



## Tales from the Terminal Room March 2011, Issue No. 95

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
    - Google decides that coots are really lions
    - Google includes your social circle in search
  - UK crime data as clear as mud
  - Business Information
    - M&A Portal – free information from BvD
    - Company registers: Slovenia and Croatia
  - Twitter Notes
  - Meetings and Seminars
    - All about Google, Newcastle
-

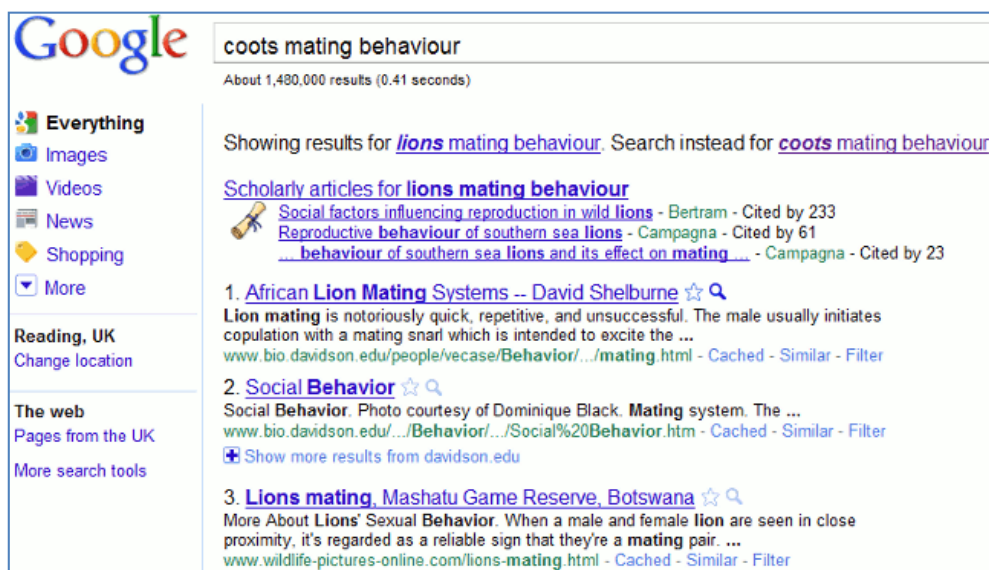
## Search tools

### Google decides that coots are really lions

First of all let us make sure we all know the difference between lions and coots. As far as I can recall, lions are huge, snarly, growly, land animals that are liable to eat you if you cross their path. This appears to be confirmed by Wikipedia (<http://en.wikipedia.org/wiki/Lions>) but of course Wikipedia could be wrong. Coots are medium sized water birds (<http://en.wikipedia.org/wiki/Coots>) and the worst that could befall you should you antagonise one is a severe pecking.

I was walking by the Thames in Caversham today and took several photos of the birds on the river. One was of two coots having what appeared to be a minor domestic or an argument over territory, but a friend suggested to me that what I saw was coot mating behaviour. What do you do in a situation such as this? You Google.

My search on coots mating behaviour came up with:



Where the [expletive deleted] did the lions come from? I just do not understand how Google managed to replace coots with lions. One is a water bird with wings, feathers, and a beak and the other a large, aggressive land mammal with fur, claws and big teeth. But Google, yet again, has decided to go off and run its own search. (See my posting [Oi! Google – you have seriously overstepped the mark](http://www.rba.co.uk/wordpress/2011/01/03/oi-google-you-have-seriously-overstepped-the-mark/))

