Making search social
Unleashing search for the mobile generation

Taptu White Paper
February 2008

"Taptu not only recognizes the importance of injecting human preferences and human judgments into computer algorithms to pinpoint truly relevant information; it has taken the lead in allowing people to communicate those results easily to their peers.

In the past, if users wanted to make themselves heard they created a webpage. What we're seeing now with Taptu is the potential for communities to form based on their search behavior and passion for results that reflect what humans want and not what algorithms dictate"

Peggy Anne Salz, Publisher and Chief Analyst, MsearchGroove

taptu.

taptu.

Table of contents:

Acknowledgements

My thanks are due to my colleagues Chris Moisan, Stef Butlin and Bob Last who gave me valuable feedback to my early drafts, and whose constructive criticism has shaped many of the arguments and premises in this White Paper. Thanks also to Celia Jones and Lynsey Tucker who proofread the document and helped polish it into a better state.

taptu.

1. Introduction

In early 1998 Larry Page decided to publish his first research paper, on the topic of the PageRank algorithm. It was not an auspicious start. The paper was rejected in its initial form and published later as part of a Stanford digital libraries project[1]. Shortly afterwards, a company called Google was set up to commercialize this technology.

The subsequent rise of Google has been well documented by John Battelle in his definitive book "The Search". By 2002 Google was generating revenues of over $400m despite the collapse of the dotcom bubble. In 2007 its global market share of search worldwide reached 65% equating to 300 bn searches per annum, or nearly a billion searches a day. Ten years after its birth Google became the most valuable company on the NASDAQ.

Surprisingly, search was not seen as an exciting business opportunity back in 1998. Search was frustrating to users, who often had to click down through many pages of results to find what they were looking for. AltaVista, the early market leader, was focussing more on turning their site into a Web portal than improving their algorithms. The key missing ingredient was relevancy of search results. Drastic improvement in relevancy, delivered by PageRank, led to search becoming a mass-market service that sat at the center of users' Internet home pages.

Google were among the first to see the potential of search applied to the embryonic but growing mobile Internet. In 2001 Google launched a new search service on mobile phones, with "users throughout the world gaining untethered access to the 1.6 billion web documents in our growing index"[2].

Mobile search grew slowly over the intervening years, but by the beginning of 2008 had only reached 2% of the volume of desktop search – 10bn searches/annum compared to 490 bn on the desktop. Yet the number of mobile Internet users has recently begun to approach the number of PC Internet users – 800m vs 1.2bn.[3]

Today's higher-end mobile phones have performance capabilities (processing power, memory, network bandwidth, browser features) on a par with 2000-era PCs. Why is mobile search today not nearly so relevant to the needs of phone users as it is to PC users? Is it just about slow networks and small screens or are there other factors at play?

In Section 2, we will propose that mobile search cannot make the journey to mass market on its own. Certain conditions must exist first in the wireless ecosystem. Even when these conditions exist, search must evolve to be relevant to the needs of today's consumers who, as we observe in Section 3, are very different in behaviors and beliefs to the PC users of 1998. In the remaining pages, we will propose a new roadmap for mobile search, which recognizes these new realities.

taptu.

2. A new wireless ecosystem

In 2003 my previous company embarked on a technical partnership with Qualcomm, who subsequently acquired us a year later. I had the good fortune to share many hours in discussion about the future of the mobile industry and our two companies with a senior vice president at Qualcomm, Johan Lodenius. Johan had travelled around the world for a decade evangelising the benefits of a new 3G wireless technology called CDMA, which was invented by Qualcomm and adopted, by a first wave of mobile operators in Japan, the USA, South Korea and Israel.

Johan and I agreed that several significant and unstoppable changes in the wireless world were already underway. They would culminate in a tectonic shift in the otherwise conservative and slowly evolving world of mobile network operators. Johan called this "the new wireless ecosystem" and had identified its important characteristics. They resonated with me because of my own experiences in negotiating deals with mobile operators and manufacturers in the existing ecosystem.

From our vantage point here in 2008, the impact of these megatrends is now clearer, now that they have had several years to take effect. These trends have been underway since 2000, and will continue to around 2010, when they will have reached mass-market penetration. The impact on consumer behavior will lag the technology by a year or two, but will be profound. A horde of start-ups are creating the new mobile services that will ride on the back of these trends, as mobile blogger Rudy De Waele has documented[4].

There are four key megatrends that are combining to create the New Wireless Ecosystem:

- Ubiquitous mass storage + processing power on handsets
- From mobile narrowband to mobile broadband communication
- A revolution in mobile UI
- From walled garden operator portals to open gardens

a) Ubiquitous mass storage + processing power on handsets

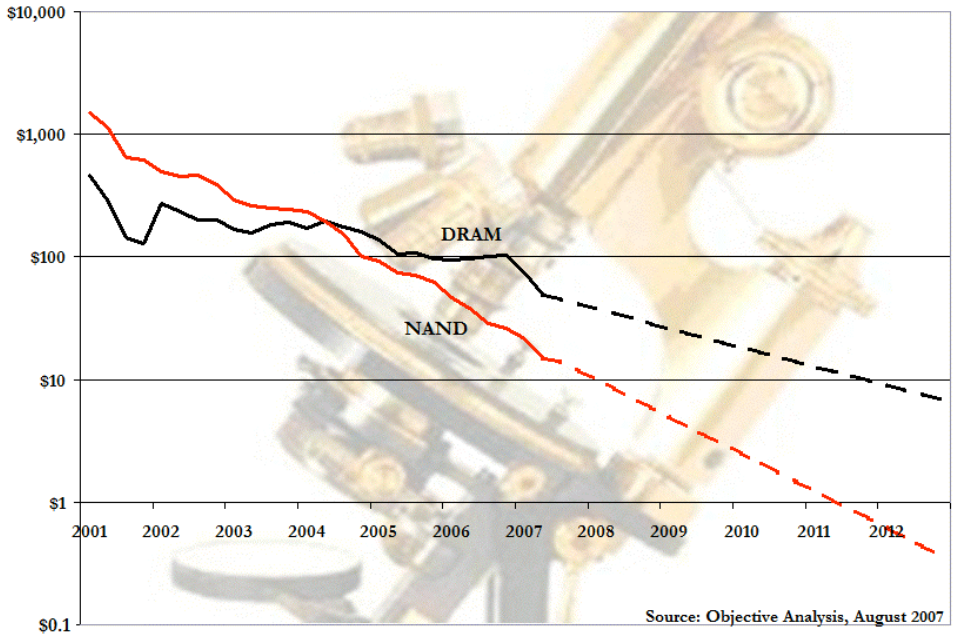

Fig 1: the declining price of flash memory. Source: Objective Analysis[5]

The availability of mass storage plus processing power on handsets at mass market prices is creating an enormous new market for mobile phones as entertainment devices. This is being driven by the continuing advances in semiconductor technology.

Let's talk MHz and Gigabytes. In the year 2000 I founded a company to develop a new user interface technology for mobile phones. By 2002, the initial system worked well on a PC, but struggled even on a high end mobile phone - a Nokia 7650 with an ARM-9 series CPU running at 104 MHz CPU clock, with 4 MB of RAM and 16 MB ROM memory. Low-end mobile phones ran with 20 MHz CPUs and 0.5 MB total RAM and ROM. It took considerable engineering resources to craft our code to run efficiently enough to make it ready for mass-market phones.

