedia, Fotopedia, Google and You Tube. It is quite good for places, personalities, or basic factual, historical, geographical or biographical information, or information about



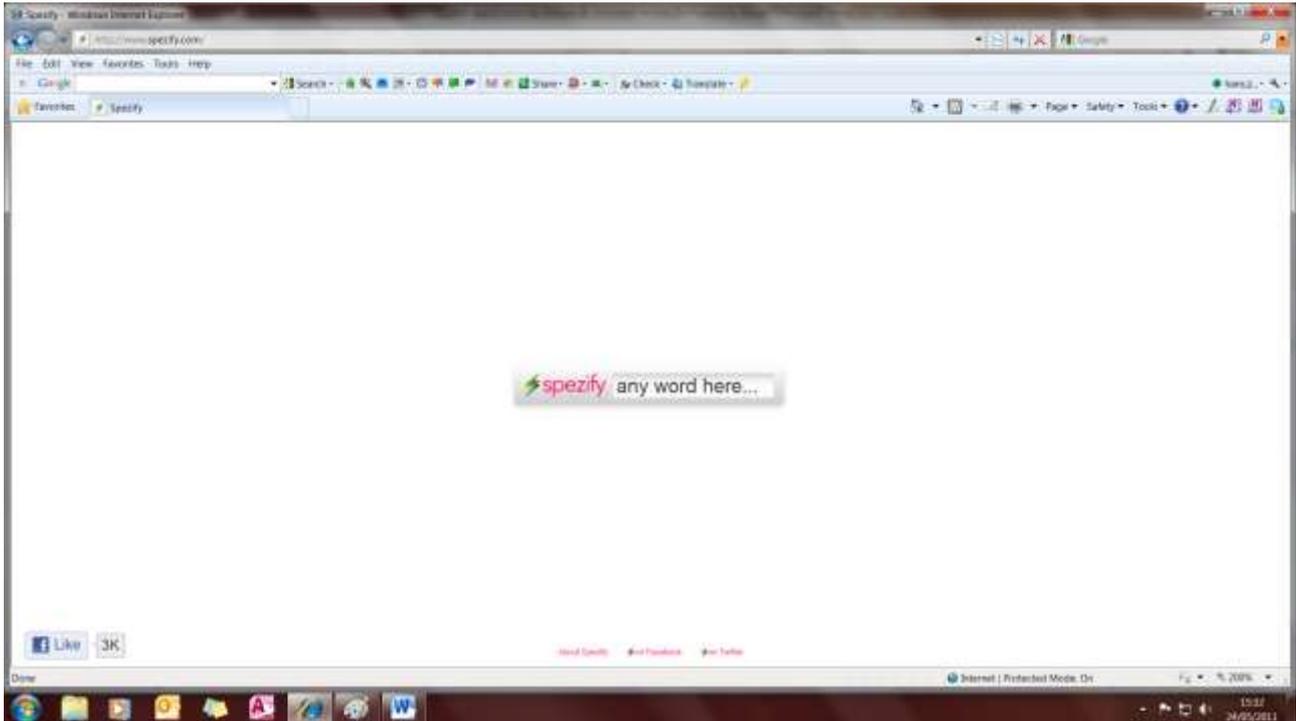
Qwiki. An impressive new visual search engine

animals, the natural world, and so on. Try it for “shieldaig”, “wester ross”, “haggis”, “shinty”, or “ullapool”, for example. It has even got a visual and audio entry for the Raigmore Hospital in Inverness, or the Plockton railway station. Some people may find the flat, computer-generated audio voice a bit irritating (and there are a small number of mispronouncements here and there), but one rather more irritating aspect is the fact that some links to related topics currently come up with a “sorry we couldn’t find a Qwiki by that name” message.

Spezify <http://www.spezify.com/>

Stockholm-based Spezify is an interesting new search tool presenting results from a large number of websites in different visual ways. It displays the results as a kind of visual collage of thumbnails of screen shots, images, and sound and videos, spread out on your screen. It pulls in results from Amazon, Digg, Yahoo, Twitter, Flickr, YouTube and other search sources but it is possible to filter the type of content

displayed. You control what you see via icons in the top-right corner. Spezify is bright and colourful and quite good for social and multimedia content, but less impressive for more serious topics and research, when results can be very hit and miss. Try it out for "applecross" or "amy macdonald" for example.



Spezify. Presents results from a large number of websites in different visual ways

TagGalaxy <http://taggalaxy.de/>

Created by Steven Wood, a freelance German graphic designer, TagGalaxy is a stunning Flash application (you will need the latest version of the Adobe Flash Player) and visual search engine that lets you browse through Flickr photos in 3D via a virtual planetary system, an explorable 'galaxy' of Flickr tags in a space-like setting. Type in a search term or 'tag' and the main term will then appear at the centre of your search 'galaxy', surrounded by a cluster of planets showing related themes and topics. You then click on the planet of pictures and it twirls around with your results enlarged as you select them. Try it, for example, for "scotland highlands", or expand this to "scotland highlands mountains landscape" and/or add any more related terms; or do the same for "applecross" for example. After assembling the results TagGalaxy automatically contacts Flickr and will show you how many photos there are on Flickr fitting the search topic and/or sub-topics (in the case of Applecross more than 6,000 images), each with a picture description and the name of the photographer, together with a link to the actual Flickr page. This is an exciting and clever visual search engine.

Yahoo! <http://uk.yahoo.com/> or <http://uk.search.yahoo.com/>

Yahoo can be a good alternative choice to Google from time to time, and sometimes it can track down sites or resources which Google doesn't necessarily pick up. Yahoo claims to be "the world's most visited home page" and that may be true as a portal, but not in terms of its Web search engine query volume, where it is still miles behind Google. Yahoo and Microsoft, in an attempt to take on chief rival Google, recently signed a multi-million pound search deal whereby Microsoft's Bing search engine will power the Yahoo website and Yahoo will in turn become the advertising sales team for Microsoft's online offerings. **Yahooligans**, its search engine for children and teenagers, can be found at <http://kids.yahoo.com/>.

Meta search engines

Meta search engines (also called meta crawlers) are those which allow you to enter keywords in the search box and your search query is then simultaneously sent to several of the most important search engines. While each search engine has its own method of searching and each will return different results, by using a meta search engine, instead of getting results from one search engine, you will be getting the best combined results from a variety of engines and/or directories, and you can then compare results.

Meta search engines operate on the premise that the Web is far too large for anyone search engine to be able to index it all, and that more comprehensive search results can be obtained by combining the results from several search engines. However, there is also a downside to meta search engines. As most only retrieve (or display) the top 10-50 results from each search engine, the total number of results retrieved could well be considerably less than found by doing a direct search on one of the top search engines, e.g. Google.

Here is a selection of meta search engines:

Dogpile <http://www.dogpile.com/>

This is arguably the best meta search engine currently available. Using Dogpile means that rather than getting results from one search engine, it will display results from several of the leading search engines, Google, Yahoo! Search, Bing, and Ask.com, as well as social media sites such as Kosmix and movie sites such as Fandango. Dogpile looks at all of them, decides which are most relevant to your search, eliminates duplicates and reveals them to you. This can sometime provide a better Web search experience. Before you start, read its "Dogpile Tricks of the Trade". Dogpile also offers quite good video, images, and news meta searches.



Dogpile. Arguably the best meta search engine currently available

Ixquick (also known as **Startpage.com**) <http://eu.ixquick.com/uk/>

A meta search engine that touts itself as "the world's most private search engine". It is a proxy service that allows users to browse websites anonymously in complete privacy, without passing on any private, personally identifiable information to the websites they view. When you perform a search, you will find a clickable 'Proxy' option below each search result. When this option is selected, Ixquick acts as an intermediary to retrieve the page and display it in a privacy-protected Ixquick window. As a result, this proxy option offers complete anonymity since the user never makes direct contact with the third-party website, and the user's IP address is invisible to the viewed website. In addition, the website cannot see or place cookies on the user's browser. However, one drawback is that pages can load more slowly since

Ixquick must first retrieve the contents and then redisplay them. An Ixquick result is awarded one star for every search engine that chooses it as one of the ten best results for your search. So a five star result means that five search engines agreed on the result (although, in practice, this seems to happen only rarely). If you are greatly concerned about privacy Ixquick might be an option, but results may be highly incomplete and not very satisfactory as Ixquick only searches → **Blekk**, EntireWeb, → **Exalead UK**, → **Gigablast**, the Wikipedia, and the Open Directory.

Mamma <http://www.mamma.com/>

Originally launched as the first meta search engine in 1996, Mamma bills itself as “the mother of all search engines”. Mamma retrieves results from a dozen or so major search engines and directories, but it is nowhere near as good as → **Dogpile** above, and the numerous “sponsored” (i.e. paid for) listings in the string of search results are a huge irritant.

Polycola <http://polycola.com/> (formerly GahooYoogle.com)

Lets you search Google and Yahoo at the same time displaying the results side by side; or you can choose to search for side-by-side results from any two of seven search engines including Ask, Altavista, AOL, Live, as well as the Dogpile search engine above. However, sometimes results can be fairly unsatisfactory.

