

The logo for Knowledgefaber, with 'Knowledge' in dark blue and 'faber' in a light green color. The background features several overlapping circles in shades of blue and light blue, and thin blue lines radiating from the top left corner.

**Knowledgefaber**

## **Real-time Search “Going Mainstream”**

A Knowledgefaber Study

A Knowledgefaber article on inclusion of real-time  
“twitter/facebook/myspace/Other” updates in search results

Written by Amit Goel, Jai Arvind and Ranjan Rajput

**8/19/2009**

Real-time search is not only a “buzz word” but has become a promising opportunity with the exploding popularity of services like Twitter, which lets users blast short messages, and Facebook, which lets people send out updates to friends.

To fine-tune their search results more, search giants started getting into deals with Twitter and Facebook (Yahoo went ahead and partnered with OneRiot, another RT search engine) whereby minute-by-minute updates will be used. This real-time inclusion is aimed to be developed into the coming days in a big “Google way” (*Google already forayed into the market while we were writing the article 😊 -- all this for really “real-time”*). How does the business side of this new development work out? Discussed below are some of the ideas that can be used to utilize real time search results.

Let us consider a Twitter account where constant updates are posted which can be accessed through search results now. Once the Real-time search kicks in suddenly we will take the “online world” to “offline world” as well in a sense. Even non-users of social networking sites would benefit from these results. Awareness and significance about social networking will increase and the entire idea will become extremely mainstream. An account in Twitter will be able to market itself (and its offerings) to people from all walks of life, even if they aren’t savvy enough to be on Twitter or Facebook. Taking a holistic view, we can say that the money that can be made in the search market remains in advertising and will continue to be so, and whoever delivers the best, most appropriate results, will have most users, market share and ad money. Add to that mobile advertising which is picking up (Read our article on monetization of content on advanced wireless technologies)

## Real Time Search

---

Real-time search is fast emerging as a popular way to find information online. The traditional search engines comprise static content wherein the web pages are indexed depending on some algorithm. In contrast, real time search enable internet users to access the most recent information faster. This becomes significant in case of trending events and breaking news. As such, the established players as well the new entrants in the search engine business are trying to grab a share in this potential market. The growth of real time search is due in large part to the popularity of micro-blogging service, Twitter, which lets users broadcast 140-character messages, and also other social bookmarking sites. Besides Twitter, Digg and Facebook messages are other major sources of user generated content which forms the base for real time search.

One critical issue for real-time search is the noise or irrelevant content that is generated. Stream everything without any filtering, and a search could bring up a lot of irrelevant chatter. Filter too strongly, and a search might omit important trends that have been picked up in the last hour. Thus it is important for real time search engines to strike a balance between popularity, relevance and immediacy.

## Evaluation of the business models of the Real-Time search engines

---

Real-time search engines market has seen a series of launches from start-ups recently. As Twitter already has an in-house search engine through its acquisition of Summize last year, the newer start-ups have to differentiate their products by filtering content based on relevancy and popularity. At this time, Twitter search only produces results by how recently they were published.

### OneRiot

---

OneRiot focuses on the actual content users are sharing via Twitter and Digg instead of their tweets and commentary. The company says it factors in 26 different criteria in its search engine, including a link's freshness, its domain authority or how reputable the Web site is, and the velocity or the speed at which a link is shared through the community.

OneRiot's advantage is the level of spam guards the company has built into the engine and its scalability. Another useful feature is search within a specific domain name so you can figure out what's shared most heavily from a certain Web site.

- Sources: Twitter, Digg, Delicious.
- Two rounds of funding totalling \$20 million from firms including Spark Capital, Appian Ventures and Commonwealth Capital Venture

### Scooper

---

Scooper is a real-time search engine that aggregates and organizes content being shared on the internet as it happens, like eye-witness reports of breaking news, photos and videos from big events, and links to the hottest memes of the day. Thus it gives the most relevant results, updated in real-time. Scooper gives a result page divided into two columns, one with unfiltered live content and the other with results sorted by popularity. It gives a nice balance of what's happening right at the moment and what content has surfaced to the top via retweets and shares in the last few hours.

- Sources: Twitter, Flickr, Digg, Delicious and other social media.