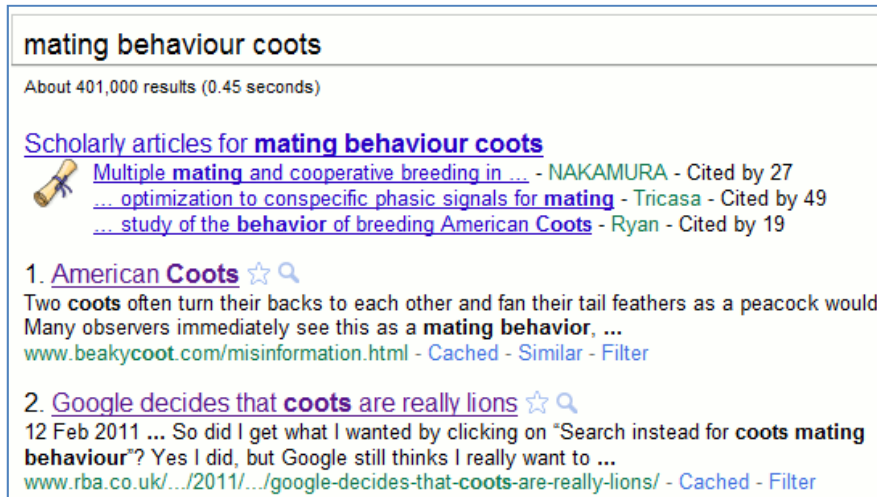
So did I get what I wanted by clicking on “Search instead for coots mating behaviour”? Yes I did, but Google still thinks I really want to search for lions and asks “Did you mean: lions mating behaviour”. Using the plus sign before coots also gave me the expected results but Google persisted and asked “Did you mean +lions mating behaviour”.

The first response to my blog posting on this was a comment from Arthur Weiss (<http://www.rba.co.uk/wordpress/2011/02/12/google-decides-that-coots-are-really-lions/comment-page-1/#comment-14207>). He suggested that Google was treating coots and lions as synonyms (both are living creatures). I thought that was pushing synonyms too far even for Google. (Sorry, Arthur). I then had two comments in quick succession from Susanna Winter via Twitter (@Mrs\_Figaro).

The first is at [http://twitter.com/Mrs\\_Figaro/statuses/36714410223341568](http://twitter.com/Mrs_Figaro/statuses/36714410223341568):

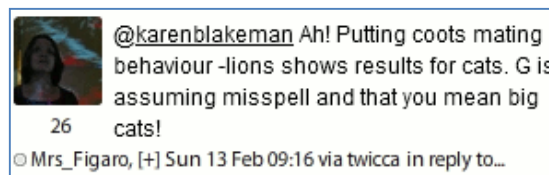


So moving coots from the beginning to the end of the strategy resulted in an exact match and not a single lion in sight:

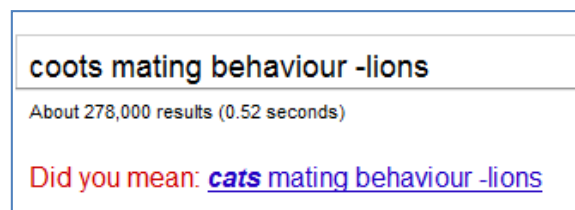


Changing the order of the search terms is a trick I often use to change the order of my results or bring up pages that might be buried in the hundreds or thousands, but I have never seen such a dramatic change such as this.

Susanna's search strategy 'coots feeding behaviour' came up with an exact match that muddled the waters even more. Perhaps there is a search frequency algorithm coming into play? Are there more searches for lions mating behaviour than for coots, but not lions **feeding** behaviour? I am not convinced that this explains Google's insistence on looking for lions rather than our animal of choice. Susanna's next tweet suggests what is going on ([http://twitter.com/Mrs\\_Figaro/statuses/36715389190676480](http://twitter.com/Mrs_Figaro/statuses/36715389190676480)):