By 2008 the Nokia N95 phone in my pocket now runs a 332 MHz OMAP ARM 11 CPU, with 64 MB RAM, 160 MB FLASH, and a 4 GB SD memory card. In six years, processing power has increased by (just) 3X, but memory capacity has increased by 10X. Flash memory is being transformed from an exotic, niche product at $2,000 per gigabyte in 2001 to an everyday commodity at $2 per GB in 2010.

Bottom line: even the low-end handset in 2010 in developed markets will carry most consumers' entire music and photo collections, plus a nice selection of video. The phone is no longer just a voice device. This is the first pre-requisite for the New Wireless Ecosystem.

b) Mobile narrowband to mobile broadband communications

Mobile phones have been able to receive and transmit data from the early days, but only at rather low data rates and high prices. In the GSM operator world, GPRS packet data was first introduced in the year 2000, with a typical speed of 10 Kbps. Since then, progress in increasing data rates has been impressive. With the latest version of 3G, called HSPA, widely deployed by 2010, we will see mobile phones able to communicate at 10 Mbps in real world conditions. This represents a 1000 X improvement in performance over a 10-year period. In 2010 we will also see trial implementations of 4G technologies such as LTE, running at up to 100 Mbps. Even more important will be the dramatic reduction in latency for data traffic that 4G offers. These performance improvements come with a drastic reduction in price per megabyte transmitted – from $20 per MB to $2 per MB today. With new flat rate mobile data tariffs, prices look to be headed even lower.

Malcolm Gladwell in his book The Tipping Point[6] explained how certain trends achieve exponential popularity while others sputter and fade into oblivion. For consumer devices a key point is reached in the adoption curve when 10 to 20% of the population within a country uses a particular device or technology. At this point the adoption curve becomes self-sustaining. Enough people are using the technology and talking about it to their friends and family that positive network effects can kick in to increase adoption still further.

When consumers use 3G phones, their usage of the mobile Internet increases significantly. According to M:Metrics, a 3G phone user is twice as likely to use email, nearly four times as likely to access information via the browser and to use mobile search, and six times as likely to download videos or play music on the handset (see Figure 2 below).

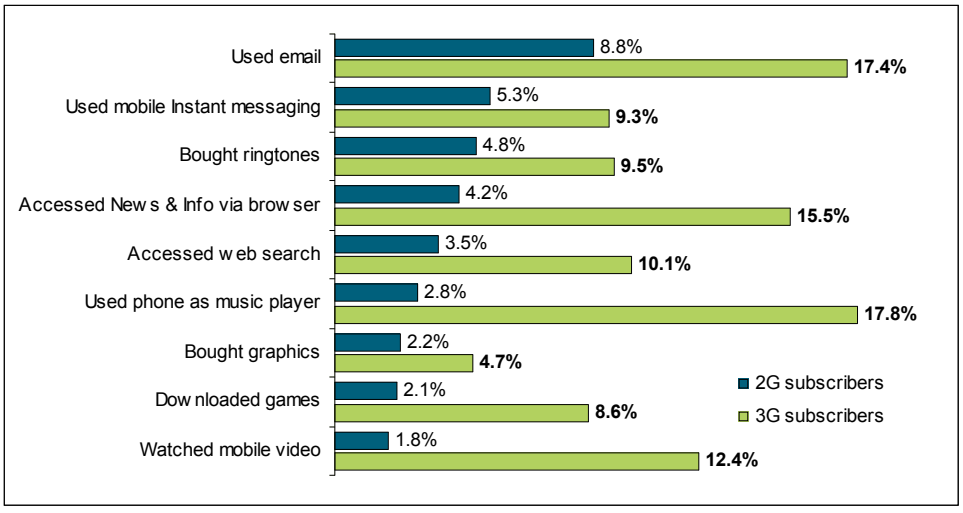

Fig 2: 3G is important: penetration of selected content types for 3G users vs 2G users in Italy, Q3 2006. Source: M:Metrics[7]

Region	Wave 1 2005	Wave 2 2006	Wave 3 2007	Wave 4 2008	Wave 5 2009	Wave 6 2010
North America	Canada, USA					
Asia	Japan, N.Zealand, S. Korea	Australia, Singapore	Hong Kong, Taiwan			
Middle East	Israel		Qatar, UAE	Bahrain, Saudi Arabia	Kuwait	Syria
W. Eur		Italy, Ireland	Austria, Denmark, Finland, Germany, Greece, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK	Belgium, France, Iceland		
E. Eur			Estonia	Czech Republic, Hungary, Latvia, Lithuania, Poland, Slovenia	Bulgaria, Croatia, Romania, Russia, Slovakia	Belarus, Serbia
South America	Venezuela					Argentina, Chile
Africa						S. Africa

Fig 3: Countries exceeding 20% 3G penetration by population
Source: Cowen Wireless Equipment Survey 2007[8]

Let's take a look at the countries in the world where 3G handset ownership now exceeds 20% of the population. The first wave was centered on Asia, starting in South Korea and Japan, which were the first countries to implement 3G on a wide scale. By the end of 2007, 24 countries had reached 20% 3G ownership. By 2010 Eastern Europe, Russia and chunks of South America will have followed suit.

The availability of Mobile broadband communication at mass market prices is the second pre-requisite for the New Wireless Ecosystem.

c) A revolution in mobile UI

As the performance of mobile technologies has accelerated, new phone functions have proliferated. This trend started with address books and alarm clocks, and has moved on to encompass calendars, games, cameras, music and video players, and most recently GPS location and navigation.

Silicon Valley author Alan Cooper coined the term “cognitive friction” which he defined as "the resistance encountered by human intellect when it engages with a complex system of rules that change as the problem permutes."[9] As mobile phones morph into cameras, music players & navigation systems, cognitive friction escalates. It's easy to get frustrated with a mobile phone. The promise that it will put you in total control of your life and make it run like clockwork is usually a shallow one.

As mobile phones add more and more data-centric features this has led to a parallel increase in the complexity of the phone UI. From the mid-1990s Nokia were the pace-setters in mobile usability. However, as Russell Beattie observes of their recent smartphones: “the GUI just hasn't scaled with all the new functionality and options available and as a result is a complete mess .... this has been a problem for a while, but with the amount of functionality now available, it's just gotten completely out of control”[10].

Fig 4: The iPhone from Apple, Inc.

The need has been building steadily for a massive simplification in the user interface of mobile devices. But it was not until 2007 that there was a revolutionary step forward, and it came from Apple with the iPhone. The iPhone incorporates a novel kind of touch user interface optimized for small mobile screens. The entire front face of the mobile device is a touch screen, which can recognise different gesture patterns performed with one or two fingers. Killing two birds with one stone, the iPhone user interface increases screen area, and increases touch interaction area, while keeping the device small and thin. Its not just a mobile phone, it's the "casual computer", as user experience designer Christian Lindholm calls this new breed of mobile device[11].

This new style of user interface will trickle down rapidly to mass market phones, from Apple and from other manufacturers. In the words of Dan Appelquist, chair of the W3C Mobile Web Best Practices working group: "the iPhone will herald a whole generation of advanced Web-capable mobile devices… it's the Web, but not as we know it"[12].

Just as new icon and windows-based interfaces made desktop PCs usable for the mass market, this big step forward in mobile UI is the third pre-requisite of the New Mobile Ecosystem.

d) From walled garden operator portals to Open Gardens

During 1998 to 2002 there was a false dawn of optimism for the mobile Internet, promoted by mobile operators and manufacturers around a technology called WAP. WAP was a proprietary version of existing Internet technologies optimized to work on the constrained networks and devices existing in the mobile industry at that time. Mobile operators followed the business model of the US Internet company AOL from a decade earlier and created "walled gardens" of selected mobile content that could be accessed from WAP browsers in a quasi-internet model.