- Launched in to the public in May 2009
- Y-Combinator funded company: Series A - \$0.75M

## Twitter

---

Twitter's real-time search is probably one of Twitter's most valuable features, as it allows keeping track of an event as it unfolds in real-time. Twitter just announced that it is slowly releasing a new interface to a subset of its users that will put Twitter Search and Trends right on users' profiles. Until now, Twitter's real-time search function, which was acquired from Summize last year, lived on a separate sub domain was not fully integrated into Twitter. Clearly, Twitter has realized that real-time search is one of its core features if it wants to monetize its service successfully.

Twitter's search only shows the most recent tweets without any filtering for relevancy or popularity. It also has trending topics, or the 10 most widely tweeted terms on its right-hand column. Twitter users took advantage of this by creating hashtags - terms or phrases like Iran Election with a # sign in front of them for easy searching.

- Sources: Twitter.
- \$55 million in three rounds of funding with investors including Institutional Venture Partners and Benchmark Capital.

## Social Mention

---

Social Mention aggregates user generated content from across the universe into a single stream of information and allows to track and measure what people are saying about any topic across the web's social media landscape in real-time. Social Mention monitors more than hundred social media properties like Twitter, Facebook, FriendFeed, YouTube, Digg, Google etc.

- Currently provides a point-in-time social media search and analysis service

## Collecta

---

Collecta monitors the update streams of news sites, popular blogs and social media; so that the results can be shown as they happen. It pulls its results from all over the Web, not just from Twitter, but also from blog posts, comments, Flickr, news feeds and more. Collecta has some nice filtering options, so you can just see blog posts, or photos, or you can remove all the updates from Twitter.”

It uses the XMPP instant messaging protocol (the same technology powering Google’s communication and collaboration tool Wave), which allows it to show information that’s truly in real-time, rather than items that are simply recent.

- The company raised a little less than \$2 million last year under its old name Stanziq.
- Investors: True Ventures

## Topsy

---

Topsy scans only Twitter and gives more weight to a source’s authority and how many times a piece of content has been shared. It ranks the influence of individual Twitter accounts by measuring the fraction of their tweets that attract responses and re-tweets from followers. Results from more influential Twitterers are ranked higher.

- Sources: Twitter, although the company may incorporate other sources
- \$15 million in equity and debt to date, including \$900,000 in seed funding from angel investor

## Tweetmeme

---

Tweetmeme only searches Twitter, looking at keywords, article relevance, level of retweets and a tweet’s timestamp.

On the right-hand column, users can narrow searches to a certain time period or level of retweets. Tweetmeme is a product from U.K.-based social media company Fav.or.it. (The company also launched a very compelling real-time product called TweetTabs Tuesday, which is browser based, and thereby wins serious points for ease of use: It allows you to open several tabs with different Twitter search queries, something you might not be able to do if you are with a desktop client. And clicking on any link provides an overlay with information about the URL, including blog posts. It also provides easy ways to retweet, or reply to a tweet.)

- Sources: Twitter
- Funding: 650,000 pounds in two rounds of angel funding.

## almost.at

---

Almost.at has the most unique interface of all the real-time search companies. It has a results page divided into three parts where you can see a live stream of tweets, shared links and photos from an event. Almost.at aims to have people who are actually on

location reporting back to the larger community. It does this by having users nominate Twitter accounts that appear to be tweeting live from the event for others to follow.

One possible business model for Almost.at is to create special embeds or widgets for event organizers who want to give followers who can't physically attend a stream of content. It could also offer premium accounts for users who want to track terms outside the three or four events Almost.at follows each day.

- Sources: Twitter, YouTube, Flickr.
- Self funded

### dailyRT

---

DailyRT search is more reliant on retweets and sharing. The results are the most shared tweets on a search term in the last hour, 24 hours, seven days, or of all-time.

- Sources: Twitter

### Twazzup

---

Twazzup focuses on creating a real-time stream of tweets, but it gives a few extras like widely shared links and photos. It also suggests influential Twitterers.

- Sources: Twitter
- Twazzup is self-funded by the three co-founders

### Friendfeed

---

Friendfeed, which was co-founded by Gmail creator Paul Buchheit, launched a redesign earlier this year with real-time results. The search function ranks results by how recently they've had activity such as receiving a comment or a “like.” In advanced search, a user can rank results by the number of comments or “likes” an item has earned, letting popular or heavily shared content surface to the top. Although the results aren't sorted solely on timestamp, the search produces a lot of tweets that aren't completely helpful.

