

New Google, New Challenges

Karen Blakeman, RBA Information Services

karen.blakeman@rba.co.uk

Just when you think you've worked out how Google works it all changes, not just the commands but the whole way Google searches and displays results. Semantic search, artificial intelligence, extreme personalisation and quick answers that save you the bother of having to read any further are just the beginning. Add a dash of EU legislation and things can start to go very wrong.

It all started to go awry with the closure of Google labs and many innovative search experiments. Then slowly but surely services and search options were dropped from Google itself: starred results, page previews, time-lines, synonym search, the plus sign, translated foreign pages search and Google Reader to name but a few. The most recent to get the chop was Reading Level. Although I did not use it that often it was a nice little tool that enabled you, when necessary, to change the emphasis of results from basic, introductory information to more research biased documents. All done, as Google often says, to provide us with "a simpler, more intuitive Google experience". ([Toward a simpler, more beautiful Google.](#)) In reality, if a function is rarely used, doesn't bring in much revenue and requires significant human time and effort to maintain then it goes.

Remember that Google is a stock exchange listed company and its main aim is to make money for its shareholders. Much of that revenue comes from advertising and, yes, plenty of people do click on those ads. Google is not interested in those of us who use their services to conduct in-depth research; we rarely click on the ads, many of us have ad blockers and anyway some of our searches are so esoteric that no ads match what we are looking for anyway! Google is interested in the person looking for cheap flights to Prague; the person looking for cheap hotels in Manchester; the group of people who are attending a conference and are searching on their mobile devices for a local pub to meet up in. These are the people who will click on an ad or sponsored link occupying most of the screen that is "good enough" for their needs. It is these types of searches that make up the bulk of Google's enquiries and create revenue for Google.

What serious researchers are left with is a reduced core of commands for refining searches. Getting the most out of this limited toolkit often requires lateral thinking but it is well worth the effort. Judging from the Top Tips suggested by participants of the UKeiG Google workshops, 'site:' and 'filetype:' commands are top of the list enabling us to filter results by type of organisation (academic, NHS, government) and focusing on document type (PDFs for research papers and official documents, spreadsheets for statistics and raw data). Getting to know how to use the Google commands and search filters is the key to better, more reliable results.

For Google, mobile is where it is all happening and that is where it is concentrating most of its efforts. ([It's Official: Google Says More Searches Now On Mobile Than On Desktop.](#)) Obviously the layout of the information is different depending on whether you are using a desktop computer or a smartphone, but so is the actual content. This can present problems when trying to talk someone through a Google search because as well as past searches and browsing behaviour the results are personalised according to the type of device and operating system being used. Now Cards also seem to be a major part of Google's future mobile strategy. There are already over one hundred Cards for android users with more in the pipeline ([Google Adds 70 More Now Cards.](#)) Now Cards remind users of local traffic, the weather, special offers from local businesses, public transport departures, delays and arrival times and seemingly just about everything else without being prompted.

Some mobile search features eventually migrate to the desktop. "Facts" and information that immediately answer your question (Quick Answers) now regularly appear at the top and to the side of results. The topics range from local football results to symptoms of medical conditions. It is not clear how Google chooses the sources for these and they can change from day to day. The quality of some of them is dreadful and sometimes dangerous. I once came across a quick answer from an appalling site on food allergies. The person responsible for the site openly stated that they were not medically qualified and proceeded to give "advice" that was misleading and could have caused a life threatening reaction in an allergic individual. Looking at the site in more detail it became clear that its main purpose was to sell food diaries, books, and allergy testing kits. At least with that Quick Answer it was possible to click through to the original website.

For many answers no source is given and when tackled about this Google said that it doesn't provide a source link when the information is basic factual data and can be found in multiple places on the Internet. ([When Google Shows A Source Or Credit For Quick Answers & Knowledge Graph.](#)) Unfortunately, repeating something many times does not necessarily make it true and there are many examples of Google getting it wrong. ([The quality of Google's results is very strained.](#)) Never,

never trust Google's Quick Answers or so-called facts. Always check them with another source.