As the mobile analyst and blogger Dean Bubley points out very eloquently: "Mobile operator portals are not the Mobile Internet or any of that nonsense. Internet is short for internetwork, which means the connection of two or more *separate* networks. Accessing an operator's portal from that same operator's customer's phone is no more an internetwork than a set-top box talking to a cable operator's head end, or my switching on the electricity at the wall for my toaster. It's one network, not several. There's no *inter*. "[13]

Many mobile operators have restricted access to the Internet for subscribers, even operators that have invested heavily in 3G network infrastructure. On the real Internet there has been an explosion in the choice of content – increasing to billions of items in the last decade. Google's index reports approximately 20 billion web pages.

On the quasi-internet offered by many operators, there may be only 200,000 content items. John Strand has highlighted the difficulty of trying to operate these quasi-internets: "When we look around the world we don't see one single operator that has been able to create a successful position as a mobile media company"[14]

After the failure of WAP, some operators started to experiment with a new model for the mobile internet – the open mobile access model – where mobile subscribers are not restricted from accessing the full internet. Ajit Jaokar has used the term "Open Gardens" to describe this model[15]. Slowly, this open Internet access model is gaining traction, even in markets like North America, which have historically been very operator-controlled.

Some mobile operators are opening their portals but at the same time introducing transcoding services provided by suppliers like Novarra and OpenWave. These are supposed to pre-process desktop Web pages before they reach the mobile screen to make them easier to display. Sometimes they interfere with the ability of mobile Web developers to optimize their services for popular devices. Where this happens, it represents not so much an Open Garden as a "chaperoned garden" which impedes progress to the real open mobile Web.

Powerful handsets, fast 3G and great user interfaces are not enough to ensure a transition to a real mobile Internet. Open Gardens are the 4th pre-requisite for the New Wireless Ecosystem.

3. From X to Y: the demographic factors

The Internet as we know it today was built and initially used by Generation X types, born roughly from 1965 to 1976. For Generation X the 'wow' factor of the Internet became the ability to find anything you want. This skill was mastered by just a few in the early days of the internet, but was unlocked and transferred to a much wider population within Gen X by the arrival of Google's powerful search box.

Assuming the four megatrends described in Section 2 come to pass, the mobile Internet will be first be used en masse by a new generation – Generation Y. Generation X may be playing a key role in helping design and build it, but mass-market adoption is coming from people born between 1977 and 2000 – Generation Y.

Who or what is Generation Y? Other names that have been used to describe this generation include the Echo Boomers, the Mobile Generation, the iPod generation and The Millennials. There is no consensus over the exact birth dates that define Gen Y. But the broadest definition generally includes the more than 70 million Americans born 1977 to 2002, together with their counterparts around the developed world.

Some observations about Generation Y:

- This is a generation that is totally at home with technology. The mobile Internet will hold no fear for them.

- Young people become celebrities in their own worlds by posting videos on YouTube, posing like a model on MySpace or creating an online reality show featuring themselves. The Pew Internet & American Life Project found 54% of those 18 to 25 have used social networking sites such as MySpace or Facebook; 44% have created a profile featuring photos, hobbies or interests[16].

- In contrast to ultra-individualist X-ers, Millennials are group-oriented -- meaning that they are less interested in an "army of one" and more interested in the "watch me become we" alternative. Group-oriented concepts such as "leave no one behind" may emerge from the movies (2002 movies *Lilo and Stich* and *Black Hawk Down* both used this phrase) and go mainstream[17].

- Millennials appear to be using rapid-fire communication via the Internet and other peer-to-peer media to build a newly inclusive "one" from their wildly diverse origins[18].

- They are far more technically competent than their school and uni teachers. Sometimes they come to school bleary-eyed after sitting up half the night on chat[19].

- Generation Y is the first generation to grow up not knowing what it was like before the Internet and before the mobile phone. They have grown up with the Internet as an assumed normal part of everyday life.

Their ability to discover weird and wacky sites is not new to them – it's just the normal Internet. This whole array of new media – blogs, chat rooms, IM, social networks – represent tools of self-expression, of customization and of personalization for Generation Y kids. "A primary goal of people my age is not necessarily to become famous but to become distinctive," says Jason Barg, 24, a 2004 graduate of Penn State University[20]. To them it's not new media, its just media.

For Generation Y, the ability to find anything on the Internet is taken for granted. Some of them will never be aware of a time before Google. The new "wow" factor is how the Internet connects them to their social networks and how it allows them to express themselves. The Internet and the PC were the tools that got them started along this particular road

In the next section, we will argue that it is the mobile phone that is the device that will take them to the next destination on this particular journey. Mobiles are truly social devices, used first for voice calls and then more recently for texting, for personalization with ringtones and wallpapers, and for file sharing through Bluetooth. The nature of search in this new world of mobile Internet devices will shift. This is because the journey that Generation Y is taking on the Internet is more concerned with social expression than finding information.

taptu.

4. Mobile search made social

Let's look more closely at the social context of search. Search on the desktop tends to be a solitary activity. Whether you are checking the weather, looking for an address, or catching up on the news you are in a solitary mode. When you are using Facebook, MySpace or Digg you flip into a social mode. People will sometimes group round a PC screen and search for something together but it is not the usual mode.

In the previous section we proposed that while the PC has evolved to become a social device, the mobile phone is a supersocial device. It has always been used for social purposes, which have only multiplied as the devices have evolved to support texting, personalization, P2P file sharing and Internet access.

Not surprisingly, as search engine vendors have migrated their services to the mobile platform, they have focussed on familiar user scenarios, usually solitary, extended to the mobile context. They have worked to optimize their existing search back ends to the following kinds of activity from your mobile:

- Finding pizza
- Checking the weather
- Catching up on the news
- Finding ringtones to download
- Checking prices online while shopping
- Finding mobile games
- Finding out sports scores

However, today's mobile search engines perform rather poorly at tasks that may have a more social context:

- Finding a friend's phone number
- Finding photos on Facebook
- Sharing a search result with a friend
- Finding new music your friends like

This may explain why Generation Y, who sees the mobile as a social device first and an information device second, is not using today's mobile search as much as expected. But Generation Y is using mobile phones to access social networks. M:Metrics measured Facebook as the UK's most popular mobile social network by October 2007, with nearly 680,000 monthly users. MySpace was second with just over 417,000 users, and Bebo was third with just over 310,000.

Mobile phones will become the optimum way of accessing social networks because they are always with you, and they are less subject to workplace usage rules than PCs.

If search is to move forward in the new medium of the mobile internet then it will need to perform well in the social dimension and not just in the solitary dimension.

Users will want to search for content and information that has social significance or lends social currency. When they find it, they will want to action it. When they find a phone number they will want to dial that number, as they do when searching the phone's local address book today. When they find a song, they will want to download a preview and play it, and share that result with friends. On mobile phones, search needs to be made social and results need to be made actionable.

There are four key aspects of "mobile search made social" which we will review in the next four sections of this White Paper:

1. Social-assisted relevancy scoring
2. Easy sharing of search results
3. Human editing of search results
4. Crawling and indexing social info

5. Social-assisted scoring

Fig 5: Logos of selected Web 2.0 companies

The phrase "Web 2.0" became popular following the first O'Reilly Media Web 2.0 conference in 2004[21]. Although the term suggests a new World Wide Web, it doesn't refer to any update to the technical standards of the Web, but rather to the way people use the Web. Web 2.0 sites have an "architecture of participation". They gain effectiveness through inter-personal connections and network effects, so that they grow in effectiveness in proportion to the people that make use of them.