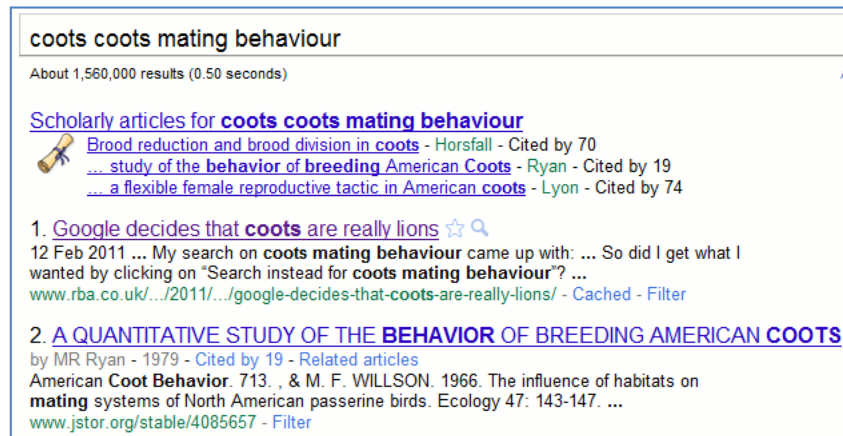
Type in the search and you get:



So Arthur **was** on the right track. (My apologies, Arthur). What probably happened with our search is, as Susanna said, that Google first assumed a typographical error and then did a synonym search on cats. What puzzles me, though, is how Google arrived at cats from coots. Surely coyotes or goats would be nearer?

I have two final variations on our search to confuse you even further.

The first is repeating coots at the start of the strategy. An exact match:



Now move one of the 'coots' to the end of the strategy and Google asks "Did you mean lions mating behaviour coots":



Initially, this seems an amusing confusion on Google's part but it does have serious implications for researchers. Coots and lions are very different animals and one would hope that people would not confuse the two. But let us assume that Google decided to include other water birds in the search as synonyms and you were in a hurry. Would you spot what had happened? And if Google is going to start messing around with company or people's names you are going to have to look very carefully at the results to ensure Google has not looked for variations. Placing names within quote marks no longer ensures an exact match (see "Oi! Google - you have seriously overstepped the mark" in the January 2011 issue of TFFTTR or the blog posting at <http://www.rba.co.uk/wordpress/2011/01/03/oi-google-you-have-seriously-overstepped-the-mark/>). At present using the plus sign before terms and phrases seems to work, although Google still tries to push what it has decided is really your intended search.

And the photo that started it all? That can be found on my Flickr account at <http://www.flickr.com/photos/rbainfo/5438769506/>. I think you will agree that coots are very different from lions ([http://commons.wikimedia.org/wiki/File:P\\_I\\_Bleyenberghi.jpg](http://commons.wikimedia.org/wiki/File:P_I_Bleyenberghi.jpg))

## Google includes your social circle in search

Google has been including search results from your social circle for quite a while. If you are logged in to a Google account it sometimes includes a box either at the bottom or in the middle of your search results page with a couple of "Results from people in your social circle":

