



## Tales from the Terminal Room August 2008, Issue No. 84

Editor: Karen Blakeman

Published by: RBA Information Services

ISSN 1467-338X

<http://www.rba.co.uk/tfttr/>

---

Tales from the Terminal Room (TFTTR) is an electronic newsletter that includes reviews and comparisons of information sources; useful tools for managing information; technical and access problems on the Net; and news of RBA's training courses and publications. Many of the items and articles will have already appeared on Karen Blakeman's Blog at <http://www.rba.co.uk/wordpress/>

Tales from the Terminal Room can be delivered via email as plain text or as a PDF with active links. You can join the distribution list by going to <http://www.rba.co.uk/tfttr/> and filling in the form. You will be sent an email asking you to confirm that you want to be added to the list. TFTTR is also available as an RSS feed. The URL for the feed is <http://www.rba.co.uk/rss/tfttr.xml>

---

### In this issue:

- Search Tools
  - TouchGraph Google Browser
  - Yahoo drops Boolean NOT operator
  - AllPlus
  - Cuil not so cool
  - Top Search Tips
- Internet Librarian International 2008
- Online Information 2008
- Information Resources
  - Guernsey Registry
  - Renewable Fuels Agency (RFA)
  - UK Coal Authority
  - Say no to 0870 – alternative telephone numbers
- Gizmo of the Month
  - Tagclouds for analysing documents
- Meetings and Workshops
  - Being a Digital Citizen, Leeds
  - Market Research on the Web, Manchester
  - Business Information on the Internet, Manchester

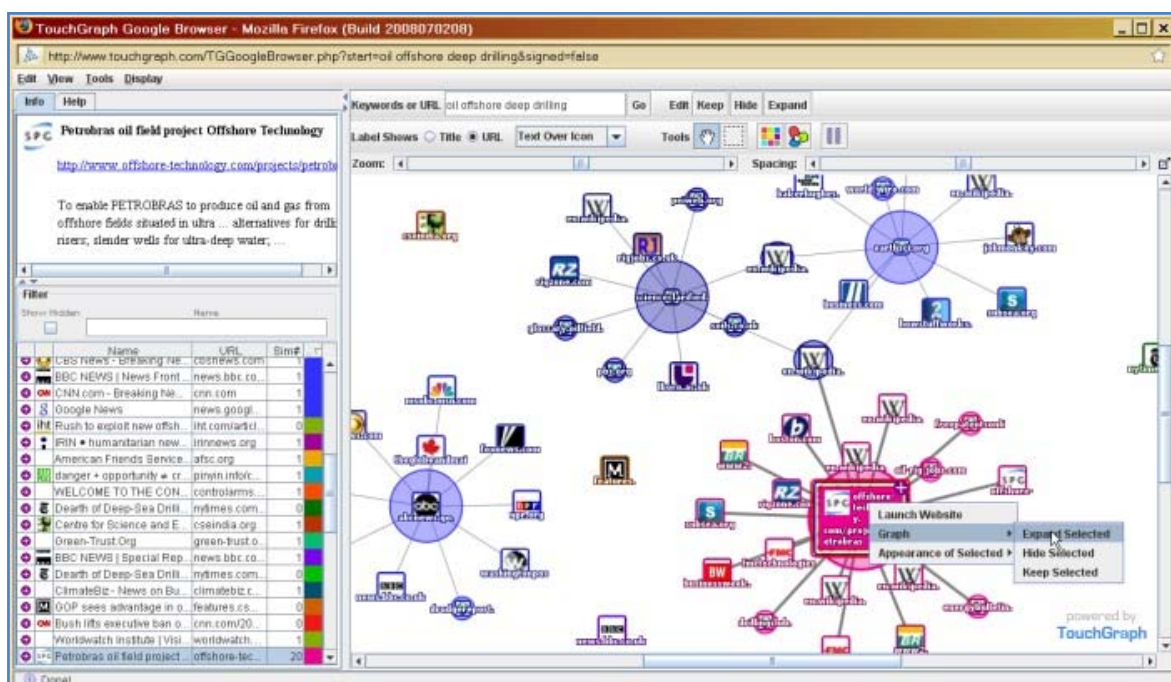
## Search Tools

### TouchGraph Google Browser

<http://www.touchgraph.com/>

Thanks to [Phil Bradley](http://philbradley.typepad.com/phil_bradleys_weblog/2008/07/touchgraph-google-browser.html) ([http://philbradley.typepad.com/phil\\_bradleys\\_weblog/2008/07/touchgraph-google-browser.html](http://philbradley.typepad.com/phil_bradleys_weblog/2008/07/touchgraph-google-browser.html)) for the alert on this search visualisation tool).