Turboscout <http://www.turboscout.com/index.php>

A search engine comparison tool that may be able to save you time as you only enter keywords once, then getting and comparing original results from over 90 search engines across seven categories: Web, Images, Reference, News, Products, Blogs and Audio/Video. For example, using Web search, after you hit the search button, the first result is for Google, and a panel on the top of the page will then give you



TurboScout. A useful search engine comparison tool

links for results from 20 other general purposes search engines, while for a News search it will display results not only from Google, Yahoo, etc., but also those from major news sites such as the BBC, CNN, and some news search engines. It doesn't replace Google, and it looks a bit faded, but it is worth a try because it can quickly tell you how good, or how bad, other search engines are in finding relevant results.

Yippy <http://search.yippy.com/> (formerly Clusty)

Uses clustering technology to organize matches in hierarchical folders and categorizes search results; i.e. it shows search results grouping similar items together. This was an innovative concept and was helpful if you were not entirely sure of what you're after, and/or you were searching for a fairly broad term to start off with, but then eventually wanted to fine tune the search. Unfortunately the new Yippy – who acquired Clusty in May 2010, and which now aims to be a secure and family oriented Web browser and search engine – has a draconian privacy as well as a censorship policy <http://search.yippy.com/censorship>. Yippy says it may censor any search results, Web domains or IP addresses it doesn't like or considers

inappropriate, not just pornographic sites or those not suitable for children (which can be filtered out by most search engines in any event), but also those containing "Anti-Semitic views or opinions", "Anti-Christian views or opinions", "Anti-Conservative views or opinions" or "Anti-Sovereign USA views or opinions". Does this mean, then, that those offering anti-Socialist or anti-Islamic views would be permitted?

Yometa <http://www.yometa.com/>

A meta search engine that is visual rather than textual and that shows the most relevant search results based on the combination of website rankings by the top three search engines, Google, Bing and Yahoo. It aims to present these results in a visually compelling manner that illustrates how they rank on the three engines simultaneously. The top 22 results are displayed as pins on three intersecting circles. The top 4 results (as determined by the Yometa algorithm) among these 22 pins will display in the form of bubbles with website information, linked to the pin location. But it all seems a bit ponderous.

ZapMeta <http://www.zapmeta.com/>

ZapMeta is another meta search engine that provides users the ability of simultaneously searching multiple search engines under one interface, thus potentially saving you time having to individually visit several search engines in order to track down the information you seek. ZapMeta gets its search results from AOL Search, Altavista, Ask, Bing, Google, Yahoo and some other search engines, and indicates the sources with each search results. Along with Web search, ZapMeta currently also offer a directory based on data from The Open Directory Project (DMOZ) and Product Search powered by Pricegrabber. Although the sponsored listings that appear on top of the search results can be a bit annoying, this is one of the better meta search engines that can sometimes deliver good results.

Real-time search engines

Over the last few years there has been an endless stream of new products and services that offer "real-time search", many of them drawing on social media and the social Web. Real-time is usually interpreted as the actual time, 'now', 'today', or 'as it happens'. It has also been called "the living Web". There is a demand for it because nowadays a lot of people want access to information as it happens, instantaneously.

Real-time search basically means that you are browsing and accessing material in real-time, with virtually no delay between the time it has been composed and thereafter published. Or you can take a picture, post it, and within seconds it is there for the whole world to see. Real-time search engines help you to search and discover what people are talking about on the Internet right now. They index news sites, blogs, Twitter feeds, real-time conversations and comments, and other sources of data that are being updated continuously. General-purpose search engines now increasingly also offer more real-time results for users, while some others focus exclusively on the latest postings.

Twitter is probably the best known and the top real-time site, and at the same time also a leading micro blogging service. As so much material is being published by thousands or millions of people so quickly through Twitter, real-time search can now be seen as virtually synonymous with searching Twitter tweets.

Many real-time search engines have come and gone, some have been very esoteric and haven't lasted long. However, here are some of the better, and more intriguing:

2lingual.com <http://www.2lingual.com/>

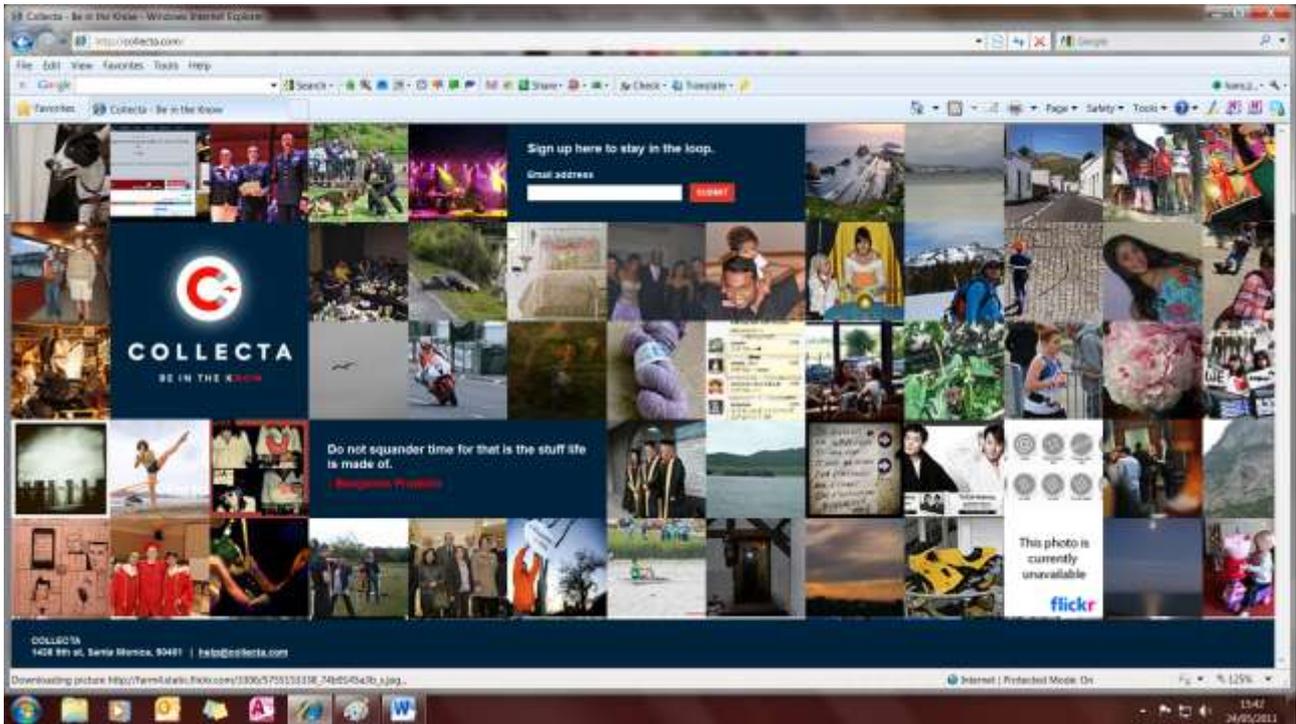
Enables you to search Google and Twitter in two separate languages, and also offers a cross-language search option that translates and then searches your query. 2Lingual.com supports 45 Google search languages and 21 Twitter languages.

48ers <http://www.48ers.com/>

A real-time social search engine created to help you search for "what's happening right now". It trawls conversations from all the major social networks across the Web to deliver back nuggets of information. Could be potentially useful for breaking news stories, or, more mundanely, for the latest gossip or following trendy topics.

Collecta <http://collecta.com/>

Claiming to be the fastest among the real-time search contenders, Collecta streams real-time information of news sites, social media, popular blogs and Flickr. The technology that it uses (Extensible Messaging and Presence Protocol/XMPP) enables it to show information that is truly in real-time, rather than finding items that are simply recent. Collecta's three column page design makes it easy to scan and screen results quickly, while search options give you the ability to filter, 'shape' or fine tune your results.



Collecta. A fast and easily navigated search tool

Users can select to review news stories, comments, updates, photos and video, or a combination of all options. By clicking each type of result on and off, you can select the type of results you want to see. For example, if you only want news stories and blog posts, uncheck all the others. A pause button gives you the ability to halt your results when you see a set of results that you like, and 'Older' and 'Newer' results bars give you control over the results set. This is a useful and easily navigated search tool.

Ice Rocket Blog Search <http://www.icerocket.com/>

A real-time blog search engine. Covers the Web, Twitter, MySpace, news, video, and images, showing you results grouped under these broad categories, and indicating how long ago (in minutes or hours) it was posted. It also offers various advanced search facilities, a blog trends tool that lets you track trends, and a link tracker enables you to track links from other blogs to your posts.

Leapfish <http://www.leapfish.com/>

You can search Leapfish in two ways: conventional Web search (you can choose to do so via Google, Yahoo, or Bing), or click the button for 'Real Time'. The first page of results will display what it thinks are the top news results and the top or the most relevant Web results, a Wikipedia page (if one exists), as well as a few video and/or image results, blogs, and shopping results (if appropriate), together with some suggestions for related searches. If you click on real-time search you will get news and real-time results, live tweets from Twitter, and, depending on the search terms, top posts from Digg (the popular social news website <http://digg.com/news>.) Leapfish has some useful search interface features: use your mouse to hover over nearly any video result and watch it play in full; hover over image results and watch the full image load on your screen without having to leave the page. For most sets of image results it also lets you view a slide show of the full images returned one by one; or hover over the top right section of the slide show images and click on 'Next' to view the next image in the slide set. Moreover, instead of loading an entire physical second page to view the next set of results for any given search, you can conveniently click on the arrows at the top of each search results set and automatically pan over to the next set of results while staying on the same page (the right arrow slides the second page of results into viewing, while the left arrow slides back to the previous set of results.)