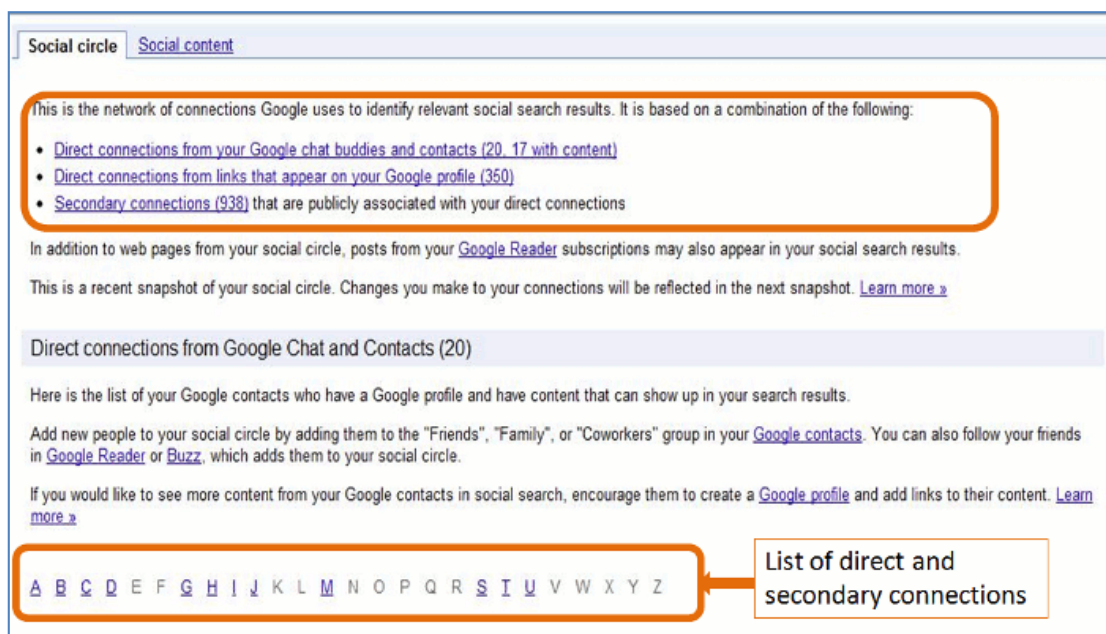


Click on the link and it will show you more from your circle.

If you are using Google.com and open up the search options in the side bar to the left of your results there is a "Social" option that will do exactly the same thing.



Want to know who is in your social circle? Head straight to your Google dashboard at <http://www.google.com/dashboard> and scroll down to Social Circle. This not only tells you which of your social networks Google is using but also lists who.



Any social networks that you have mentioned in, for example, your Blogger profile or your general Google profile such as Twitter will be included as will contacts in Google Reader, Google Buzz, Google Contacts and Picasa. In addition to your direct connections Google also searches the content of secondary connections that are publicly associated with your direct connections. The A-Z reveals all!

Until now the Social option has been kept separate but Google has started integrating all of these results with the rest of your search starting with Google.com (see Official Google Blog: An update to Google Social Search <http://googleblog.blogspot.com/2011/02/update-to-google-social-search.html>). So if you are using Google.com and are logged into your Google account you will now start seeing results automatically from your social circle.



Underneath the social search entry it will say something like "Joe Bloggs shared this on Twitter". Hover over this and it tell you how you are connected.

I frequently use social media to search for information and advice but I much prefer to choose when and how I incorporate it into my strategy. I do not always want Google unilaterally deciding to add opinions from all and sundry in my network. Remember, this is not just the direct connections that **you** have chosen to make but others who are connected to them and over which you have no control. The only way you can shut this off is to log out of your Google account. I will reserve final judgement until I have used it a bit more but after only 2 hours of experimentation I am already finding that some searches are dominated by results from Twitter, Google Reader and Flickr. The results look even more cluttered than before and I fear it gives Google even more opportunity to completely mess up the search.

---

## UK crime data as clear as mud

I'm a nosy neighbour. I like to know what's going on in my area: who's bought the house next door, local planning applications, is there any dodgy activity going on? My husband and I are both self employed so there is usually at least one of us out and about in Caversham during the day. That means we have the chance to chat with our local postman, workmen digging up the road, Police Community Support Officers doing their rounds and with people in the local shops, bank and post office. Crime, not surprisingly, is a major topic on our "watch list" and just over two years ago police forces in England and Wales started to provide access to local crime statistics via online maps. The service allowed you to drill down to ward level and view trends in burglary, robbery, theft, vehicle crime, violent crime and anti-social behaviour.

The format varied from one police force to another. For example Thames Valley Police provided a basic map and tables of data. Others such as the Metropolitan Police included additional graphical representation of the statistics, for example bar charts.