[TouchGraph](#) is a Java application that loads into your browser window and displays the networks and connections in your data or a search results list. There are three free demos that tie into Google, Amazon and Facebook. I found the last two very cumbersome to use and limited in their usefulness but the Google one is worth a try. You type in your keywords or a URL and Google's results are presented as groups of interconnected 'blobs'. You can zoom in and out, hide or expand individual groupings and filter results, although I must admit to not having yet mastered the last feature.



An interesting way to explore search results but I still prefer [Cluuz's](http://www.cluuz.com/) (<http://www.cluuz.com/>) visualisation of web based data.

### Yahoo! drops Boolean NOT operator

I go away on holiday for three weeks and on my return disaster strikes. "[Yahoo! drops Boolean support](http://www.pandia.com/sew/699-yahoo-drops-boolean-support.html)" (<http://www.pandia.com/sew/699-yahoo-drops-boolean-support.html>) was the headline that jumped out at me as I fired up my RSS reader. The story, carried by Pandia Search Engine News, refers to a report by Greg Notess in Online Magazine. I was not able to access the original in Online so had to make do with Pandia's summary and test out the claims myself.

It turns out that it is only the NOT operator that has vanished. Not such a catastrophe after all as you can still place a minus sign in front of the term that you want excluded from the search results. The posting goes on to say that nesting with parentheses has also been dumped. That was definitely not the case when I carried out test searches on both the UK and US versions of Yahoo. All of my nested Boolean searches worked, once I had replaced the NOTs with minus signs.

I shall wait a while before assuming that the change is permanent. Earlier this year, the link and linkdomain commands went AWOL. They were no longer available on the main Yahoo sites but

still worked on AltaVista. After a couple of days they reappeared on Yahoo as suddenly as they had vanished. Hopefully NOT will be back - it is easier to remember to use only Boolean operators rather than a mixture of Boolean and math signs - and fingers crossed that none of the other commands disappear.

## AllPlus

[AllPlus](http://www.allplus.com) (<http://www.allplus.com>) is yet another search tool that searches Ask, Google, Live and Yahoo in one go. As well as web pages, you can also search the News, Image, Video and Blog databases of the search engines. Moreover is added to the mix for News. In addition to displaying the results for your selected type of information, it gives you items from the other resources on the right hand side of the page. On the left hand side of the screen, it organises your results into a 'cluster tree' and you can also view a 'cluster graph'. Clicking on the topic in the tree or graph brings the pages in that cluster to the top of the results list and is a useful way of focussing your search.

## Cuil not so cool

[Cuil](http://www.cuil.com/) (<http://www.cuil.com/>), pronounced 'cool', claims to have the largest database of all the search engines and more relevant results. A detailed article on the company's background, together with a critique of its PR, can be found on Search Engine Land - [Cuil Launches - Can This Search Start-Up Really Best Google?](http://searchengineland.com/080728-000100.php) (<http://searchengineland.com/080728-000100.php>)

Cuil's index of 120 billion web pages might indeed make it the most comprehensive search engine on the web. Google no longer discloses the size of its index, although it claims to "know about a trillion unique web pages". Techcrunch has put Cuil through its paces ([Google Beats Cuil Hands Down In Size And Relevance, But That Isn't The Whole Story](http://www.techcrunch.com/2008/07/27/google-beats-cuil-hands-down-in-size-and-relevance-but-that-isnt-the-whole-story/) <http://www.techcrunch.com/2008/07/27/google-beats-cuil-hands-down-in-size-and-relevance-but-that-isnt-the-whole-story/>) and says:

"It seems pretty clear that Google's index of web pages is significantly larger than Cuil's unless we're randomly choosing the wrong queries. Based on the queries above, Google is averaging nearly 10x the number of results of Cuil."

Many of Techcrunch's searches are single word searches; Google's hit count is never accurate for those and is usually grossly overestimated. Add in more terms and the hit count starts to approach

reality. My test searches found approximately the same number of pages in both Cuil and Google, apart from those queries that had more than three terms. For example, my standard 'gin vodka sales UK' search came up with:

We didn't find any results for "gin vodka sales UK"

Some reasons might be...

- a typo. Please check your spelling.
- your search includes a term that is very rare. Try to find a more common substitute.
- too many search terms. Please try fewer terms.

Finally, try to think of different words to describe your search.