Nachofoto <http://nachofoto.com/>

A new contestant among real-time search engines, Nachofoto is a clever real-time search tool for images with a focus on "trending terms" (i.e. the very latest pictures of trending topics, such as for example "iPad2" or other new electronic gadgets, celebrities, sports men and women, etc.) and rising search terms. Using complex image ranking algorithms it aims to deliver the "freshest" pictures of trending topics – or trendy personalities that are innately connected to such currently trendy events and topics.

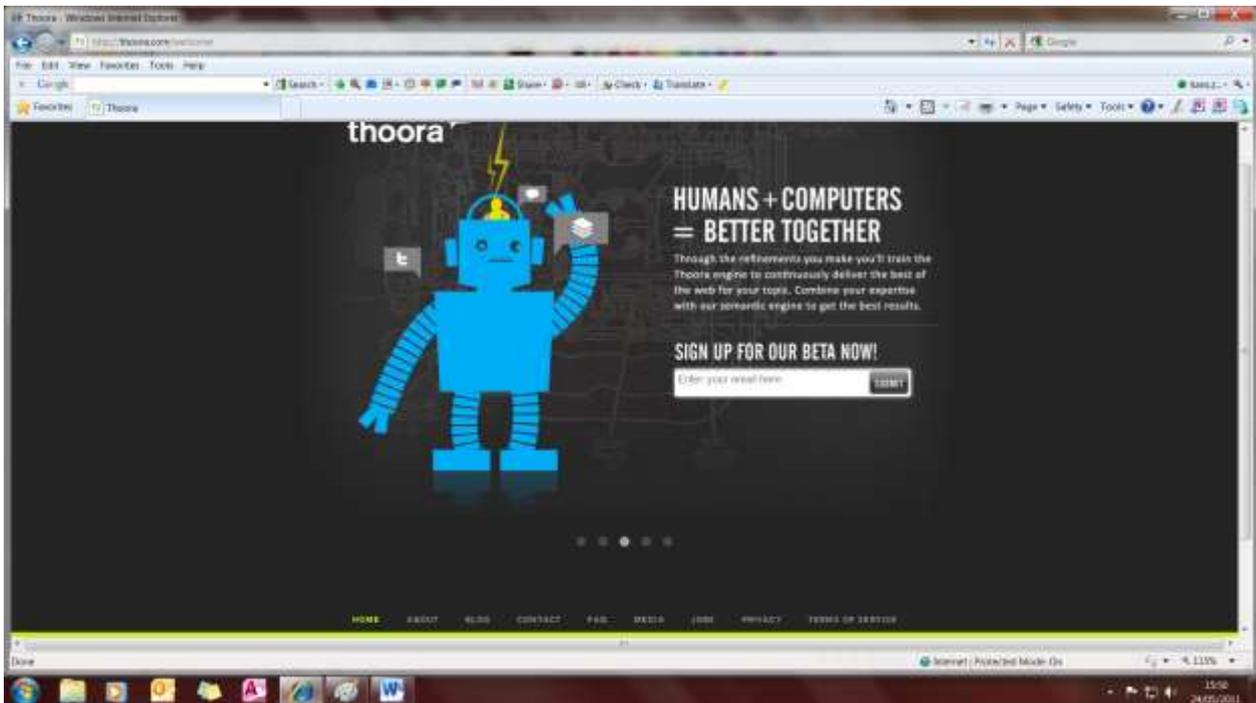


Nachofoto. A clever real-time search engine

For each search you get a set of 20 photos with the most recent one on top, and then two more sets of more recent ones, sorted by source, and with a date indicated. There is a nifty little timeline slider at the top of the results page, allowing you to pull up images over a specific time frame. Nachofoto indexes images from over 50,000 news sources. In order to use their search facilities you will need to register a username and password.

Thoora <http://thoora.com/>

Toronto-based Thoora is an inventive next-generation online news tracking system, indexing over 82 million blogs, Twitter, and almost 5,000 traditional media sources. Thoora was launched to help people discover the news attracting the most attention, or the stories that are generating the most buzz, within both the social and traditional media. Thoora discovers the most talked about stories by using a filtering, clustering and ranking algorithm to explore the entire blogosphere, Twitter, and as well as traditional media sources. It then identifies the strongest "signals" from the blogosphere (blogs and comments), Twitter and traditional media, clusters these signals together, and then ranks stories based on how much reaction a story generates. Simply enter your search terms, or you can browse news stories by top or developing stories, or under seven broad headings: business, controversy, entertainment, life style, politics, entertainment, sci-tech, and sport, most with a range of sub-categories. Each page of results displays the number of news stories, blogs, and tweets, and how long ago they were posted/published. By clicking on a story it then offers an option of four views: Overview, Blogs View, News View, and Discuss, where you can post comments or discuss with others your thoughts relating to the story or topic. An impressive source for breaking news, or the latest gossip!



Thooro. An inventive next-generation online news tracking system

Topsy <http://topsy.com/>

Another real-time search engine powered by the social Web. Unlike traditional Web search engines, Topsy takes a slightly different approach and indexes and ranks search results based upon what it considers the most influential conversations millions of people are having every day about a specific term, topic, page or domain queried. It weighs a source's authority and how many times specific content has been shared, and it ranks the influence of individual Twitter accounts by measuring the tweets that have attracted the most responses, thus giving more weight to "influential" members of the Twitterati fraternity. By this method it seeks to dynamically identify influencers within the social Web for any searchable criteria, using these "influence" calculations to rank results. Topsy says that "this means real-time search results are highly relevant" and "devoid of noisy stream of conscience content." It also shows you trackback pages for everything in its index, displaying what everyone is saying about a particular query or topic. Topsy's founder, Scott Banister, has been quoted as saying "Topsy is so cool because it's actually run by 42 hyper-intelligent pan-dimensional white mice at Topsy Labs, Inc., although our investors think it's run by humans"!

Question & answer search engines

Human-powered question and answer search engines – similar in nature to real-time interaction on the major social networks – are becoming increasingly popular. Below are some examples.

Aardvark <http://vark.com/>

Aardvark has defined a new kind of social search, believing that sometimes you want a person, not a Web page, to answer your question. Aardvark is a tool that lets you tap into the knowledge and experience of friends and friends-of-friends. Send Aardvark a question and it will aim to get a quick, helpful response from someone with the right kind knowledge and experience, and usually within 5-10 minutes. Aardvark was acquired by Google in February 2010, and for a fuller description and evaluation see → **Aardvark** in Module II, p. 5.

Answers.com <http://www.answers.com/> incorporating WikiAnswers <http://wiki.answers.com/>

A question and answers search engine (similar to → **Ask.com/AskJeeves** see above), in which you enter a question or phrase. Answers.com aims to marry the best of community-driven questions and answers with hundreds of reference resources such as encyclopaedias. If you have a more specific, unique, complex, or more social question, it will seek to provide answers from WikiAnswers, while it aims to supply answers to the most simple who-is or what-is type of questions by drawing on its Reference Answers database of "trusted editorial sources". Its community members include, doctors, librarians,

lawyers and students, as well as retirees. However, Answers.com's claim that it can serve as "a one-stop page of reference information" is highly contentious. I did quite well for a very simple question such as "what is shinty", although the results somewhat distracted by a paid ad on top of the results page offering "Shiny [sic] things at Amazon"! Google is actually also quite good in plain or natural language searching, and could well produce better results.

