

Sorting the Wheat from the Chaff: Using FeedWatch with Google News and Twitter

Simon Ellison-Bunce
Fellsoft Limited

May 2010

Introduction

One of the potentially most valuable applications of FeedWatch is to use it to regularly search sources such as Google News & Twitter for the latest news related to key contacts, and then to make that content easily and immediately accessible to users by creating activities in InterAction. It's an enticing prospect, but at first sight extracting high quality contact specific news from the web might seem like the proverbial hunt for a needle in a haystack. The emphasis should definitely be on finding quality rather than quantity, with the aim of bringing in only high value content, and not in so much volume that users are overwhelmed.

This paper aims to show that by making full use of the search options provided by both Google and Twitter, it is actually very simple to focus in on only the most relevant and valuable content for your key contacts.

For any web search there are two key factors: first, to specify as closely as possible **what** you're looking for (and often just as important, what you're not looking for); second, to specify **where** you want to look for it. The best way to combine these factors for any particular search will depend entirely on exactly what the search is.

We'll take a look at how to apply these principles to Google News and Twitter, with some practical examples. For all these examples, you can try out similar queries on the Google News or Twitter web sites interactively to check the quality of the results for your own key contacts. Once you have identified the best search options interactively, you can simply paste that query into FeedWatch to create a feed with the same results, using FeedWatch's integrated support for Google News and Twitter searches.

Google News

What are you looking for?

Suppose you were looking for news stories about Barclays... you might start by going to the Google home page, clicking the News link at the top of the page to get to Google News, and then simply searching for the single word "Barclays". Of course the exact results you'll get back depends on what stories are current, but you'll probably find that some of the stories relate to football, because of Barclay's sponsorship of the Premier League.

The first helpful technique you can use with Google is to exclude items that pollute the content you want by including terms in the search preceded with a "-". For example, if you run a Google News search on "barclays -football -premier" you should now find that the football related items are excluded and that all the results are about Barclay's business activities.

You can further refine the results by specifying that you only want items returned where your terms appear in the headline of the article rather than in the body. To do this, add the option "intitle:" before any word you require to be in the title - for example, "intitle:barclays -football -premier".

You can combine these options in multiple different ways to focus on the results you really want. For example "intitle:barclays -intitle:capital -football -premier " will find stories where Barclays is in the headline but Capital is not (so excluding stories about the Barclays Capital operation), and also excluding any mention of football. As another example "intitle:Barclays private equity -football" will find articles with Barclays in the headline and including mention of private equity but not football.

You can also use the option allintitle:, which requires all the search terms to be in the title, so for example "allintitle:Barclays Bank" will only return results where both the words Barclays and Bank are present in the headline.

Where do you want to look?

A key challenge when searching for news on the web is that of duplication. When a news story such as a merger or takeover announcement is released from a single original source, it will be picked up and repeated by multiple news providers, meaning that a simple search on relevant terms can often return multiple essentially identical results.

Google News goes quite a long way to resolving this challenge, because it will automatically "cluster" similarly worded stories together into a single item. One source will be picked as the primary link for the item, and others will be shown as alternate versions within the same item. When FeedWatch processes this item, it will create just a single activity for the item, but the activity notes will contain links to all the different versions found by Google so that a range of sources can be checked.

Another very powerful technique is to tell Google News that you only want to see results from a single web site, by using the "site:" option. For example, the search "site:telegraph.co.uk intitle:Barclays" will find only stories from the Daily Telegraph web site that include Barclays in the headline, and the search "site:ft.com intitle:Barclays" will only return stories from the Financial Times web site.

You can only include a single site: option in a Google News search, but it's very simple to set up several FeedWatch searches from different sources, using different site: options for each feed.

As well as the more general news sources such as newspaper web sites, you can get very useful results by searching on specific specialised vertical news sites that are relevant to your contacts. Google doesn't provide a list of sites it indexes for Google News, but some representative examples that are included currently are listed below:

- cnplus.co.uk (Construction)
- propertyweek.com (Property)
- campaignlive.co.uk (Media & Advertising)
- mandadeals.co.uk (Mergers & Acquisitions)
- accountancyage.com (Tax and Corporate Finance)
- insurancedaily.co.uk (Insurance)
- theregister.co.uk (IT & Technology)
- hemscott.com (General business and finance)

To check whether a particular site is indexed by Google News, just run a search in Google News for "site:<domain>" with no other search terms - if any results are returned, the site is part of Google News.