This is a simple search for which all of the other major search engines find a highly relevant answer from the Gin and Vodka Association (yes there really is one). Reduce the number of keywords to a maximum of three and Cuil recovers but returns irrelevant results. It looks as though Cuil cannot cope with long search queries and this could be related to the way it ranks pages. It claims not to just catalogue keywords on a site and then rank the site based on its importance; it also tries to understand how words are related. This seems to work best with one or two word queries. Results are shown in three columns and on the right hand side of the screen it offers options to refine your results by category. Apparently, Cuil also recommends related searches in tabs across the top of search results but I never saw these in any of my test searches. Oh, and most of the images that appear next to some of the results appear to have been picked at random and are not linked to the page in any way.

So far I was not impressed and for me the final straw was the total absence of any advanced search features. No file format searching, no limiting to media types, no domain or site searching. I could be charitable and attribute the absence of results for multi term strategies to the service being new and not being able to cope with the traffic generated by the launch announcement. The relevance of the results from my test searches, though, was not sufficiently high for me to be able to dispense with advanced search. Cuil is going to have to try a lot harder if it is to persuade me to return. At present this is not a Google beater and neither does it out-perform the other top search

### Top Search Tips

I ran another advanced search workshop (Google and Beyond) for UKeIG on June 11th, this time in London. Twenty people attended the event and came up with the following list of top search tips at the end of the day.

**1. Use the Advanced Search screen.** There are lots of goodies to be found on the advanced search screens: options for focussing your search by file format (e.g. xls for data and statistics, ppt for expert presentations, pdf for industry or government reports); site and domain search to limit your search to just one web site or a type of organisation (e.g. UK government, US academic); and in Google there is a numeric range search.

**2. Google Custom Search Engines (Google CSE)** at <http://www.google.com/coop/cse/>. This made its first appearance in the [Top Tips from the Liverpool workshop](#) earlier this year. Ideal for building collections of sites that you regularly search, to create a searchable subject list, or to offer your users a more focused search option.

**3. See what Google does with your search string.**

a) If you use the default search box and Google comes back with odd results, click on Advanced Search to see what it has done with your search terms.

b) If you use the Advanced Search screen and fill in the boxes, see how Google formats the search strategy by looking the search box at the top of the results page. By learning the

commands and prefixes you can build more specific searches more quickly on the default search page.

**4. Cached copies.** Look at the search engines cached copy of a web page if you can't find your search terms in the document or if the page is nothing like the description in the results list. You will see the version of the page that has been used by the search engine for indexing and with your terms highlighted.

**5. Use tools such as [Intelways](http://www.intelways.com/) and [Zuula](http://www.zuula.com/)** for quick and easy access to a wide range of search tools covering different types of information. Enter your search once, click on the tab for the type of resource for which you are searching (video, images, reference, news etc.), and then work your way through the list of search engines.

**6. [Alacrawiki](http://www.alacrawiki.com/)**. The Alacra Spotlights section is a good starting point for evaluated sites and information on industry sectors. It is also a good example of what to look for when assessing the quality of a wiki and how easy it is for anyone to edit the pages. In the Spotlights sections there is no edit option, not even if you register for an account and login. Only the Alacra editors can edit the pages.

**7. Open access journals.** Google Scholar sometimes leads you to copies of journal articles in institutional repositories and open access journals, but there are also directories of open access journals. For example: <http://www.doaj.org/>, <http://www.wsis-si.org/oa-journals.html>, <http://www.abc.chemistry.bsu.by/current/fulltext.htm>. This is not my area of expertise so comments on other directories are welcome.

**8. Social bookmarking sites.** Try social bookmarking sites, not only for creating your evaluated lists of sites but for searching other peoples. For example [FURL](http://www.furl.net/), [Del.icio.us](http://del.icio.us/), [Connotea](http://www.connotea.org/), [2Collab](http://www.2collab.com/). Connotea (owned by the Nature Publishing Group) and 2Collab (owned by Elsevier) are aimed at researchers and scientists.

**9. Search results visualisation.** Try out some of the newer search tools that present results and search options in a different way. For example [Cluuz](#), [Kartoo](#), [Kvisu](#), [Quintura](#). [Some of the participants specifically mentioned Cluuz and Kvisu].

**10. The Internet Archive (Wayback Machine)** at <http://www.archive.org/> for pages, sites and documents that have disappeared. Ideal for tracking down lost documents, seeing how organisations presented themselves on the Web in the past, and for collecting evidence for a legal case (e.g. 'passing off', copyright infringement).

---

## Internet Librarian International



Celebrating its 10th birthday this year, Internet Librarian International 2008 takes place at the Novotel London West on the 16<sup>th</sup> and 17<sup>th</sup> October 2008.

The Thursday keynote speaker is Guy Cloutier speaking on people at the heart of innovation and Friday's is the Shanachie Story presented by Erik Boekesteijn, Jaap van de Geer and Geert van den Boogaard from Delft Public Library.