None of them pinned down incidents to individual streets or addresses but they did give you an idea of the level of crime in a particular neighbourhood, how it compared with the same period the previous year and whether it was high, above average, average, below average, low or no crime. They were short, though, on detailed definitions of what each category of crime included. I looked at these maps out of personal curiosity rather than using them for any serious business

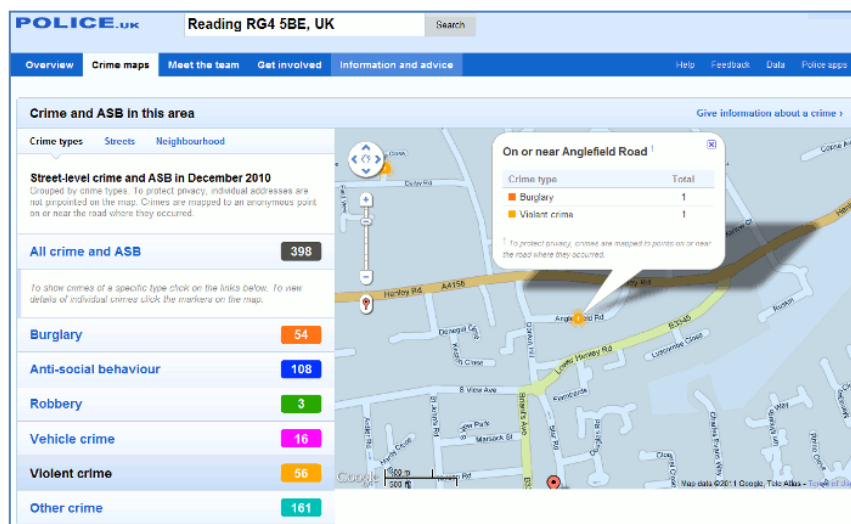
application, and I made certain assumptions such as murder being included under ‘Violence against the person’. That may not have been the case.

Some police forces placed obvious links to the information on their home pages whilst others buried the data in obscure corners of their web sites. The crime maps were then all moved to the CrimeMapper web site – the Thames Valley Police map can still be seen at <http://maps.police.uk/view/thames-valley/> – but that has now been integrated into the Police UK website, which “includes street-level crime data and many other enhancements”.

All you have to do is go to <http://www.police.uk/>, type in your postcode, town, village or street into the search box and “get instant access to street-level crime maps and data, as well as details of your local policing team and beat meetings”. The first screen looks good with news of local meetings, events, recent tweets, YouTube videos and – as the home page promised – information on my local policing team.



When I focus on the map to look at the detail there are markers for the location of the crimes and clicking on them gives you a brief description of the crime:



In this example, the detail box had listed two crimes “on or near Anglefield Road” and this is where I started to become confused. Were the burglary and the violent crime part of the same incident or totally separate? Furthermore, if you look in the left hand panel of the screen you will see “To protect privacy, individual addresses are not pinpointed on the map. Crimes are mapped to an anonymous point on or near the road where they occurred.” Fair enough, but I would like to know how near ‘near’ is. 100, 200, 400 yards? Half a mile, a mile? And does the focus shift from one street to another from one month to the next? If it stays put then a street could gain a reputation

from crime that it does not deserve but if it shifts there is no way one can compare data from one month or year to another, which brings me to my next question.

Why is there only one month's data? Previous versions of the crime maps gave you three months data for the current and the previous year for comparison. It could be because this site has only just started and the data is not comparable with that on previous maps. There is nothing about this in the Help section of the Police UK site but the Guardian reports:

*"police forces have indicated that whenever a new set of data is uploaded – probably each month – the previous set will be removed from public view, making comparisons impossible unless outside developers actively store it."* (Crime maps are 'worse than useless', claim developers [http://www.guardian.co.uk/technology/2011/feb/02/uk-crime-maps-developers-unhappy?CMP=twl\\_iph](http://www.guardian.co.uk/technology/2011/feb/02/uk-crime-maps-developers-unhappy?CMP=twl_iph))

In fact, January's data is now available for download as well as December's but there are still no trends or comparisons. Again, this may be because they do not feel that there is enough data yet to do this but it could mean that if you want to run comparisons over time **you** will have to download the files and store them on your own system each month, or find someone else who is already doing it.