Relevancy is everything to a search engine. There is no point in returning 1.47m results in 0.135 seconds if the most relevant results are buried 50 pages down the results list. Especially on a mobile phone, the top performing search engine is the engine that can return the results that are most relevant to the user within the top few search results.

Google were able to harness their powerful innovation in relevancy – PageRank – to overtake an established incumbent, Altavista, in Internet search. PageRank measures the significance of a page by measuring the degree to which a web page is linked and referred to by other web pages, which themselves will have varying degrees of significance. This works less well on mobile devices compared to PCs, for two main reasons:

- It takes no account of the mobile-friendliness of a web page, so song searches tend to return song results that cannot be played or previewed on a mobile phone.

- Web pages on the emerging Mobile have very few external links to other sites, making link strength an unreliable indicator.

Web 2.0 brings new opportunities for improving relevancy. Web 2.0 sites tell us which content is most popular by measuring participation of the community. MySpace, for example, is the world's largest social web site for music artists. The number of artist profile views on MySpace is a good indicator of artist popularity. The number of song plays shown in the MySpace player on the artist's profile page (see below – Lily Allen example) is an excellent measure of song popularity.

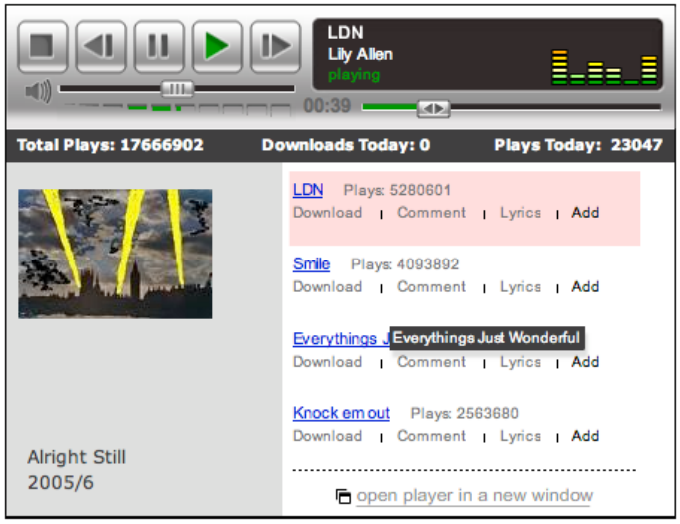

Fig 6: Lily Allen song "LDN" displayed with number of plays on MySpace.com

A search engine that could crawl music sites like MySpace and extract key metadata about artists and songs, including popularity data, which could then rank search results for artists and songs by popularity, and provide access to mobile-friendly song summaries, would perform much better at music search on a mobile than a traditional mobile search engine. Instead of calculating relevancy from link strength measured from Web pages, it calculates relevancy from social scoring. MySpace contains over 10m songs as of January 2008, so the performance of such a search engine across the Long Tail of music content would be expected to be very good.

Social-assisted scoring is a strategy that can be extended across many content categories. Humans are good at generating content, classifying it and aggregating it into comprehensive social sites like MySpace and, Wikipedia. Machines are good at crawling and indexing and ranking the content contained and listed by these sources – that's what a search engine does. A search engine based on social-assisted scoring can combine these processes to give you the holy grail of very pure and relevant search results and very good Long Tail coverage.

The price you would pay compared to algorithmic search would be a reduction in the variety of results. You might end up with an index of 1 billion pages rather than the 20 billion plus index of the full Web. On the desktop this price may not be worth paying. But on mobile, where the user gets punished much more heavily when they inspect candidate results of low relevancy (by having to perform many more clicks, taps and scrolls because of the small screen size, and sitting waiting for the wireless network to respond) this price IS worth paying.

This strategy can be extended beyond music to any category of social Web content where popularity of content is measured on a Web site. Video content on social sites like YouTube (see below) can be scored by the number of video plays, and by user-generating ratings for each video clip.

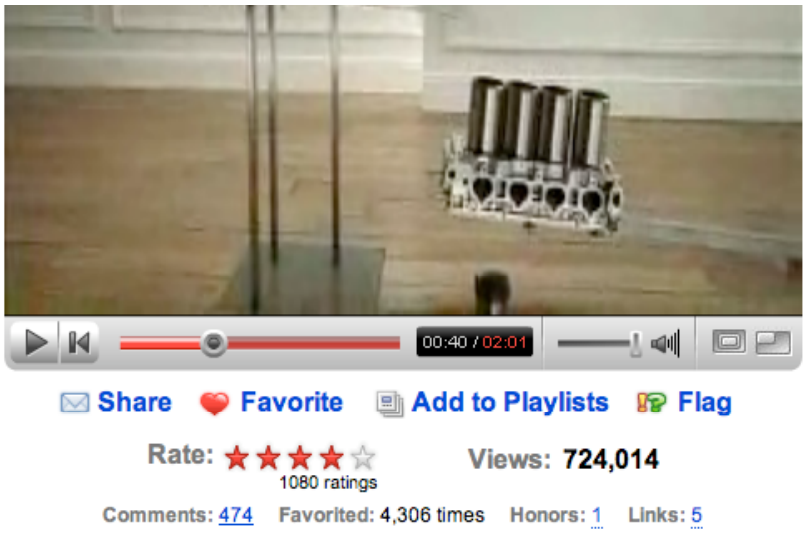

Fig 7: Honda viral video as rated and viewed by users on Youtube.com

Topical Web content can be scored by the number of times a content item is commented on and shared within a community. A good example of such a community site is Digg (see below). Social-assisted scoring applied to this category of content would measure the number of social bookmarks created ("Diggs"), the number of comments, and the influence of the users who are sharing the bookmarks.

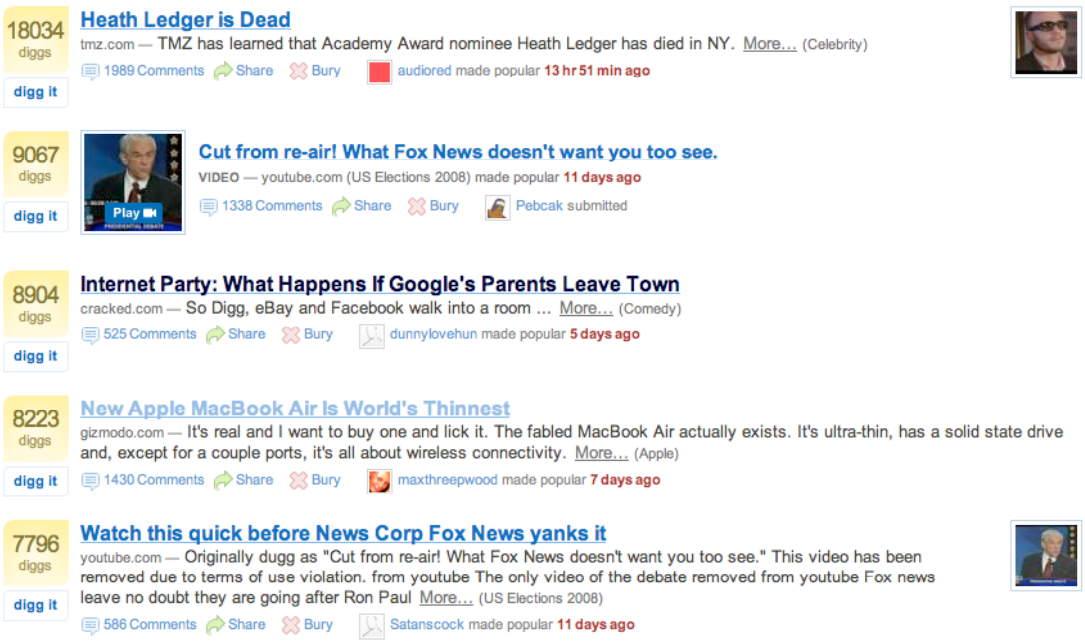

Fig 8: Most popular items in last 30 days on Digg.com, January 23rd 2008

With social-assisted scoring, relevancy can now operate at 2 different levels:

Level 1: Relevancy in general – what the world thinks is important
Level 2: Relevancy in your friends group – what your friends think is important

A mobile search engine that operates at Level 2 requires a search engine that can access your social graph and boost the ranking of search results that your friends have deemed to be important. It would preferentially show results that were more fun and more relevant to you in your social context. For Generation Y, this would be a significant benefit.