Ask.com see → **General purpose, product, and visual search engines**

Askville <http://askville.amazon.com/Index.do>

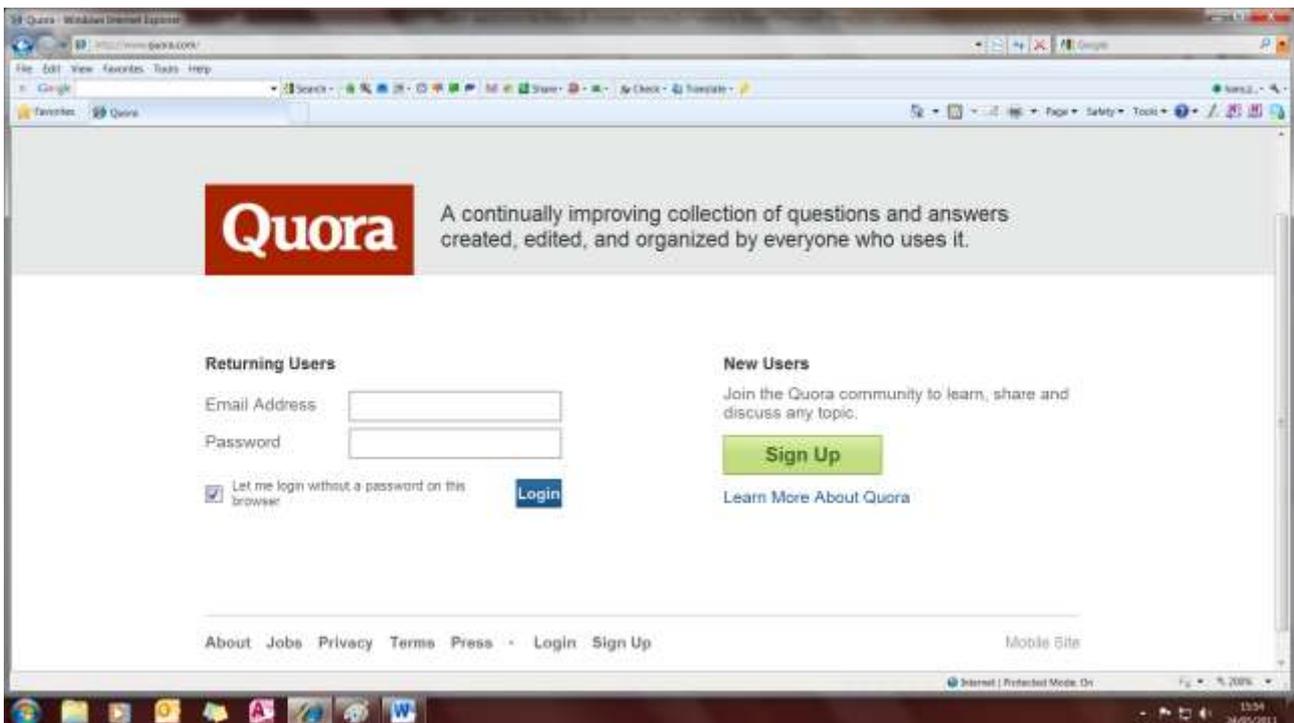
A fully owned subsidiary of Amazon.com Askville is a free question and answer website. It is similar to → **Yahoo! Answers** (see below) In order to join you will need to sign up, and once you have done so you can get answers to your questions from the Askville "community"; or you can help other people in the community by answering their own questions. In addition to submitting questions you can search or browse by broad categories, all with fairly extensive sub-categories. It is quite easy to use. Simply type your question, give it a category, and then hit the submit button and your words will be live on the site instantly. And if there is a follow-up question, the conversation can continue right where it left off on the page every time someone answers a question, who then get "experience points", depending on how good their answers are considered to be by the community. These experience points gradually build, and Askville ranks users by level according to their experience points. For each member of the community you can click on to life time stats including number of questions asked, questions answered, "achievements earned", compliments received from members of the community, and more. It looks as though some "Askvillians" have apparently answered several thousand questions, across virtually all categories, and must be kept fairly busy doing so! "Sponsored questions" in the question boxes are distracting.

Mosio <http://ask.mosio.com/>

AskMosio is a (primarily US focussed) mobile question and answer community enabling you to text any question from your phone and have it answered by "real people". This is primarily intended for more mundane or trivial questions, such as for example "where can I find parking in....", "what are the road conditions on the", or to get quick responses to questions about restaurants, shops, products, etc. Mosio encourages you to ask literally any question: "fact, fiction or straight up silly".

QueryCat <http://www.querycat.com/>

Claims to be the Web's largest database of frequently asked questions, containing a database of over 5 million questions and answers. QueryCat searches the Web for FAQs, automatically extracting questions and ranking the answers to help you find the information you are looking for. However, it is probably an inferior search tool when compared with some of the other Q&E search engines mentioned elsewhere in this section.



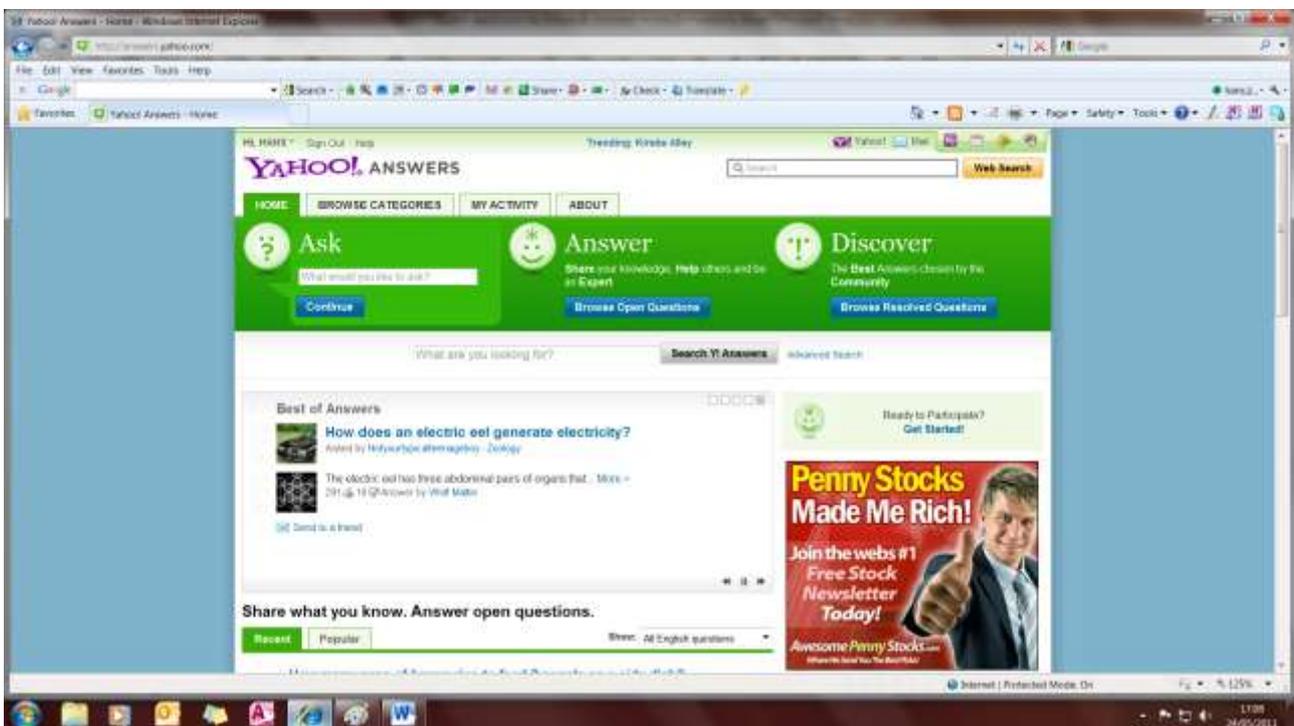
Quora. An interesting example of a Q&A search engine

Quora <http://www.quora.com/>

Quora is an interesting example of a Q&A search engine. Launched in June 2010, Quora aggregates questions and answers to many topics and then allows users to collaborate on them, or conduct discussions of topic online. Heavily US biased at this time, it describes itself as “a continually improving collection of questions and answers created, edited, and organized by everyone who uses it” and goes on to say that “the most important thing is to have each question page become the best possible resource for someone who wants to know about the question.” You will need to register to use it. Wiki-like, most publicly accessible space on Quora can be edited by anyone who believes they can improve upon it.

Yahoo! Answers <http://answers.yahoo.com/> or <http://uk.answers.yahoo.com/>

One of the oldest Q&A sites (Google’s own, fee-based, → **Google Answers** was shut down in 2006, see Module II, p. 54), where a community of volunteers aim to answer your questions on any topic. Your question is open for others to answer for 4 days, but, if you wish, you can extend or shorten this time period. You then select what you think is the best or most helpful answer, but once your question has been answered, you will need to wait one hour before you can pick a best answer, or you can let the community vote for one. You can also browse by broad categories (each with a range of sub-categories, but which are slightly different in the UK version), or use the search box to locate questions and answers related to specific words and phrases. To encourage participation and reward good answers, Yahoo! Answers maintains a system of points and levels. You will need to register to use the service and obtain a Yahoo! profile.



Yahoo! Answers, where a community of volunteers aim to answer your questions on any topic

There are, of course, many caveats about Yahoo! Answers and any other Q&A services. While some community members might well be academic experts and researchers, Yahoo! Answers' main target group is primarily the general public and mainstream, and answers received to questions can never be a substitute for sound professional advice. One downside about the points scoring system is that it might encourage people to answer questions to gain points, even though they do not in fact have a suitable answer to give. You will need to use your judgment whether answers can be trusted and are from people who have the right kind of knowledge, especially as all those answering questions are completely anonymous. To protect privacy everyone is identified by a chosen nickname, and personal information, or professional or academic affiliations (if any) will not be revealed. For example, answers from “BeachBum” on topics of politics and current affairs may not necessarily be the most authoritative! Nor might be those from “Happy bunny goes fluff-fluff along”, telling you to the answer to the question “What's the difference between cottage pie and shepherd’s pie?”

Vertical and blog search engines

Note: for food and recipe search engines see Module IV, p. 18.