The Guardian article also says:

*"the Information Commissioner's Office (ICO) advised that tying crime reports to postcodes or streets with fewer than 12 addresses would render the individuals involved too identifiable. The police have also decided to remove data about murders or sexual assaults".*

With respect to the latter the help file on the Police UK site suggests otherwise:

*"Crimes have been grouped into six categories following advice from the Information Commissioner's Office. This doesn't mean that the crimes listed under 'other' are not seen as important. Rather it ensures that for some of the more sensitive crimes there is even greater privacy for the victims."*

So which is it: murders and sexual assaults are not included at all or aggregated under "other"? Jonathan Raper says on his blog Placr News ("Five reasons to be cautious about street level crime data" <http://placr.co.uk/blog/2011/02/five-reasons-to-be-cautious-about-street-level-crime-data/>)

*"Some data is redacted eg sexual offences, murder. The Metropolitan Police has already released this data to ward level though... and it is easy to cross-reference one murder in one ward to reports in the local press at the same time"*

Data visualisations and mashups are becoming increasingly popular making it considerably easier to assess a situation and view trends. The Guardian Datablog (<http://www.guardian.co.uk/news/datablog>) for example, encourages people to take data sets, mash them up and create their own visualisations, and upload a screen shot to the Guardian Datastore on Flickr (<http://www.flickr.com/groups/1115946@N24/>). It is vital, though, that the source of the data, whether the full data set or just a selection has been used, and whether or not it is going to be updated is clearly spelt out. All too often one or even all of these are missing from the accompanying notes, and in some cases there are no notes at all.

An example of **good** practice is "UK transport mapped: Every bus stop, train station, ferry port and taxi rank in Britain" (<http://www.guardian.co.uk/news/datablog/2010/sep/27/uk-transport-national-public-data-repository> ). The posting clearly states the source (<http://data.gov.uk/dataset/nptdr>) and its coverage.



*“A snapshot of every public transport journey in Great Britain for a selected week in October each year. The dataset is compiled with information from many sources, including local public transport information from each of the traveline regions, also coach services from the national coach services database and rail information from the Association of Train Operating Companies”*

It then goes on to specify the time period (5-11 October, 2009) and the tools that were used to create the visualisation.

Another is the “Live map of London Underground trains” (<http://traintimes.org.uk/map/tube/>). This shows “*all trains on the London Underground network in approximately real time*“. The source is a live data feed from Transport for London (TfL) and the notes state that a “*small number of stations are misplaced or missing; occasional trains behave oddly; some H&C and Circle stations are missing in the TfL feed.*” It would be helpful to have a list of those missing stations, but the site has at least brought the issue of potential missing data to the users’ attention.

Returning to the Police.uk crime data, there are three major problems with the site for me as a researcher:

1. Are all crimes included in the database, or are some such as murders and sexual assaults excluded altogether or aggregated under “other”? More detailed and unambiguous scope notes on this please.
2. The street data level is useless. The markers are not exact locations but “near”, there is no definition of “near”, no information on how the position of the marker is calculated or the geographic radius that it covers. It would be better to return to aggregated data at the ward level.
3. There are no options comparing time periods. You have to devise some way of doing that yourself or find a site that is already doing it.

The new crime data web site is a retrograde step. We need transparency and clarity rather than the muddle and confusion that has been generated by the lack of information on what is being provided.

---

## Business Information

### M&A Portal – free information from BvD

Bureau van Dijk’s M&A portal (<http://www.mandaportal.com/>) is a gateway to news, events, research and analysis on mergers and acquisitions. The information on the portal home page is free of charge and there is a free search option for tracking down deals and rumours contained in BvD’s Zephyr database. The deals can be sorted by value, date or status.