### ILI Early Bird and CILIP discounts

Book before 24<sup>th</sup> September and you will benefit from the Early Bird discount. CILIP members are entitled to a special discount. Further details of the programme and how to book a place are on the ILI web site at <http://www.internet-librarian.com/>

## Online Information 2008



December may seem far away but we are already planning for the annual UK online jamboree at Olympia. The dates for your diary are December 2<sup>nd</sup>- 4<sup>th</sup>.

Online Information is **the** event for finding out what is happening in electronic information. Details of the conference are available at <http://www.online-information.co.uk/>.

As well as the main conference there is an exhibition with a free seminar and masterclass programme.

---

## Information Resources

### Guernsey Registry

<http://www.guernseyregistry.com/ccm/portal/>

The new [Guernsey Registry](#) was launched on 1 July 2008 following a review of the Companies (Guernsey) Law and the existing company processes. The Registry also includes the Bailiwick of Guernsey Intellectual Property Office which administers all Trade Mark and Design registrations and activities, and provides information on the Intellectual Property Laws enacted in the Bailiwick. As well as providing e-filing services for companies, there is a [searchable database](#) (<https://www.greg.gg/webCompSearch.aspx>). You can search by registered number or company name (current and previous). Company number, name, type, status, registered address, activity, and date registered are available free of charge. Company profiles, certificates of good standing and of incorporation are priced as are the annual reports. There is a link to the Companies (Guernsey) Law 2008, which can be downloaded free of charge. Be warned - it is 569 pages long!

### Renewable Fuels Agency (RFA)

The [Renewable Fuels Agency \(RFA\)](#) (<http://www.dft.gov.uk/rfa/>) has been set up by the UK Government to implement the [Renewable Transport Fuel Obligation \(RTFO\)](#) (<http://www.dft.gov.uk/rfa/aboutthertfo.cfm>), which came into force on 1st April 2008. The RTFO obliges fossil fuel suppliers to ensure that by 2010 biofuels account for 5% by volume of the fuel supplied on UK forecourts. The purpose of the RTFO is to "reduce the UK's contribution to climate change and its reliance on fossil fuels". The RFA publishes updates on the progress of the RTFO. These include monthly reports on progress on achieving compliance with sustainability criteria and quarterly reports to the Department for Transport and annual reports to parliament. All reports are available on the web site.

With serious questions being raised about the impact of biofuels on food prices, farming and the environment in general, it will be interesting to see how long this all carries on. The RFA's first monthly report has just been published and covers the period 15th April - 14th May 2008. The press release contains some good summary statistics for those of us who need to get hold of such data in a hurry. There are 'associated files' (PDF and an Excel spreadsheet) that contain more detailed information.

### UK Coal Authority

The UK [Coal Authority](#) (<http://www.coal.gov.uk>) aims "to facilitate the proper exploitation of the Nation's coal resources, whilst providing information and addressing liabilities for which the Authority is responsible, in a professional, efficient and open manner".

If you own or are purchasing a property in a coal mining area there is a search service that will check whether or not the property might be affected by past or existing mining activity. The gazetteer (under Information Services) gives an indication of places in Great Britain that may, or may not, require a mining search to be performed. If you are thinking of moving within the UK it would be as well to check this site as well as the Environment Agency's '[In Your Backyard](http://www.environment-agency.gov.uk/maps/)' (<http://www.environment-agency.gov.uk/maps/>), which tells you if the area is at risk of flood and whether there's a landfill site near your house.

The Coal Authority site also has reports on mining activity in the UK and statistics on production. There are links to sites that cover coal mining related topics and information on technologies. The latter includes coal gasification, cleaner coal technologies and environmental issues. These technology 'fact sheets' are a good starting point if you need to get up to speed on the industry jargon and how it all works.

## Say no to 0870 - alternative telephone numbers

<http://www.sayno0870.com/>

"Say No to 0870" has been around for a while but it is only recently that I have had to make serious use of its database. I have been researching hotel business meeting facilities and was amazed at how many still use 0870 numbers. My middle name ought to be 'Scrooge' because I resent having to pay for a phone call when I could be using my mobile's very generous free minutes (geographic numbers only) allocation. When that is used up, I have various VOIP services that I can use for a few pence.

This web site gives alternative geographic numbers for 0800, 0808, 0844, 0845, 0870 and 0871 numbers. Search on the company name or its non-geographic number and you are offered alternative geographic numbers. Sources include web sites and contributions from users. You may wonder why freephone numbers are included: these often take you to a central call centre or redirect you to a branch location based on your telephone area code (if available). You then spend what seems to be forever trying to get connected to the right branch or location. "No I don't want to book a room .... and, yes, I know I am phoning from Newcastle but I want to contact the Events Sales Manager for Whizzo Hotels in Oxford!"