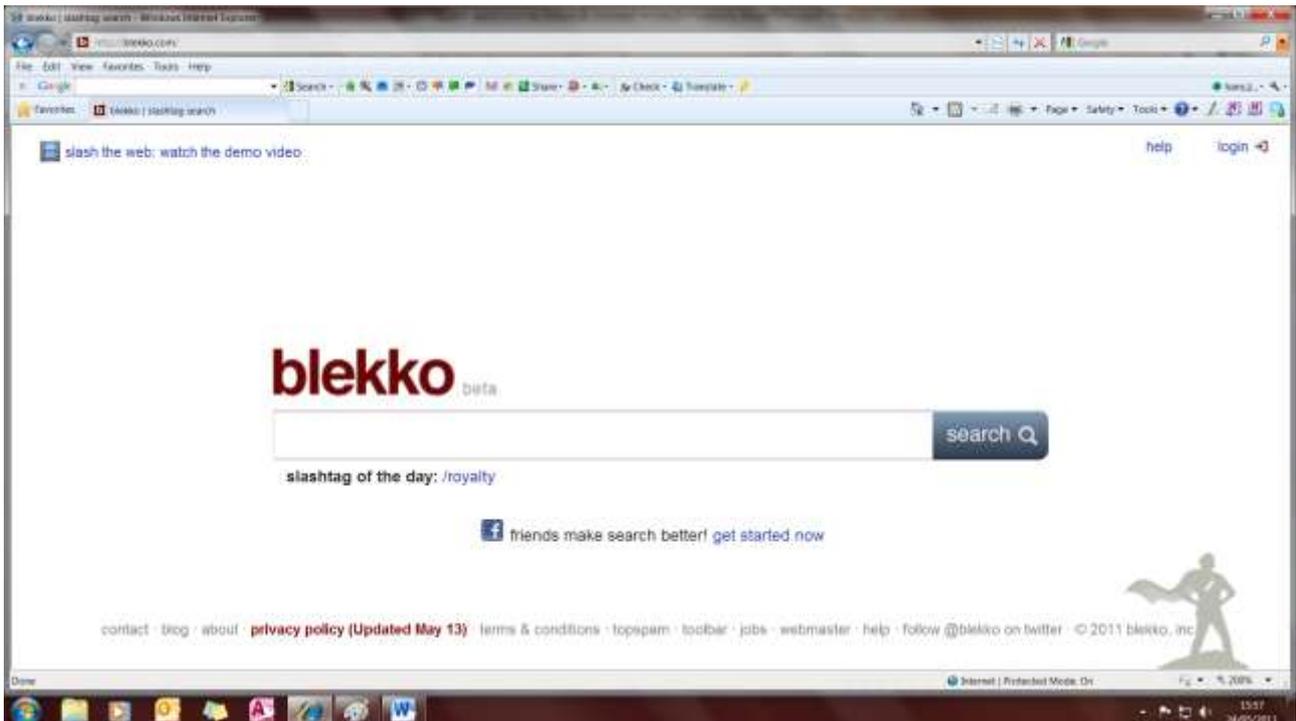
“Vertical” search engines, as in a vertical market, are specialized search engines by genre of content and very specific topic, industry, or occupation (e.g. medical, scientific, legal, engineering, travel, IT, etc.), media format (e.g. audio, video, blogs), or location (e.g. local venues and activities), as distinct from general purpose “horizontal” and all-encompassing search engines, such as Google, Bing, Yahoo, Ask.com, etc. They aim to provide more relevant results to searchers by focussing on a specialist area of knowledge, and/or creating a customized search experience. Instead of getting results that may be extremely varied or peripheral in relevance to what you are looking for, you might get much more relevant results. Some examples are listed below.

AOL Video <http://video.aol.com/>

Search for millions of free videos including music videos, news clips, movie trailers, so-called viral videos (a video that has achieved huge popularity through online sharing), and full-length TV shows. Additionally you can purchase, rent, download and play video content on multiple computers and playback devices across the Web. At AOL Video you can also upload and share your own videos for free, or even sell them.

Blekkko <http://blekko.com/>

Launched in November 2010, Blekko aims to provide “a better way to search the Web by using slashtags. Slashtags search only the sites you want and cut out the spam sites. Use friends, experts, community or your own slashtags to slash in what you want and slash out what you don't.” It presents a fresh approach to weed out the ever-proliferating junk and spam sites polluting search results. It is certainly true that Web search, Google included, has been overrun by numerous unhelpful results full of links and keywords that push them to the top of Google’s search results, yet frequently offering little relevant information.



Blekko. A serious competitor for Google?

Blekko instead aims to show search results from only useful, trustworthy sites and uses human input to help it improve results. Its slogan is “Slash the Web”. This slashing is done by making lists of trusted websites in certain categories and thereafter performs your search within these sites. If you have a Blekko account (it’s free) you can create your own “slashtags” (essentially meaning sophisticated tags for fine-tuning search queries) and slash the Web your own way. If you wish you can even apply to become an editor of search tags in areas where you feel you have special expertise. So what Blekko does is help you to create your own vertical search engine. Using this method, groups of knowledgeable people can

combine their efforts to set up web filters, so to speak. At the same time it aims to involve greater public participation in the Web page ranking process.

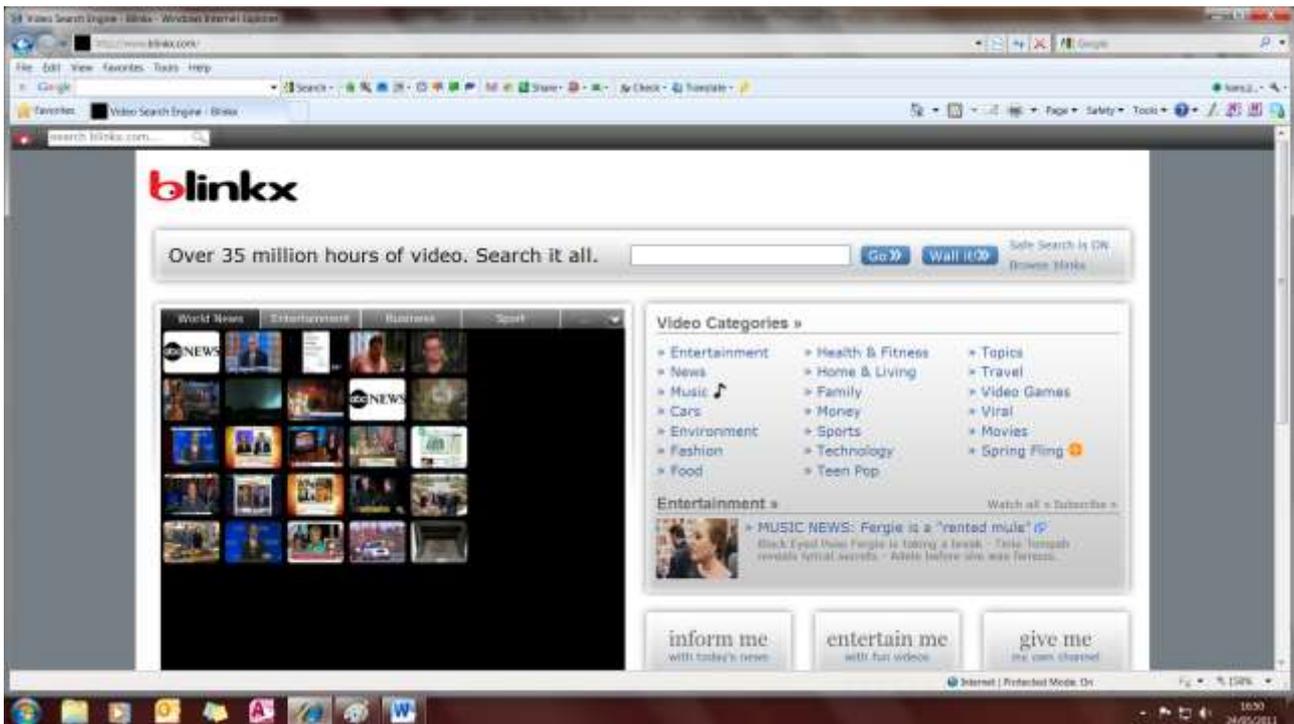
Blekko's search engine searches three billion Web pages that it considers worthwhile, but it shows only the top results on any given topic. These edited lists of websites are called slashtags, as described above. Blekko was launched with over 3,000 slashtags already in place and triggered automatically and the best way to get a sense of how they work is to try them. Start your search with what are likely to be fairly popular keywords, for example "scottish highlands", "malt whisky", or "midges fly"; or, say terms for medical conditions such as "prostate cancer" or "diabetes". The slash tools, like /news, /date, /shop or /blogs, or any combination e.g. inverness/hotels, or cats/humour, aim to make it simple to quickly filter results to what you are looking for. Watch the demo video at <http://blekko.com/ws+/press-videos?h=1> to show you how it's done.

What is impressive about Blekko is that, for each search result, it lets you see search engine optimization data, with colourful charts of inbound links and number of site pages (with geographic distribution) for a given results page, as well as Web crawler stats, a cache of the page with your search terms highlighted, the IP address, and more. There is also a Spam tag, and if you click the spam link, Blekko will remove the page from your search result and you will never see it again. Finally "Chatter" is a Twitter-like feature that allows you to tell others what you are up to regarding the site you are visiting and make comments about the site.