Basic information is free of charge but you can purchase the full details from the Zephyr database. The cost of the reports varies depending on the amount and type of information available. The portal also has M&A news feeds for 36 industry sectors. Registered users – registration is free – can select the sectors that interest them most and create a tailored news tab.

REFINE SEARCH RESULTS		11 Deals	Deal value - Descending
Geography			
North America (11)			
Industry - Major			
Metals & metal products (1)			
Other services (10)			
Deal types			
Acquisition (1)			
1.	Deal headline Target Acquiror Vendor Deal type	Gryphon Gold confirms announcement by Bolero Resources (De Gryphon Gold Corporation (Gold mine developer - United States Of A Bolero Resources Corporation (Precious and base metal properties e	Acquisition 100%
2.	Deal headline Target Acquiror Vendor Deal type	Bolero Resources issues units to MineralFields (Deal No 160124 Bolero Resources Corporation (Precious and base metal properties e **** (Investment services - Canada)	Minority stake 11.39%
3.	Deal headline Target Acquiror Vendor Deal type	Bolero Resources completes offering with Mineralfields Group (De Bolero Resources Corporation (Precious and base metal properties e **** (Investment services - Canada)	Minority stake 19.301%
4.	Deal headline Target Acquiror Vendor Deal type	Bolero Resources closes non-brokered private placement (Deal Bolero Resources Corporation (Precious and base metal properties e	Minority stake 14.514%

## Company registers: Slovenia and Croatia

I have added two more company registers – Slovenia and Croatia – to my Official Company Registers page at <http://www.rba.co.uk/sources/registers.htm>.

### Slovenia

[AJPES – Agency of the Republic of Slovenia for Public Legal Records and Related Services](http://www.ajpes.si/) (<http://www.ajpes.si/>)

AJPES provides access to the Slovenian Business Register online (ePRS), and Annual Reports of companies, cooperatives and sole proprietors (JOLP) free of charge. ePRS currently has around 211,000 entries and includes companies (partnerships and corporations), sole proprietors, legal entities governed by private law, societies, natural persons performing registered or regulated activities, subsidiaries and other divisions of business entities and main offices of foreign business entities.

You can search on a range of criteria: name, identification number, tax number, address, activity, sector and legal form. The amount of information available on each business entity varies depending on its legal form but typically includes identification number, company name, tax number, details of representatives and founders, etc. Credit rating reports (eS.BON) are also available for a fee. The interface to the site is available in English, Slovene, German and Italian. The information is available in English and Slovene. Free registration is required to view the free documents.

### Croatia

[Biznet](http://www.biznet.hr/) (<http://www.biznet.hr/>) is maintained by the Croatian Chamber of Commerce and provides information and statistics on the Croatian economy, and access to the Register of Business Entities. The Biznet home page is in Croatian and the link to the Register is under [Registar poslovnih subjekata](#). You can view data for a single company or a selected set of companies grouped according to certain criteria for example industry, company size. Available information includes identification number, short and long name, county, municipality, address, phone, fax, E-mail, company size, main activity, legal form, number of employees, membership of professional associations. The Register is available in both Croatian and English and is free of charge.

Many thanks to Tomaz Lajovic for the information on both services.

Although much of the data on both sites is available in English some of the detailed reports and guides are in the local languages. Google's Chrome automatically offers to translate texts and the Google toolbar for both Internet Explorer and Firefox has a translate option. If you are not allowed to install the toolbar then you can copy and paste the text or the URL of the page you wish to translate into Google's translation tools at [http://www.google.co.uk/language\\_tools](http://www.google.co.uk/language_tools). The quality of Google's translations has improved significantly but you should still treat the results with caution. If the information is mission critical then find someone who is fluent in the language concerned to double check Google's effort.

## Twitter Notes

The following are some of my recent tweets and retweets. They are selected because they contain links to resources or announcements that may be of general interest. I have unshortened the shortened URLs.