We will probably be seeing more "facts" in the future and it has been suggested Google could start to use them as an additional ranking factor. If a page contains data and information that is repeated across the web then it will be deemed to be authoritative. If it has conflicting statements it could be regarded as untrustworthy. So where does that leave a comprehensive scientific review article that tries to give a balanced overview of a topic and cites contradictory evidence? ([Google wants to rank websites based on facts not links - 28 February 2015 - New Scientist](#) and [Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources.](#))

Then there is Google's habit of dropping terms from our searches without asking and, on occasion, completely rewriting your search. I wrote a blog posting about the problem that was picked up by Google's Dan Russell ([Dear Google, stop messing with my search.](#)) He commented:

"When you do a multi-term query on Google (even with quoted terms), the algorithm sometimes backs-off from hard ANDing all of the terms together. It's a kind of "soft" backoff. Why? Because it's clear that people will often write long queries (with anywhere from 5 to 10 terms) for which there are no results. Google will then selectively remove the terms that are the lowest frequency to give you some results (rather than none). Bear in mind that 99% of searchers have no idea why they'd want to hard AND, and just get frustrated when they get no results... But I see what you mean about wanting to know if there are NO hits to a given query. I'll pass this information along to the Google design team and see if we can't do something with this."

Google's reaction was to introduce Verbatim, which runs your search with no omissions and no variations on your terms. It is one of the most powerful tools for regaining control of the search process and essential as Google's search rewrites become increasingly bizarre.

There are times when no matter how many advanced commands you use you could still be missing important information as court cases and legislation begin to dictate what can and cannot be displayed. The EU ruling on the so-called [right to be forgotten](#) gives individuals the right to ask for links to information about themselves that is inaccurate, excessive or out of date to be removed from search results. This affects searches that include a person's name and all search engines that have offices or operate in EU countries plus Norway, Switzerland, Iceland and Lichtenstein. An individual has to send in a request to the search engine stating which links they want removed and why. It is then up to the search engine to

decide whether or not the request is valid. If the search engine refuses the request the individual is entitled to take the matter further by applying to the relevant Data Protection agency.

According to its own [transparency report](#) Google has so far received 265,571 requests to remove about 964,533 links, of which 41.3% have been removed from its search results. Remember, it is just the links to the information that are removed; the information itself still remains somewhere on the Internet. You just have to know how and where to look for it. The easiest way to circumvent the ruling is to use a non-European version of Google such as Google.com or Google.ca that currently give you access to the full list of results. Do not assume that this will always be an option as there are some who would like the delisting to be worldwide.

Should you be using Google or general search tools in the first place? For example if you are interested in directorships held by a person, then the official company registers such as Companies House should be top of your list and won't be affected by the right to be forgotten. And, as was said earlier, never trust Google's facts and Quick Answers. Find out about the specialist sites that cover your subject area and type of information (news, statistics, company financials, government data, research papers etc.).

So what of the future? Google's Eric Schmidt recently said that the Internet will disappear by which he meant that it will become such an integral part of every device and appliance, and everything we do that we won't consider it as being something separate. Cars, washing machines, refrigerators that warn you when something starts to go off, and home environmental controls are just a few examples of this Internet of Things. Google has invested heavily in advanced technologies such as artificial intelligence and it is already being used in some areas of search, as part of robotics projects, driverless cars and even in the home. Nest, for example, is not just a home thermostat controller. It learns and adapts to your routine just as Google search personalises results by looking at your browsing patterns. If you are worried now about how much information the search giant has on you, imagine how much more it could collect through the Internet of Things.

Disconnecting and limiting the amount of data being gathered will no doubt be as difficult as it is now. "Are you REALLY sure you no longer wish to enjoy an enhanced, refrigerator experience by sharing the state of its contents with your friends?" Google will ask as I delete the fridge's profile from my Google dashboard. I await the day when, as I go past Tesco on my way home, Google flashes an alert on my Android phone warning me that the cheese in my fridge has gone off accompanied by ads for special offers on Gorgonzola. Perhaps you'd like a cheeky little merlot to go with that (£3 off)? Then Google Maps tells me the roads

near my house are gridlocked, Google News informs me that several fire engines are in attendance outside my house, and my Google room thermostats report unusually high temperatures. Cue adverts for air conditioning units, ice cream machines and fire extinguishers all connected to the Internet and, of course, powered by Google.