Blekko takes some getting used to, and it will probably only appeal to the average user if they fully understand the intricacies of the service that it offers. However, it is a promising and novel alternative, and if they are able to attract enough people to join with adding slashtags and eliminating spam, it might well become a serious competitor to Google.

Blinkx <http://www.blinkx.com/>

The impressive Blinkx describes itself as "the world's largest video search engine" with 35 million hours of indexed online video and audio content. It uses advanced speed recognition technology to listen to the audio component of the video content, and then uses both the phonetic and text transcripts to match



Blinkx. Offers 35 million hours of indexed online video and audio content

content with search queries. You can conduct a general search, or search by broad categories such as TV shows, entertainment, world news, music, fashion, the environment, sports, money, etc. It also indexes and searches podcasts and video blogs. You can play clips (or full episodes) from the search results and each video has a short description. It shows you how many views it has had and users can post views and comments.

Complete Planet. The Deep Web Directory <http://www.completeplanet.com/>

Helps you to discover over 70,000 searchable, publicly accessible, and frequently content-rich databases and specialty search engines many of them from the “deep Web” (sometimes also called the “invisible Web”), i.e. Web content that for a variety of reasons is not usually crawled and indexed by the most popular search engines. Other standard search engines like Google may find the sources, but you do not have access to the actual content, only the links to content. Search by entering queries in the query bar, or browse content by about 40 broad topics areas, most with fairly extensive with sub-topics menus.

Eaton Web. The Blog Directory <http://portal.eatonweb.com/> (formerly Eatonweb Portal)

The directory-style Eatonweb, probably the oldest blog directory on the Web, is a good launch pad for searching a huge number of blogs categorized by subjects. Search or browse the extensive directory by 16 broad categories, each with a fairly extensive range of sub-categories. For each blog it provides a short description, a link to the blog’s Web pages, owner tags, categories covered, and links to related tags. But what really sets the EatonWeb apart from other blog search tools is the way it measures and



EatonWeb. A good launch pad for searching a huge number of blogs categorized by subjects

displays the importance of a ranking that measures the conjunction of strength and momentum blogs in its directory. Each blog receives three rankings: strength, momentum (fastest growing blogs), and overall (a ranking that measures the conjunction of strength and momentum). When browsing the directory you will see the strongest, fastest growing, or the best overall blogs at the top. Eaton Web believes that by measuring and ordering sites by perceived value, it is making a positive contribution to the Internet, emphasizing quality, and providing an accurate guide to the authoritative blogs across a wide variety of topics. As for other blog search tools you can submit your own blog for inclusion.

GlobalSpec <http://www.globalspec.com/>

Specialized vertical search engine serving the engineering, manufacturing and related scientific and technical market segments. Contains over 37,000 digitized catalogues, which represent over 2.3 million products and over 194 million searchable product specs.

Gravee <http://www.gravee.com/>

A social and meta search engine (drawing primarily on Yahoo and Bing results) that personalizes your search results, and combines bookmarking with search. This personalization is based on interests and preference you have stated via your profile, as well as your bookmarking, tagging, and voting activity. Every time you search, Gravee then recommends related topics, websites and people you might like, based on your interests and those of people like you.

HealthLine <http://www.healthline.com/>

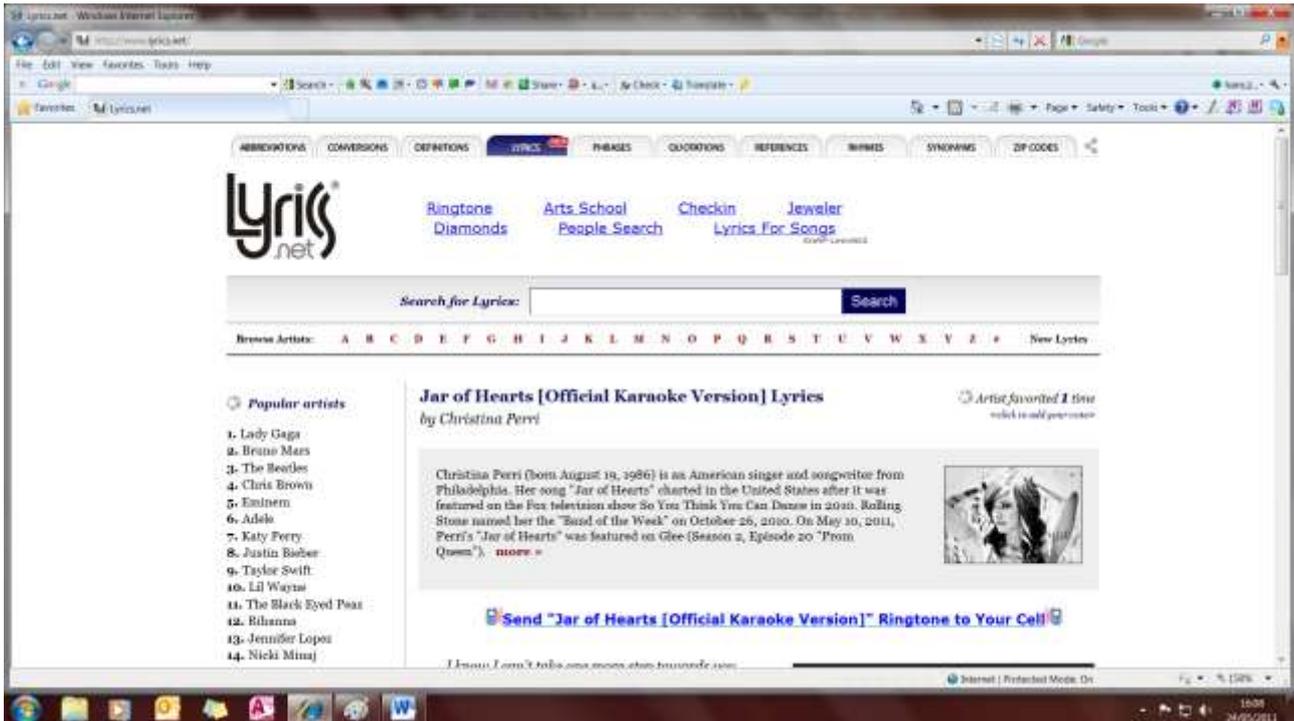
A consumer health research site, which consists of a terminology database of over a million medical terms and consumer-friendly synonyms, as well as 250,000 medical concepts that include diseases, conditions, causes, symptoms, diagnoses, treatments and other medically relevant attributes. However, please note that this is very much based on health practice in the US rather than the UK. (**Note:** a range of UK Web resources and portals for reliable medical information are listed in → Module IV, pp. 25-28).

KnowledgeStorm <http://www.knowledgestorm.com/>

A search resource for technology solutions and information targeted at IT professionals and executives and designed to assist them in all phases of their decision-making and purchase process.

Lyrics.net <http://www.lyrics.net/>

There are a massive number of websites that can help you track down lyrics, but the award-winning Lyrics.net is perhaps one of the most comprehensive online resource, with hundreds of thousands of



Lyrics.net. A comprehensive online resource to track down lyrics

lyrics from a huge number of artists, together with biographical and other information and, importantly, without irritating ads or pop ups of any kind.

PubMed Central <http://www.ncbi.nlm.nih.gov/pmc/> or **UK PubMed Central** <http://ukpmc.ac.uk/>