To summarise, social-assisted scoring brings the following benefits to mobile search:

- Higher accuracy of scoring for specific categories of content
- More accurate relevancy signals in the absence of page linking
- Works well across the entire Long Tail of content within a category
- More personalized - it can measure relevancy within your social context and social group

taptu.

6. Sharing mobile search results

Sharing mobile search results today is very difficult for users. Mobile search has a social context (as discussed in section 4), so if sharing results is difficult then searching is likely to be less frequent. Conversely, if sharing results becomes easy, then mobile searches will be more frequent.

Sharing search results is very easy to do on the PC. We have two hands available, and a mouse or touchpad. We have very capable Web browsers which we can open in multiple tabs on large screens. It's easy to copy a URL into an email or drag and drop it into a Facebook Wall to share them with friends.

On the mobile phone Web browsers are more basic, usually with just one window or tab, and copy and paste is more difficult. Only a small proportion of handsets have email set up. Sending URLs in text messages to multiple recipients is not easy. It takes a lot of clicks and scrolls.

How then can we make it easier to share search results on the mobile phone? In mobile usability studies, one popular technique is to measure and then reduce click-distance. Click-distance is the sum of the number of clicks and scrolls that must be performed by the user of a mobile service to complete a given task. The frequency of use of a mobile service falls off rapidly as click-distance increases, by as much as 50% per click.

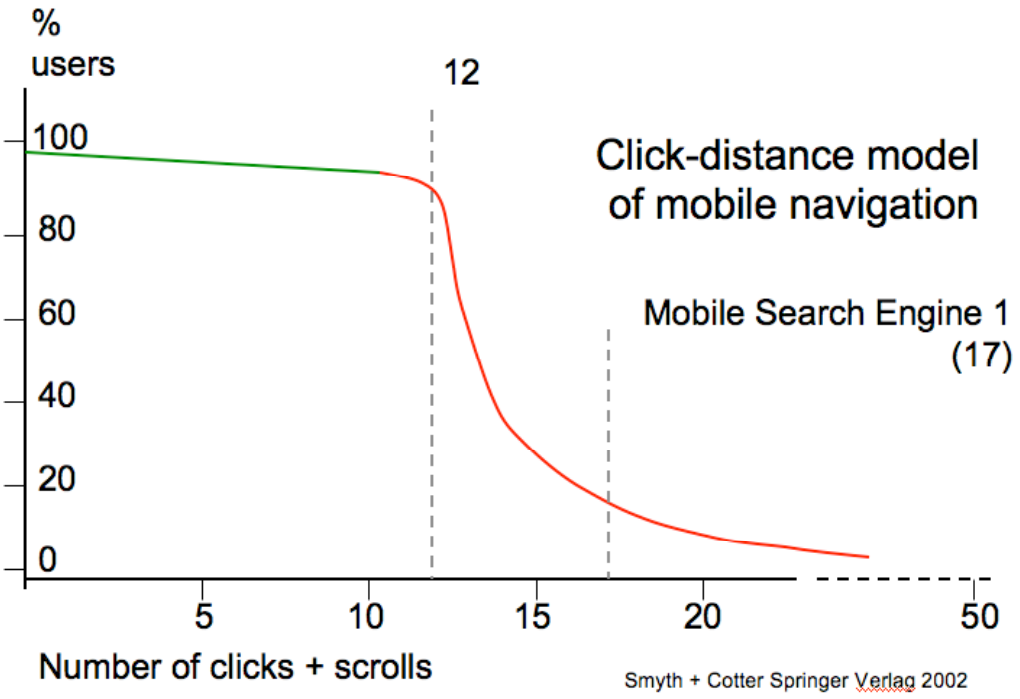

Fig 9: The Smyth & Cotter experiment, The Plight of the Navigator, 2002

In a landmark study published in 2002, Barry Smyth and Paul Cotter at University College in Dublin analysed the behavior of mobile users who were looking for mobile content on the O2 mobile portal[22]. They observed a critical click-distance threshold – 12 clicks - where users became much less willing to continue navigating to content. After this threshold, each additional click or scroll resulted in a 50% reduction in the number of users willing to continue to navigate the user interface.

For mobile results sharing, the challenge is to take a task that has a very high click distance today, and design a new user interface and screen flow that drastically reduces this. This implies that the search box and the friends' links have to sit together very closely in the user interface and its navigation flows. Mobile users will want to share search results with both desktop and mobile users. Mobile-to-mobile sharing has to work for both one-to-one sharing and one-to-many. One-to-many sharing has to work without incurring large costs from sending high volumes of text messages, which are almost always priced per user per message.

What will be needed in mobile is not just a great search engine, but a great search and share engine. On mobile, search doesn't finish when you find something, it finishes when you find and then share it with friends. They may then send you a comment back, or ask you a new question which stimulates more searches. In Section 9 we will describe a novel search result sharing system for mobile devices, called 1-Tap sharing.

Let's now move on to discuss the third key component of mobile search made social: human-assisted results.

taptu.