When using site specific searches, remember that a few sites may make the full content only available to registered users.

Once you've identified the right search conditions for a contact using the Google News site, you can have FeedWatch monitor the results of that search by simply creating a new feed subscription with the options "I know the type of feed I want to subscribe to", then "The results of a search on Google News". Click Next, then just paste in the search conditions from Google.

More information about the full range of Google search options is available via the integrated help for FeedWatch Manager.

For sites that you'd like to use with FeedWatch but which aren't part of the Google News index, you can use the free Google Alerts service. More information about Google Alerts is also available in the help for FeedWatch.

Twitter

What are you looking for?

Twitter is becoming increasingly established as the site where the latest news spreads most quickly, but the nature of the medium makes the challenge of separating the quality content from the rest significantly greater than with Google News. And on Twitter, news duplication is much more of an issue - in fact, the measure of the importance of a story on Twitter is how many times it is duplicated (or "retweeted"). This means that a simple search for a particular term on Twitter is generally unlikely to yield the desired small quantity of high value results.

Before looking at how to control where you search, it's important to understand the significance of hashtags on Twitter. A hashtag is simply a word with a hash symbol (#) at the beginning of it, and has evolved as the principle means of grouping related tweets together. For example, users who regularly tweet on the topic of intellectual property rights are likely to tag their tweets with the hashtag "#ipr". Hashtags are completely freeform, there is no definitive list of them, and Twitter users vary widely in how consistently they use them - but they can still be a useful method of filtering when used in a search. They can also help in identifying tweets that are definitely about a particular topic, rather than those in which the word is just included - for example, a tweet containing "#Barclays" is more likely to be relevant to the Barclays organisation than one that just contains the word "Barclays".

Searches on Twitter can be run directly from the home page at <http://twitter.com>, or else from the dedicated search page at <http://search.twitter.com>.

Where do you want to look?

When looking for wheat in the Twitter chaff, the principle technique to focus in on good quality content is to search only tweets from known, reputable sources. It's very simple to do this - once you have established the username for a reliable source, just include "from:username" in your search (this is the Twitter equivalent of the Google site: option). For example, the Twitter search "from:ft barclays" will only return tweets from the Financial Times (Twitter user name FT) that mention Barclays. Unlike Google News, you can include more than one from: option in the same search - so for example "from:ft from:theeconomist Barclays" will find mentions of Barclays by either the FT or The Economist (Twitter user name TheEconomist).

It's also worth noting that because Twitter is a very personal communication channel, you may find that some of your key contacts use Twitter on an individual basis. It's very easy to use FeedWatch to

track both sides of their conversations on Twitter and link that content to their InterAction contact - just create a search of "from:username OR to:username".

Once you've identified the right search conditions for a contact using the Twitter search site, you can have FeedWatch monitor the results of that search by simply creating a new feed subscription with the options "I know the type of feed I want to subscribe to", then "The results of a search on Google News". Click Next, then just paste in the search conditions from Google.

Links to further detailed information about Twitter search options is available in the online help for FeedWatch.

Qualifying results

If you are not sure about the quality of results a particular feed will return, and don't want to immediately publish every item to users, use FeedWatch to set the visibility or scope of the activities created for that feed so they're not initially visible to all users. You can then review the activities as they appear, and change the visibility of individual activities manually as appropriate to make them available to users.

Conclusion

By making full use of the search options provided by Google News and Twitter, FeedWatch can be used to pull in highly targeted news items relevant to your key contacts. Creating a carefully targeted search may take a little more time than a simple search, but with FeedWatch you only need to create each search once. Once created, the feed from the full targeted search will be monitored completely automatically, and as frequently as you require.

Fellsoft Limited was founded by Simon Ellison-Bunce in October 2009 and is focused on providing add-ons and integration products for LexisNexis® InterAction®, the leading CRM software solution for relationship-based organisations and professional services firms.

Simon started his career in legal IT and CRM when he joined Tikit in 1997. He led the early UK implementations of InterAction, and became the first InterAction Certified Professional outside the USA. Later, he was responsible for the design and development of E-Merge and ReAction Server, Tikit's very successful deeply integrated eMarketing solution for InterAction which has over 200 customers worldwide.

Web: www.fellsoft.com

Email: simon@fellsoft.com

Phone: +44 7966 241887