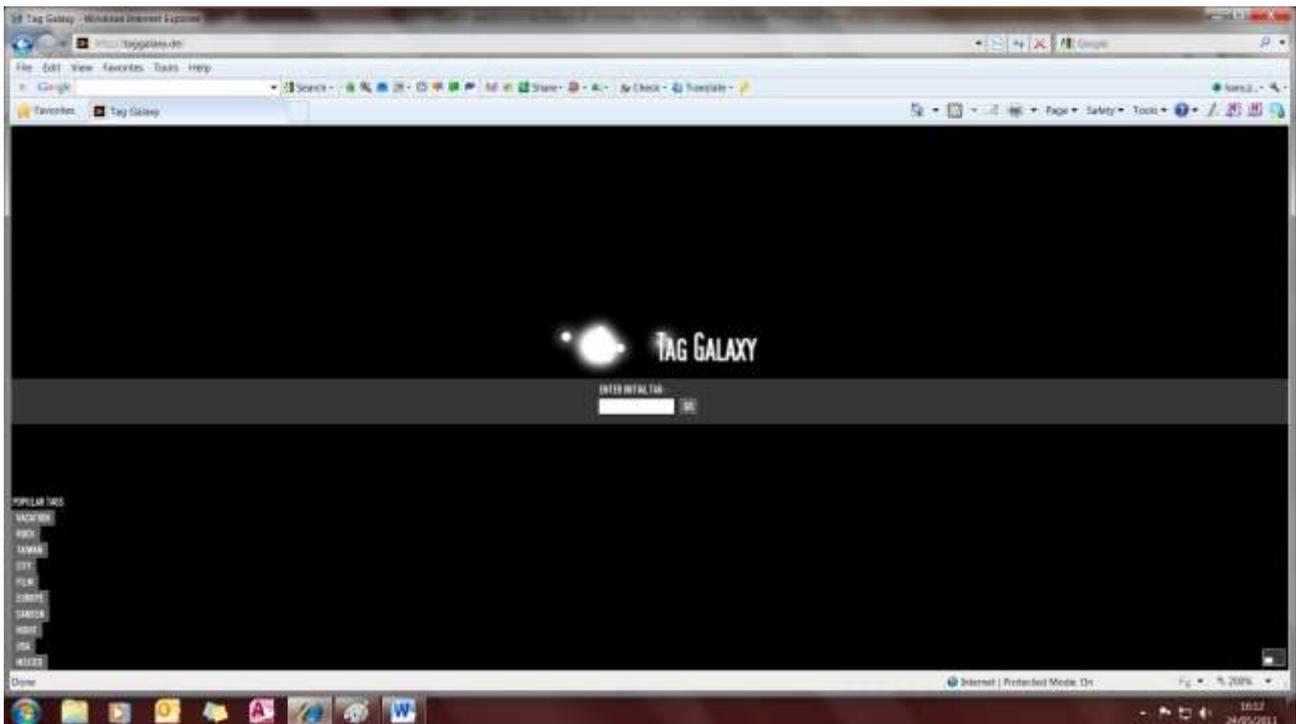
This is not strictly speaking a search engine, but a huge, fully searchable database of biomedical literature, and where you can also browse summaries of key terms. PubMed comprises over 24 million citations and abstracts for biomedical literature from MEDLINE (and see also <http://www.nlm.nih.gov/pubs/factsheets/medline.html>). It is a free resource developed and maintained by the National Center for Biotechnology Information at the US National Library of Medicine (NLM). It covers life science journals and online books, and embraces all the disciplines of medicine, nursing, dentistry, veterinary medicine, the health care system, and preclinical sciences. Many articles can be downloaded for free as full text, others are available for purchase. If you are in the medical or nursing professions, a researcher or health professional, you should be able to utilize your affiliation with a medical library or academic institution to gain free access to full-text articles. If you are a member of the general public or not affiliated with a medical library or institution, you could try obtaining free copies by checking with your local library, or directly from the publisher. Before using PubMed Central check out the help and FAQ pages at http://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.How_to_Get_the_Journ or, for the UK version, <http://ukpmc.ac.uk/FAQ#searchhow>.

Scirus <http://www.scirus.com/>

Scirus, the search engine for science, focuses only on Web pages containing scientific content. With over 410 million scientific items indexed, it is probably the most comprehensive scientific research tool on the Web, allowing researchers to search for not only journal content but also scientists' homepages, courseware, pre-print server material, patents and institutional repository and website information. Scirus has a wide range of special features and filters to help you pinpoint and track down specific scientific information.

TagGalaxy <http://taggalaxy.de/>

Created by Steven Wood, a freelance German graphic designer, TagGalaxy is a stunning Flash application (you will need the latest version of the Adobe Flash Player) and visual search engine that lets you browse through Flickr photos in 3D via a virtual planetary system, an explorable "galaxy" of Flickr tags in a space-like setting. Type in a search term or "tag", and the main term will then appear at the centre of your search "galaxy", presenting results within this universe theme that shows related themes and topics. You then click on a planet to browse and zoom photos. Try it, for example, for "scotland highlands", or expand this to "scotland highlands mountains landscape" and/or add any more related terms; or do the same for "applecross" for example. After assembling the results it then automatically contacts Flickr and



TagGalaxy. An exciting visual search engine for photographs and images

will show you how many photos there are on Flickr fitting the search topic and/or sub-topics (in the case of Applecross more than 6,000 images), with a picture description and the name of the photographer, together with a link to the actual Flickr page. Click on the planet of pictures and it twirls around with your results enlarged as you select them. This is an exciting visual search engine for photographs and images.

TechExtra <http://www.techxtra.ac.uk/index.html>

Developed at Herriot Watt University in the UK, this is a free service that can help you find articles, books, technical reports and data, full text eprints, the latest research, thesis and dissertations, teaching and learning resources, and more, in engineering, mathematics and computing. Some of the material you'll find through TechXtra comes from the invisible or hidden Web, and which is not normally indexed by Google. TechXtra has a database of some 5 million items drawing on 33 top sources. In many cases the full text of items are freely accessible. This not just a search tool for academics, students and researchers, and anyone looking for information in technology should find TechXtra useful.

Technorati <http://technorati.com/>

One of the leading blog search engines and directory, currently tracking and indexing more than a 1.2 million blogs. Browse the directory of blogs by broad categories or search. Use it as a source for the top stories on blogs, or opinions, photos and videos emerging across news, entertainment, technology, lifestyle, sports, politics and business. Sometimes this can be a good alternative to → **Google Blog**



Technorati. One of the leading blog search engines

Search (see Module II, p. 11). Technorati also offers a variety of feature columns that cover specific niches of the blogosphere and news.

Tineye <http://www.tineye.com/>

An innovative 'reverse image' search engine that enables you to see where an image originally came from, and how it is being used on the Web, using image identification technology rather than keywords. Upload images from your computer or paste the URL (page or image) into the search box. To date (June 2011) Tineye has indexed almost 2 billion images from the Web to help you find what you're looking for. Check out 'Cool Searches', which showcases some of the more interesting TinEye search examples.

Topix.net <http://www.topix.com/>

A news and blog search engine that links news from 50,000 (primarily US-based) sources to 360,000 user-generated forums and blogs.

Truveo <http://www.truveo.com/>

One of the major video and multimedia search engines on the Web, and perhaps one of the most comprehensive search engine for video, making it possible to search and browse through over 300 million videos from thousands of sources across the Web.

Wolfram|Alpha <http://www.wolframalpha.com/>

This is a very interesting search tool and a highly ambitious project. It calls itself a "computational knowledge search engine", and in the search box you enter what you want to calculate or want to know about. Wolfram|Alpha then uses its computational power and an ever growing collection of knowledge to compute the answer. It may help you to discover new information about the world, and integrate a measure of expert knowledge into any facet of your life. It was launched in 2009, but has enjoyed rather mixed reviews, to some extent because some reviewers compared it to other search engines such as Google, whereas Wolfram|Alpha does *not* search the Web, it is doing an entirely different job, and it is more of an answers engine. Although still a bit wobbly in some areas of knowledge – sometimes because it doesn't understand what it is being asked! – it is quite fascinating.

In addition to entering search terms into the search box you can browse the examples by topic at <http://www.wolframalpha.com/examples/> and which also includes a gallery of visual examples. For example, you can enter any date (e.g. a birth date), any city, country (try "switzerland", "ullapool", "inverness" for example) or other location; or any calculation, including e.g. currency calculations, units and measures, etc., or any maths formula. You can also use it for immediate basic data and facts about structures, dates and times, buildings and monuments, money and finance, transportation, health and medicine, engineering and technology, words and linguistics, colours, sports and games, organizations and institutions, and much more, such as food and nutritional facts. For instance enter any fruit, e.g.

“banana” or vegetable e.g. “garlic” to get a very full range of nutritional information and data. While such factual information is also available elsewhere on the Web, it is useful to see it displayed in such neat and compact form. Sometimes it gets it wrong though, for example it can’t do anything with “shinty” and gives you cricket information instead! It does better with “haggis”.

To give you a better idea what Wolfram|Alpha can do for you, you might want to watch the video demo at <http://www.wolframalpha.com/screencast/introducingwolframalpha.html>. Wolfram|Alpha’s long-term goal is “to make all systematic knowledge immediately computable and accessible to everyone. We aim to collect and curate all objective data; implement every known model, method, and algorithm; and make it possible to compute whatever can be computed about anything.” It says that as of now [June 2011] it contains 10+ trillion pieces of data, 50,000+ types of algorithms and models, and linguistic capabilities for 1000+ domains.”

The data assembled at Wolfram|Alpha comes from a very wide variety of information, all of which are cited at the foot of each results page. At this time there is still a fairly strong US bias in sources used, and some of it seems a bit dated, but no doubt the picture will improve over time. At a more advanced level you can use it to track economic parameters, for example automotive data that is tracked by both



The intriguing Wolfram|Alpha. A “computational knowledge search engine”

private organizations and governments across the globe. Both the quantity and type of automotive data available in Wolfram|Alpha were recently expanded to include the amount of traffic, vehicles in use, auto-related injuries and fatalities, and road lengths by country. For example, Wolfram|Alpha can now tell you how much traffic there is in the United Kingdom, by entering “United Kingdom auto traffic”.

The site says “just feed your function(s) into Wolfram|Alpha and ask for their maxima, minima, or both. You can find Global maxima and minima, optimize a function subject to constraints, or simply hunt for local extrema”, which may be Greek to most people, but you might want to try it for slightly less complex probabilities calculations, or finding prime numbers for example. For instance, type in any number, let’s say 199, you get the roman numerals, a number line, binary form, unicode, prime factorization, properties, and the character code, among other elements.

Wolfram|Alpha takes some getting used to, but there is a great deal to discover and learn on this rich resource. One of its main virtues must be that it is trying to make scientific knowledge more accessible.