7. Human-assisted results

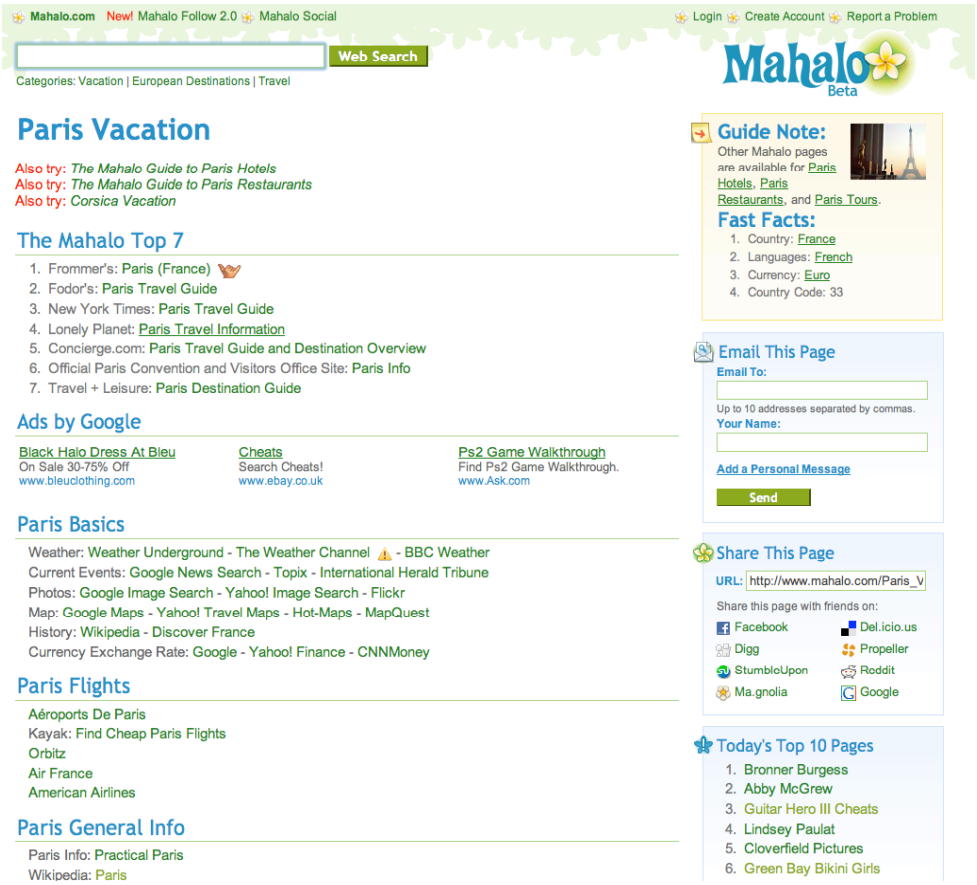

Fig 10: Mahalo display of search results for "Paris vacation", Jan 23rd 2008

Human editing has the potential to significantly improve the relevancy and usefulness of search results. Compare the example search for "Paris vacations" shown above. Apartment rental sites and package holiday sites, both in the organic results and ad sections, dominate Google's results. Mahalo's human researchers have compiled an "ideal" Paris vacation results page which most users would find much more useful in planning a vacation. They have coined the term "human-powered search" to describe their approach.

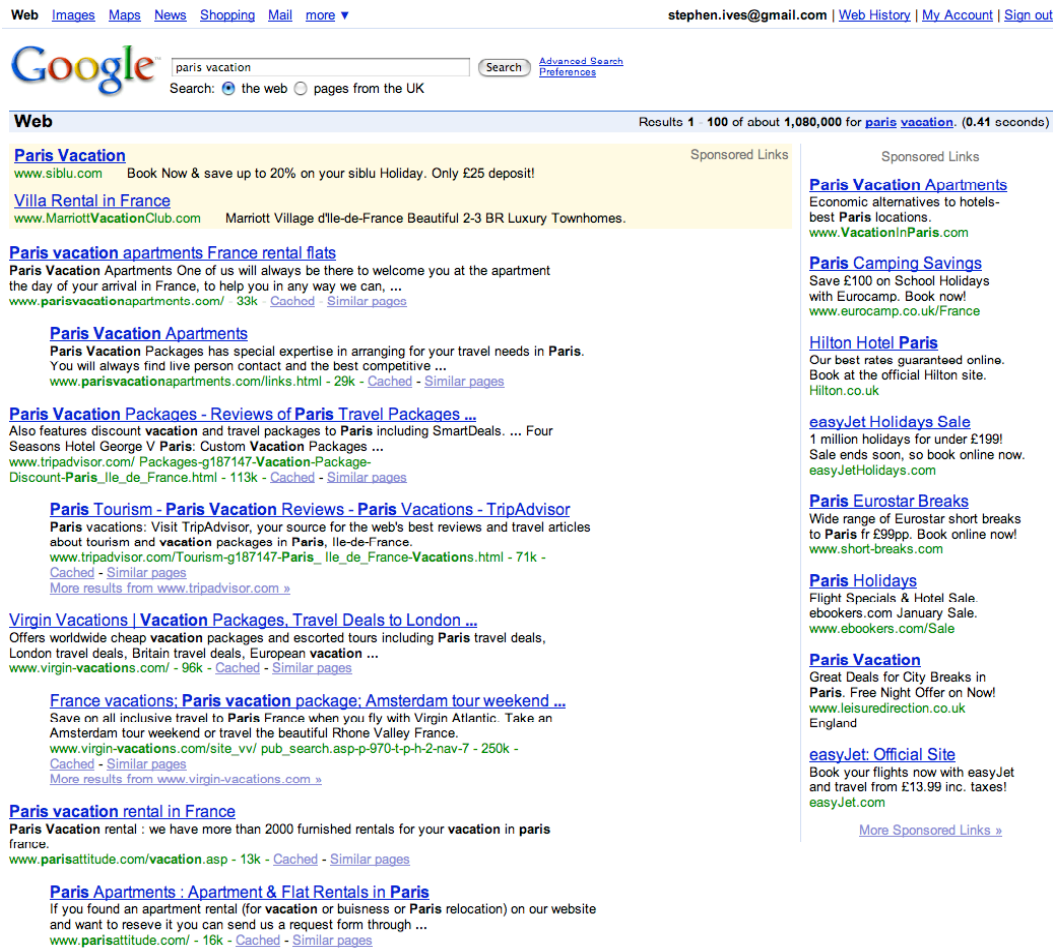

Fig 11: Google display of search results for "Paris vacation", Jan 23rd 2008

The main benefit of this approach is that human judgement can be used to optimize results and eliminate so-called "SEO spam" in search results. SEO spam refers to search results that are highly ranked only because of the manipulation of search engine behavior by SEOs. Human-assisted search works very well for popular searches at the top of the Long Tail, where competition for placement at the top of the results list is high and the motivation to "spam" the Google algorithms is also high.

The main downside is that manual result editing is time-consuming, and can only be applied to most frequent searches at the top of the Long Tail. Mahalo only optimize results for the top 50,000 search terms in a single country: the USA. For all other search terms they fall back to Google results. Mahalo are planning to ease this constraint by creating the Mahalo Greenhouse – an army of part-time researchers, a paid-for kind of social network linked in to the Mahalo site.

Human editing of search results is a controversial topic. Chris Anderson, author of "The Long Tail", posted in his blog on a recent DLD panel discussion[23]:

> Jeff Jarvis is liveblogging a DLD panel discussion between Jimmy Wales (Wikia) and Jason Calacanis (Mahalo), with Marissa Meyer (Google) in the front row. They're discussing human-driven search (Mahalo and Wikia) vs. algorithmic search (Google). And then comes this fascinating exchange: "Esther Dyson says the problem with their model is the long tail. If she were starting a company she'd do something else. "Ok, I'm going to kill myself now," laughed [Calacanis] He says that the **fat tail will be human, the medium tail social, the long tail algorithmic.** And he says that the advertising interest is in the fat tail."
>
> As Kevin Marks and Julian Bond rightly point out in the comments here, Jason slightly mangled the nomenclature, so that should actually be **"The short head will be human, the fat middle social and the long tail algorithmic".** Still, that single sentence is worth another book. I won't write it, but I'll bet someone else does.

With mobile search, it is very hard for an algorithmic search engine to decide which results are more relevant between highly ranked Web results and mobile Web results. The desktop Web result will score higher in almost all algorithmic scoring systems. Mobile sites, which are much less inter-linked and which have much lower traffic, will score lower on these measures but - being mobile-optimized - may offer a much better and more actionable result to the mobile user.

For this reason, human editing of search results offers even more potential for mobile search than it does for desktop search. Humans can differentiate between desktop results that display well through a mobile transcoding engine and those that display badly. Judging which desktop results and which mobile Web results should be mixed into the best possible mobile search results list can be based on inspection of those results by people on real devices. Humans could also optimize the page layout of results on the small screen. Expect human editing and human boosting of search results to play an important role in mobile search in the future.

In the next section we will take a look at the 4th and final aspect of "mobile search made social": the crawling and indexing of social information.

taptu.