- Sources: Twitter, Digg, Google Reader, Tumblr, Flickr.
- The company raised a \$5 million Series A round from founders and Benchmark Capital.
- Employees: 9

## Crowdeye

---

It pulls its results from Twitter and provides context to Twitter trends, but it doesn't have full access to the Twitter data, and therefore is at a slight disadvantage to other sites. It shows a graph of tweet subject trends over the past three days. It includes a list of related categories, hashtags, and common words from tweets about the search term — but displayed as a tag cloud; and popular links about your search term are also highlighted.

- Source: Twitter
- Privately funded

## Status search

---

Status Search depends on logging into your Twitter and Facebook accounts (Status Search itself doesn't actually see password). Once connected, updates from friends and followers in both places can be searched at once. This doesn't provide an “everyone” view, but it's also kind of neat to focus on what's being said by those you know or are connected with.

## Twingly

---

Twingly Micro-blog Search primarily focuses on tapping into links being shared through micro-blogging service, especially in an attempt to show you what's being shared in real time.

Source: Twitter, Jaiku and Identica and Bleeper

## Twittorati

---

Just out from veteran blog search service Technorati, Twittorati leverages Technorati's ratings of top bloggers and blogs to provide top tweeted links based on authority.

## Itpints

---

It pulls back micro-blogged content from Twitter (called lifestreaming) as well as from news, social bookmark, image and other sites.

## Monittor

---

Picks three trending topics, shows live tweets scrolling down the page. There is a provision to narrow down to a particular place.

## Twitority

---

It tries to rank tweets by the authority of the twitterer.

## Yauba

---

This pulls up pages within the site itself, supposedly annihilating tracking. The real time drop down option can be used to search against Twitter and Identica.

## Sensy

---

Sensy allows navigating the real time web and seeing what is going on right now with any person, place, or subject matter. Also the hot topics and most popular topics of late are highlighted. It also shows most popular links for any respective search query. It has implemented a spam filter to prevent irrelevant results from showing up in the search.

Websites that want to place real time results on their pages can do so through the feed for websites. The feed allows any website to display real time information for any subject matter. Web publishers can add this automatically updating information to their site via the free feed. This site stays fresh as the content updates own its own, so it's an easy to use, automated solution.

## Leapfish

---

LeapFish provides multi-media and real-time search, communication, and sharing platform that gives consumers the most convenient, fun, and personalized way to experience and share the traditional and real-time Web – the new “Living Web”. The new search platform seamlessly combines the ability to search and share real-time and traditional content in any media format in a single interface across a range of different sites.

Source: Twitter, Youtube, Twitpic, Flickr, Digg

---

## Business models of Real-time Search Engines:

As the traffic from TWITTER, FACEBOOK, FLICKR and blogs increase, the need to refine the search results has resulted in not just biggies like Google working on real-time search efforts, but also a slew of startups like ONERIOT, SCOOPLER, CROWDEYE and COLLECTA. What these players aim to track is results from social networking sites and also blogs, comments on websites, feeds and even status updates.

So how does the stuff work? Well, let us consider that you are searching for the latest PC/video game from the Harry Potter franchise, ***Harry Potter and the Order of the Phoenix***. The page displays a continuous stream of results that is updated on a regular basis. A click on anyone of the preferred results leads to the original TWEET or the comment or the feed as the case may be. Of course, a filter is included that can customised with respect to individual user.

On the technology side, data is collected and indexed though an XMPP stream, the details of which we will go into. Suffice it to say the minute any new information comes in the search results are updated accordingly. ***One of the downside to this is the fact that the results are displayed in a chronological order rather than on a best-case basis.*** But the view is that such options could be added in the future as the technology matures and the users themselves refine their searches as results pour in.

Twitter is one real-time “happening tool” that can provide valuable “keyword-rich” content especially for “breaking news” from an area or specific “queries” to know about stuff such as COP15 in Copenhagen. The belief is that though indexing Twitter is a resource-consuming and mammoth task, the results are bound to be rich.