---

## Gizmo of the Month

### Tag clouds for analysing documents

CV not getting you those all important interviews? Nobody answering your job advert? Or perhaps your corporate publicity is not doing the biz? Processing your document through a tag cloud generator might give you a clue as to where you are going wrong. Sue Hill gave a presentation at a recent City Information Group open day on CPD and skills. In passing she mentioned that they sometimes run a CV or job description through a tag cloud generator to show people why their lovingly created prose is way off the mark. The tag cloud brings to the fore your most used terms and it can be a shock to discover that you have placed the emphasis in totally the wrong area. It then struck me that you could do this with any form of literature - a web page, training publicity, membership recruitment forms.

There are dozens, if not hundreds, of tag cloud generators on the Web and most of them are free. For starters try [Wordle](http://www.wordle.net/) (<http://www.wordle.net/>), [Tagcrowd](http://tagcrowd.com/) (<http://tagcrowd.com/>), or [Tag Cloud Generator](http://www.tag-cloud.de/) (<http://www.tag-cloud.de/>). The example below is a tag cloud of the UKeiG home page generated by Wordle.



---

## Meetings and Workshops

**Workshop:** Being a Digital Citizen

**Date:** Wednesday, 10<sup>th</sup> September 2008, 2 pm - 4.30 pm

**Venue:** Weetwood Hall, Otley Road, Leeds, LS16 5PS

**Organiser:** Work AnyWare Ltd

**Presenter:** Karen Blakeman

Which Web 2.0 applications are seriously being used in business - what actually works, and what could be destined for Gartner's 'Trough of disillusionment' never to be seen again? This workshop will explain the jargon and look at how the technologies can be used effectively to promote yourself, your business and to keep yourself up to date with what is happening in your area of expertise. Participants are encouraged to come along with questions and issues that they would like to discuss. At the end of the session, the group will be asked to nominate Web 2.0 applications for the "What really works" awards. Categories are 'Yes', 'No', 'Maybe' and the John McEnroe "You cannot be serious" award.

**URL:** <http://www.workanyware.co.uk/digitalcitizen/>

**Workshop:** Market Research on the Web

**Date:** Tuesday 21 October 2008

**Venue:** Eddie Davies Library, Manchester Business School

**Organiser:** Manchester Business School

**Presenter:** Karen Blakeman

Practical guidance is offered for finding and evaluating statistical, market research and industry information, covering free and fee based services. Suitable for researchers compiling analyses of markets, looking for potential partners, clients, suppliers or who are responsible for competitor intelligence

**URL:** <http://www.mbs.ac.uk/bis-training>

**Workshop:** Business Information on the Internet

**Date:** Wednesday, 19<sup>th</sup> November 2008

**Venue:** Eddie Davies Library, Manchester Business School

**Organiser:** Manchester Business School

**Presenter:** Karen Blakeman

**URL:** <http://www.mbs.ac.uk/bis-training>



## Contact Information

Karen Blakeman, RBA Information Services  
UK Tel: 0118 947 2256  
Int. Tel: +44 118 947 2256

UK Fax: 020 8020 0253  
Int. Fax: +44 20 8020 0253

Address: 88 Star Road, Caversham, Berks RG4 5BE, UK

## Archives

TFTTR archives: <http://www.rba.co.uk/tftr/archives/>

## Subscribe and Unsubscribe

To subscribe to the newsletter fill in the online registration form at <http://www.rba.co.uk/tftr/>

To unsubscribe, use the registration form at <http://www.rba.co.uk/tftr/> and check the unsubscribe radio button.

If you wish to change your subscription from the PDF to the Plain Text version, send an email to [tftr@rba.co.uk](mailto:tftr@rba.co.uk) with "Plain Text Version Please" in the subject line.

---

## Privacy Statement

Subscribers' details are used only to enable distribution of the newsletter **Tales from the Terminal Room**. The subscriber list is not used for any other purpose, nor will it be disclosed by RBA Information Services or made available in any form to any other individual, organization or company.

---

## Copyright



Tales from the Terminal Room by Karen Blakeman is licensed under a Creative Commons Attribution-Non-Commercial 3.0 Unported License.

For permissions beyond the scope of this license contact Karen Blakeman by email at [Karen.Blakeman@rba.co.uk](mailto:Karen.Blakeman@rba.co.uk) or via <http://www.rba.co.uk/about/contactkb.htm>

---