8. Crawling and indexing social info

The most frequent mobile search that takes place today is the search for a friend's phone number in the handset's address book. This is not mobile Internet search, but it is a very important type of search nonetheless. It confirms in a very clear way the need that people have to find social information on the mobile.

From which other sources might we want to find social information from? The next most important repository of social information is our preferred social network. I'm a Facebook user and if I search for 'Bob' (see below), I want to see a results list that puts my friend Bob from Taptu at the top of the list, ahead of any Bobs that are not in my Friends list. Facebook's search function works in exactly this way. It has access to my social graph, and can rank search results accordingly.

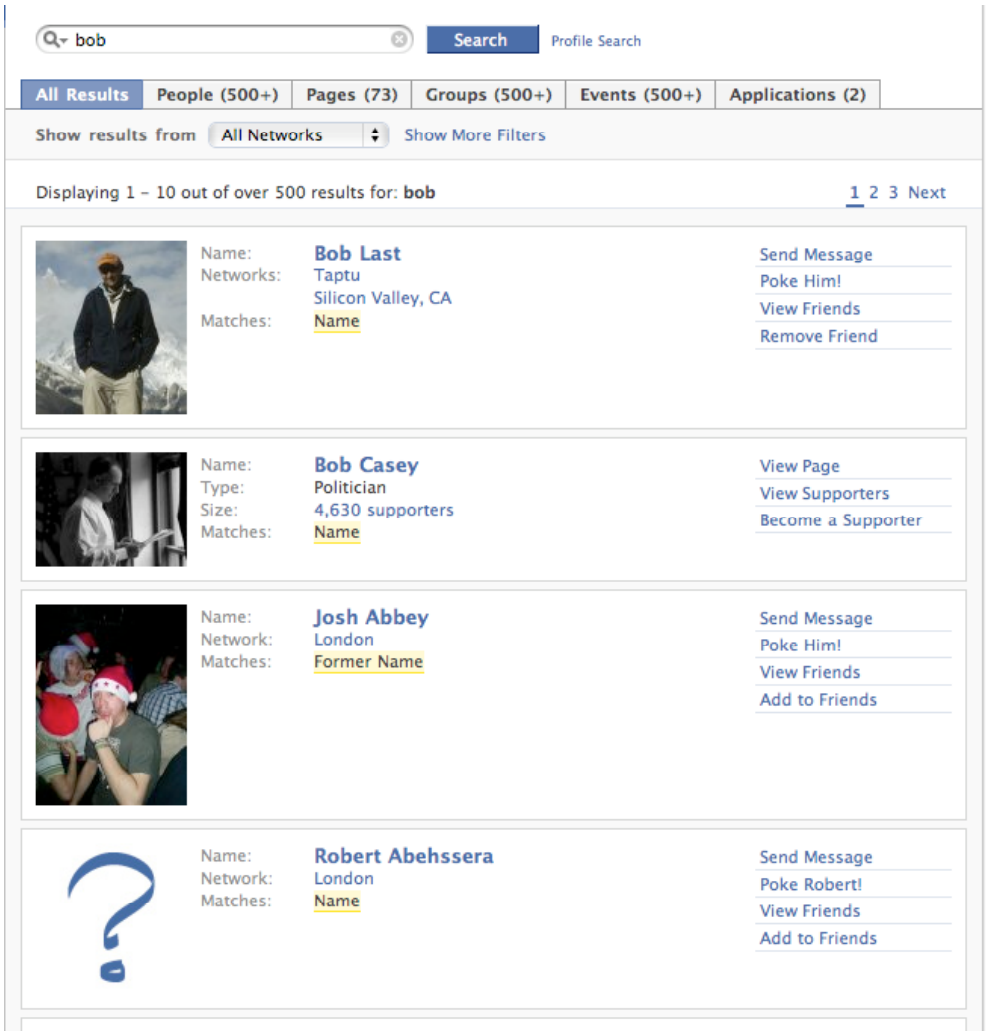

Fig 12: Searching for 'Bob' in my Facebook account (with Bob's permission ☺)

Search engines cannot do this today. Social networks don't want their users' private information to be copied and redistributed without permission, and they definitely don't want it to be indexed and crawled by public search engines. Sometimes they don't even want it to be copied even with the permission of the owner. In a notorious incident in January 2008, Facebook deleted blogger Robert Scoble from its social network after Scoble ran a Plaxo script[24]. This script logged into his account and ran screen-scraping software that exported Scoble's friends' details out into the Plaxo service.

Despite the murky waters that surround this issue, the portability of users' social information (the "social graph") is necessary and is coming. Examples of user-initiated portability that are allowed today are the bulk import routines that many applications (Facebook included) have developed for importing email addresses. Facebook does this to help users identify potential new friends when they first join Facebook, and it clearly offers a benefit to the user who would find it very time-consuming to do this manually.

There is an emerging need to search not only your own social network information, but to search friends' information (e.g. profiles, photos) and to find friends of friends, subject of course to the appropriate permissions being set by the owners of this data. Dataportability.org (http://www.dataportability.org) is a recent initiative set up to promote standards for data portability for information held in social networks. Facebook, Google and Plaxo signed up shortly after the resolution of the Scoble incident.

Once agreed methods emerge for data portability that are supported by the main social network players, then this data will be available to be crawled by search engines. Mobile search engines have the greatest need to crawl and index this information, so expect this capability to show up first on mobile.

9. Taptu: a roadmap for making mobile search social

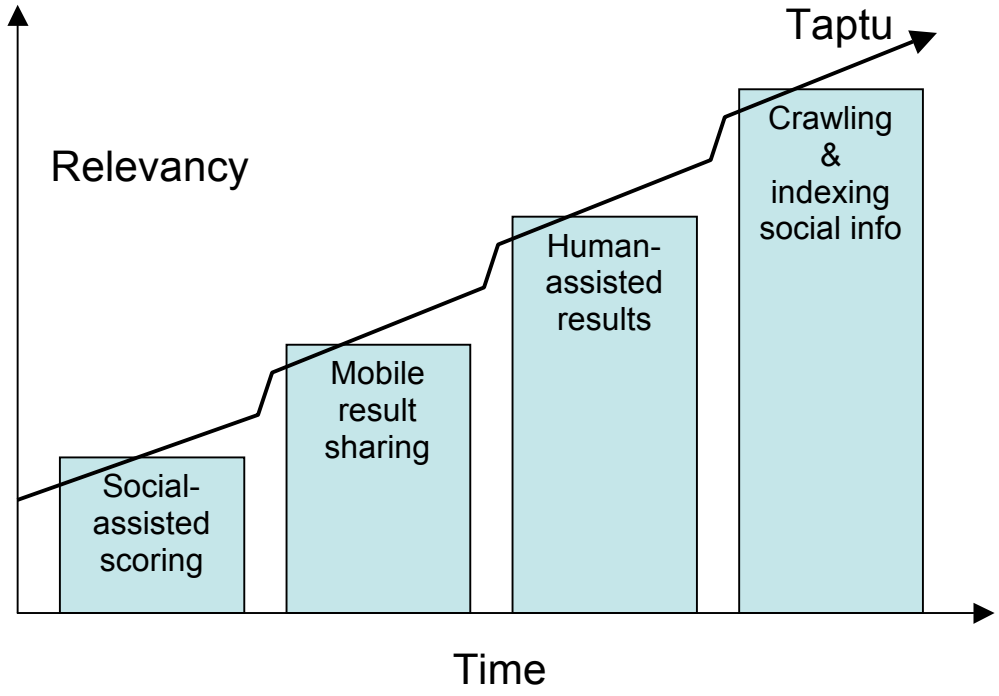

Fig 13: Taptu mobile search roadmap

At Taptu, we believe that mobile search must work well in a social context if it is ever going to cross the gap to the mass market. Unlike today, where mobile search is typically used just once a day or once a week, mass market mobile search will be used at least five times a day by a large proportion of mobile users. To bring about this transformation a number of innovations will be required.