Some of the business models that TWITTER can cash-in on are described as follows:

1. *Search* – This is the primary model that not just TWITTER but any real-time search engine can work on. This search can be powerful and *real-time* in the sense that, for example, during the recent student upheavals in Iran (with rigorous media censorship in effect) Tweets were the one the many ways in getting news out. Advertisements on the lines of GOOGLE’s ADWORDS could also be implemented. Pure ads in the form of pictures or videos or posts could be tagged from TWITTER. The ads can be linked through a local site like BREAKINGSHOPPINGNEWS.COM using hashtags to tag the ad category like motels, pizzas etc.
2. *Electronic Commerce* – Let us consider a fashion line (Armani) that wants to convey its last remaining stock details to its customers. Instead of going to e-bay,

they can post details on their TWITTER account and the products sell like hot cakes! Of course, there is the option for TWITTER to charge some commissions for such direct sales.

3. *B2B Consumer Service* – This involves personnel offering advice on anything the consumer asks details about. Only the communication is online and the medium is Tweets. TWITTER is already in a deal with BEST BUY doing the exact same thing! This could be an important part of the vendor/customer relations in the future.
4. *Network Marketing* – Given the large size of the user base for TWITTER, the potential connectivity is valuable and every conversation is beneficial. Though TWITTER has opted out of advertising (for now), others may decide to license “targeting.” This involves targeting a group of people who share a common interest and then expanding this net manifold by looking for others with the same interests. While “targeted” connectivity is surely a takeaway, privacy issues come to the fore.

### Google’s Recent Foray

The recent launch of “Real-time Search” from the leader in search business talks about how important this piece has become. Google's real-time search capability, which required the company to develop a slew of new technologies, will be rolled out over the next several days (08, Dec, Source: Total Telecom). We tried searching for “Real-time search” and got the following results with a scroller on the first page itself.

Google Launches **Real-Time Search** ✖

7 Dec 2009 ... Google **real-time search** updates as stuff is happening around the Web: live tweets, Yahoo Answers, news articles and Web pages stream in on ...  
[mashable.com/2009/12/07/google-real-time-search/](http://mashable.com/2009/12/07/google-real-time-search/) - United States - [Cached](#) - [🗨](#) [↑](#) [✕](#)

Latest results for **real time search** - [Pause](#)

**Real time search** is apparently up and running:  
 Yfrog - j4zc - Uploaded by heilpern - yfrog.com  
 heilpern @ [twitter.com](https://twitter.com) - seconds ago

[Wessler: A hiring process done by a committee of one](#)  
 Peoria Journal Star - 1 minute ago  
 The notion of a Bradley family is **real**. That reality is why so many talented people in all ... or why they leave for a **time** but feel compelled to return, ...

Like Whoa! **Real time** Google **search** just picked up my latest tweets, a little weird.

Yahoo is also turning on the real-time search war well. Among the companies providing real-time results to Yahoo on a trial basis is OneRiot (initially to only 10% of users). It is believed that Yahoo is testing a new "search shortcut" that would include real-time results at the top of search results pages. The shortcut will only appear on certain queries that will be determined by Yahoo. It has been also discussed that Yahoo had also been testing the limited inclusion of Twitter messages within search results to certain users.

While Microsoft's Bing was the first one in the biggies to start this service Google is going to give it a tough time ahead. Bing won kudos from search engine experts for its ability to do real-time searches of social networking sites like Twitter, etc., but how it does against Google going forward is yet to be seen.

### **Conclusion**

With so much happening in the "Real-time" search market, it could trigger either a mass acquisition or a mass extinction of the RT start-ups such as Topsy. It is not only important to include real-time search results but the biggest challenge is to be able to rank them by reliability of the author (not all tweets are equal), freshness (most recent), and importance w.r.t. topic searched apart from host of other things. Some of the start-ups mentioned above have great ideas around this aspect of ranking. Other have some great ideas around how results can be displayed. However, a search engine giant like Google would not buy a search company as it has internal capabilities to do it (core competency) and continue to buy mobile advertising (admob) and players in other parts of the value chain. The giants have enough talent, technology and marketing firepower that they can make real-time search a mainstream idea. The onus is on who will make the best possible integration of the technologies into its business model and who will be able to rank the best-in-class tweet/updates/blog results as compared to recent results. Whoever it is will most likely emerge the undisputable leader in the overall search market.

Disclaimer: Information contained in the report has been obtained from sources deemed reliable and no representation is made as to the accuracy thereof. Neither Knowledgefaber nor its affiliates, officers, directors, employees, owners, representatives nor any of its data or content providers shall be liable for any errors or for any actions taken in reliance thereon