### March 1<sup>st</sup>

Backup your stuff! "Google Glitch Disables 150,000 Gmail Accounts"

<http://mashable.com/2011/02/27/gmail-glitch/>

### February 27<sup>th</sup>

#2011PAD February 25th, 2011: Hydrogen not so green

<http://www.flickr.com/photos/rbainfo/5479552510/> Slide from Chris Rhodes (my husband) lecture to the local WI

### February 23<sup>rd</sup>

Twapperkeeper announces Planned Removal of Export & Download / API capability of twitter

archives March 20, 2011 <http://twapperkeeper.wordpress.com/2011/02/22/removal-of-export-and-download-api-capabilities/>

### February 19<sup>th</sup>

From @libram Ways I use Twitter to find, use and broadcast information « Roddy Macleod's Blog

<http://roddymacleod.wordpress.com/2011/02/18/ways-i-use-twitter-to-find-use-and-broadcast-information/>

### February 17<sup>th</sup>

Ofcom says 'P' logo must be shown for 3 seconds at start & end of TV programmes containing

product placements <http://www.out-law.com/page-11776> [Follows on from the announcement that UK TV programmes will be allowed to include product placements within the programme itself]

### February 16<sup>th</sup>

Looking at current & planned roadworks chaos in Reading over coming months

<http://www.elgin.gov.uk/> Roadworks Information [Elgin stands for **E**lectronic **G**overnment **I**nformation **N**etwork]

### February 12<sup>th</sup>

At least this volcano is easier to pronounce: RT @spiky7: Don't book any holidays this year...

<http://www.telegraph.co.uk/news/worldnews/europe/iceland/8311924/Icelandic-volcano-set-to-erupt.html> [Report on another Icelandic volcano that may be about to erupt]

### February 8<sup>th</sup>

RT @infoforenergy: IFEG website of the day: <http://energy.publicdata.eu/ee/> - graphics comparing European countries Co2 emissions, renewables & more

Interesting visualizations of where our (UK) taxes are spent "Where Does My Money Go"

<http://wheredoesmymoneygo.org/>

RT @kwiddows: Just discovered twapperkeeper has started charging if you want more than 2 archives - anyone know of an alternative service?

### February 4<sup>th</sup>

RT @NazlinBhimani: Ebooks - current stats and rankings:

<http://stephenslighthouse.com/2011/02/02/e-bookssome-current-stats-and-rankings/>

### February 3<sup>rd</sup>

RT @stephendale: Flickr Accidentally Deletes a User's 4,000 Photos and Can't Get Them Back

<http://www.observer.com/2011/tech/flickr-accidentally-deletes-users-4000-photos-and-cant-get-them-back> [In view of the above, you may want to consider Flickrcredit <http://sunkencity.org/flickrcredit>. See Phil Bradley's review at <http://sunkencity.org/flickrcredit>]

## Meetings and Seminars

**Workshop:** All about Google

**Date:** Thursday, 31 March, 2011 - 09:30 - 16:30

**Venue:** Netskills Training Suite, University of Newcastle

**Organiser:** UKeig

**Presenter:** Karen Blakeman

**Outline:** Google is collecting more data and adding more services all the time. For many people it is **the** place to start a search but more information and more search features does not necessarily mean more relevant results. Furthermore, Google's automatic personalisation and semantic search all too often adds to the confusion. This hands-on workshop will look at the latest developments in Google, how it works, and how you can focus your searches for better results.

**URL:** <http://www.ukeig.org.uk/trainingevent/all-about-google-regain-control-search-karen-blakeman>

## Contact Information

Karen Blakeman, RBA Information Services

Twittername: karenblakeman (<http://twitter.com/karenblakeman>)

Email: [Karen.blakeman@rba.co.uk](mailto:Karen.blakeman@rba.co.uk)

Tel: +44 118 947 2256

Fax: +44 20 8020 0253

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

## 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](http://creativecommons.org/licenses/by-nc/3.0/). 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>