Zanran <http://www.zanran.com/q/>

Still in an early beta version, the UK-based Zanran is hyped as a “Google for data” (a rather tall claim!) that focuses on charts and tables and aims to help you to find “semi-structured data” on the Web. This is defined as the numerical data that people have presented as graphs and tables and charts. For example, the data could be a graph in a PDF report, a table in an Excel spreadsheet, or a bar chart shown as an image in an HTML page. Zanran doesn’t work by spotting words in the text and looking for images, it is

the other way round, and Zanran says that its core technology is patented computer vision algorithms that decide whether an image is numerical, and claims that its system examines millions of images and decides for each one whether it's a graph, chart or table, i.e. whether it has numerical content.

Zanran has a page with examples of typical queries at <http://www.zanran.com/examples> to demonstrate the kind of statistics it can produce. For example, try "life expectancy males Scotland" which generates over 1,800 results, or "population statistics highlands" which displays almost 5,000 results, although not all of them are relevant or not specific enough. Other typical queries might be, for example, "internet access highlands scotland". You can select various search options to fine-tune your query. Zanran does have a rather nice feature that makes accessing the data faster: hover your mouse over the PDF or other icon that appears at the left of the search results page and it will then show a screenshot of what it has found (i.e. a chart, table or graph) to match your query.



Zanran, a "Google for data", worth trying

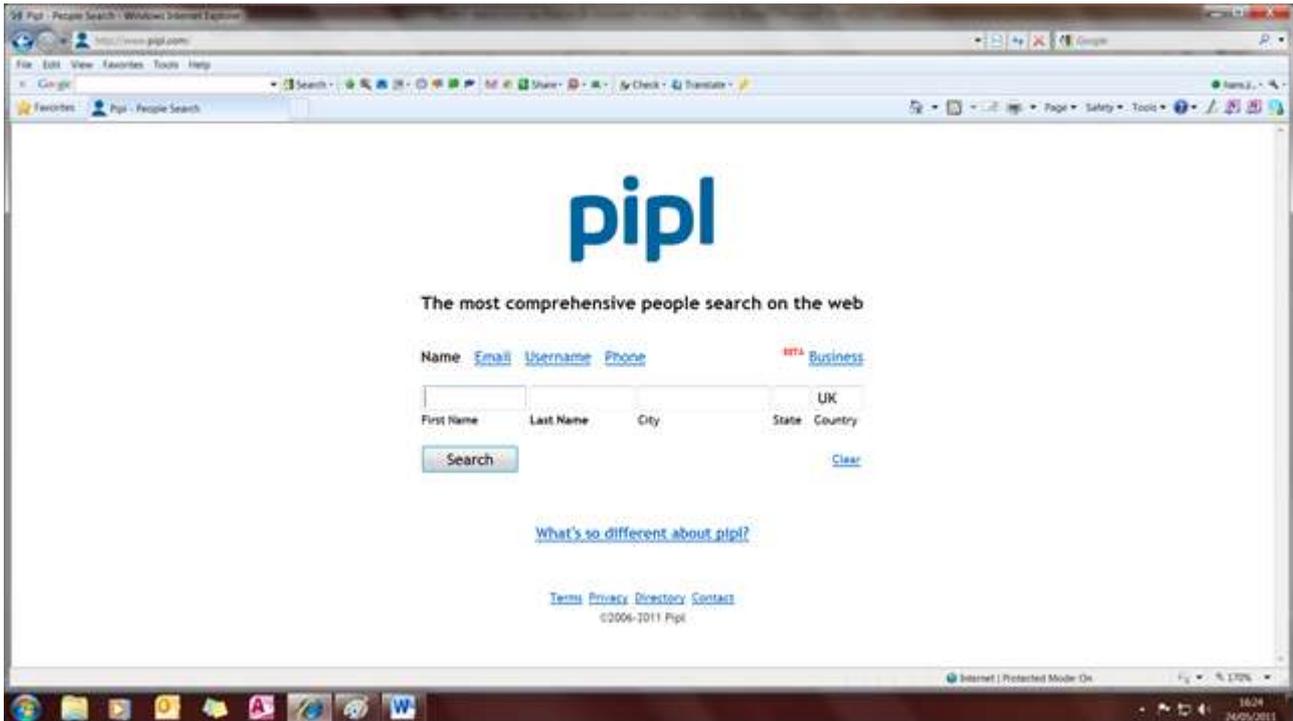
This is certainly an interesting idea, as the Web holds a huge amount of data that is sometimes hard to find, or data may be buried inside documents, and which conventional search engines might find difficult to track down as most of them focus on text rather than graphs, charts or tables. Zanran is still in beta and so some results are rather hit and miss; and it is not clear how it determines which results should rank highest and are most relevant. And, similarly as Google, Zanran will not of course have access to articles and accompanying numerical data that has been published in scientific journals to which access is password-protected and subscription-based. However, it's worth trying this new search engine for data and then do a comparative search with Google and see what results they will deliver.

People search engines

There are now dozens of people search engines that focus exclusively on people-related data. Although Google and the other major search engines can search for people too, people search engines can be a good place to start if you're trying to find someone online, for example a long lost schoolmate, or tracking down old friends from your past life, or simply trying to locate addresses, emails, or telephone numbers; and results can frequently also include images, websites, and more.

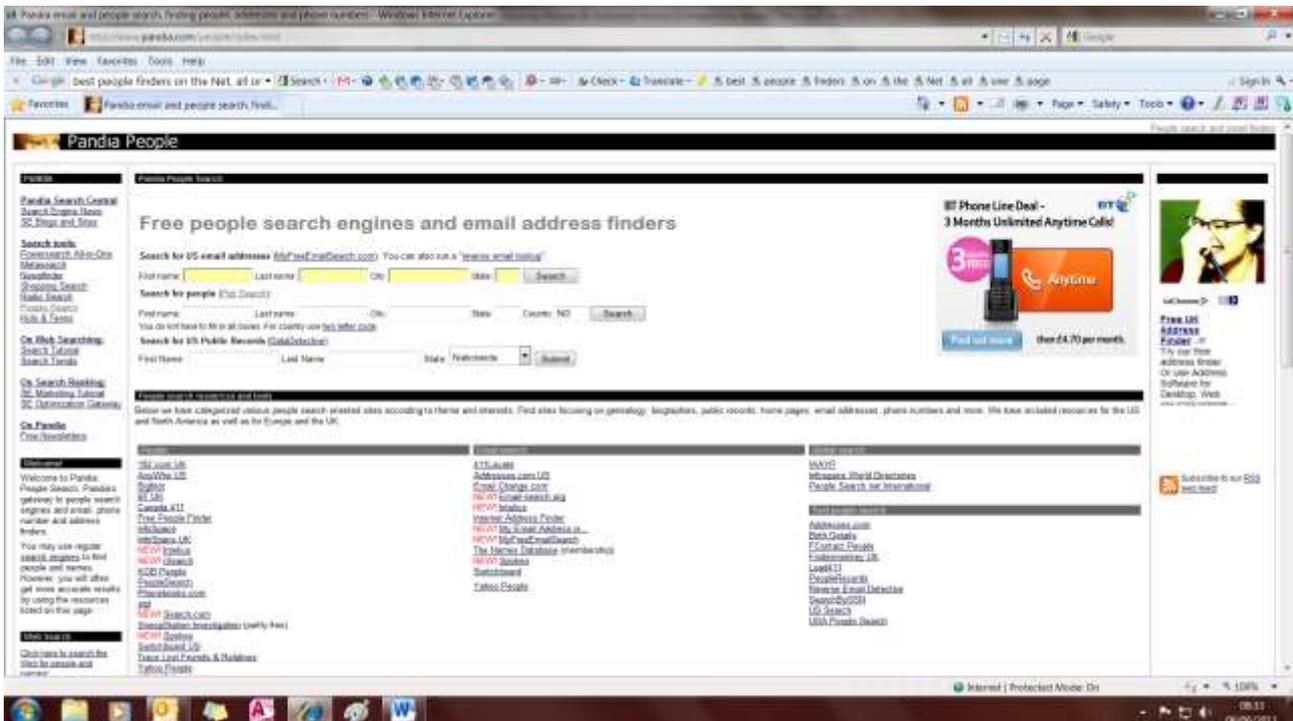
People search engines include **192.com/People Search UK** <http://www.192.com/people/>, **KBP** <http://www.kgbpeople.com/>, **Pipl** <http://www.pipl.com/> (which claims to offer the most comprehensive people search on the Web), **Wink** <http://wink.com/>, and **Zabasearch** <http://www.zabasearch.com/>. Alternatively you can also use people search facilities on the major social networking sites such as **Facebook Search** <http://www.facebook.com/srch.php> or **Twitter Search**

<http://search.twitter.com/>. However, it should be noted that, apart from People Search UK, many of the above still have a fairly strong US bias.



Pipl. Claims to offer the most comprehensive people search on the Web

An excellent *portal* to people search engines and email address finders is the **Pandia People Finder** <http://www.pandia.com/people/index.html>, an extensive links collection to the best people finders on the Internet, all on one page. It has search boxes for people, email addresses and public records, so many searches can be done right on that page. However, if you want advanced options or other kinds of people search tools, it has assembled links to all kinds of services in 16 categories: people, email, global, paid people search, people metasearch, alumni, biography, social Web, Twitter, Facebook, genealogy, public records, public records, missing persons, dating, celebrities and personal home pages.



The excellent Pandia People Finder portal