Some of these innovations sit in the underlying platform and are not directly linked to the social dimension. The ability to make song and video results more useful by providing playable summaries is one example. The provision of a rich, mobile screen-sized summary for each search result is another.

Other needed innovations are driven by the requirement to make mobile search work well in a social context. We have reviewed four key required innovation steps in the last four sections of this White Paper:

- Social-assisted scoring
- Sharing mobile search results
- Human-assisted results
- Crawling & indexing of social info

The first step, social-assisted scoring, was delivered in the first Taptu Beta service that went live in October 2007.

The second step, mobile result sharing using a novel 1-Tap sharing concept, went live on Taptu in February 2008. It is described in more detail below.

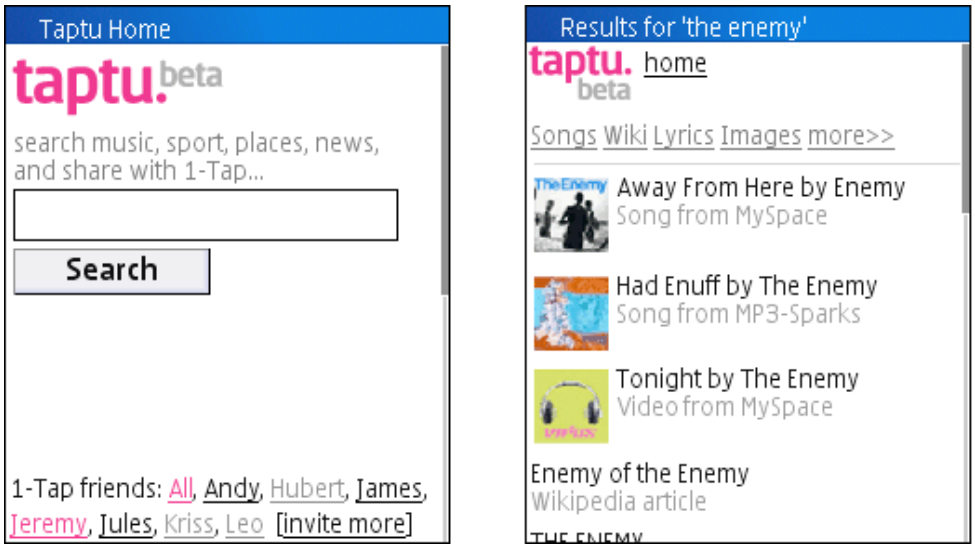

Fig 14: Taptu 1-Tap sharing: home screen and an individual result summary screen

In Taptu, when the user searches for content they first of all see a results page, and then when they click on a desired result they see a screen-sized result summary, which gives them a number of options. They can go directly to the source site; they can play a preview of the result (if it is a song or video), or then can share it with one or many friends. Each of these options is accessible with just one click.

Like e-commerce sites with 1-click buy buttons, where you first of all have to configure your credit card details, there is a configuration process in Taptu where you set up a list of 1-Tap friends. To save time, these can be imported from email systems like Gmail or mobile social services like Twitter.

Once this list of friends is set up, it is displayed in two main places in the Taptu UI:

- On the search home page
- In the action box on the search result summary screen

The Taptu service displays the friends list on the search home page to give clear notification of who has shared a search result with you in the main screen in Taptu, and to make it as simple as possible for the user to view that result.

The design of the action box on the search result summary screen makes it possible to share search results one-to-one and one-to-many. In both cases, in order to use the 1-Tap share functions the user needs to set up a user account on Taptu, and login to the service. Once logged in, the user can configure the service to auto login on subsequent visits to Taptu.

Taptu's 1-Tap result sharing system is an important step forward in creating a mobile search engine that works well in a social context.

The third required innovation step, human-assisted results, and the fourth step, the indexing of social information, are part of our future roadmap.

Concluding remarks

We stand at the beginning of the transition to a new wireless ecosystem, a transition that will unfold progressively over the next five years. This ecosystem is characterized by 3G networks, mobile phones that carry music and video collections, easier user interfaces and open mobile Internet access.

As Peggy Anne Salz, Publisher and Chief Analyst at MsearchGroove[25] observes: "The worlds of mobile social networking and mobile search are colliding, producing opportunities for companies to tap the minds of community members – both implicitly and explicitly – for much better quality results".

The generation of consumers who will be most active in this new ecosystem – Generation Y – have needs that are different to yesterday's consumers. They have learned to express themselves through social networks, instant messaging and blogs. They have grown up with Google and mobile phones and are totally at ease with technology. For Generation Y the mobile phone is becoming the supersocial device that will allow them to be constantly connected. If mobile search can fulfil its promise, then it will become "the user interface to everything" in the new ecosystem.

taptu.

References

Section 1: Introduction

[1] The Search, by John Battelle, see p82 (2005)
[2] Google Corporate Information, see http://tinyurl.com/3dk4x (2007)

Section 2: A new wireless ecosystem

[3] Mobile vs Fixed Internet Users – its all about the browser, by Dean Bubley, see http://tinyurl.com/2xkmcj (2008)

[4] Rudy de Waele presentation on the Mobile 2.0 ecosystem, see http://tinyurl.com/2dfyfg (2007)

[5] Flash Memory vs Hard Drives. Which Will Win? By Jim Handy, see http://tinyurl.com/yt2s4b (2006)

[6] The Tipping Point, by Malcolm Gladwell, see http://tinyurl.com/osc48 (2000)

[7] The Audience for Mobile Advertising, Q3 2006 presented by M:Metrics at MMA Workshop, 6 Dec 2006 (2006)

[8] Wireless Equipment Industry Outlook, Worldwide Mobile Forecasts 2005 – 2010, by M. Hoffman and J. Baxter, Cowen & Co (2007)

[9] The Inmates are Running the Asylum, by Alan Cooper, see http://tinyurl.com/2446ah (1999)

[10] Some Nokia N95 and S60 Thoughts, by Russell Beattie, see http://tinyurl.com/2bhzed (2007)

[11] The Future of Voice, presentation by Christian Lindholm at Mobile Monday London, December 3 2007 (2007)

[12] The iPhone, dotMobi and the Future of the Mobie Web, by Dan Appelquist, http://tinyurl.com/2fmaus (2007)

[13] Mobile vs Fixed Internet Users – its all about the browser, by Dean Bubley, see http://tinyurl.com/2xkmcj (2008)

[14] Alex Gibson interview with John Strand, 7 Oct 2006, see http://tinyurl.com/yv3dn4 (2006)

[15] OpenGardens blog by Ajit Jaokar, see http://opengardensblog.futuretext.com/ (2006)

Section 3: From X to Y: the demographic factors

[16] Social Networking Websites and Teens: an Overview, by A. Lenhart and M. Madden, Pew Internet & American Life Project, see http://tinyurl.com/2cxw7m (2007)
[17] Who's Filling Gen-Y's Shoes?, by P. Markiewicz, see http://tinyurl.com/2obvpd (2003)
[18] Who's Filling Gen-Y's Shoes?, by P. Markiewicz, see http://tinyurl.com/2obvpd (2003)
[19] A New Switched-On and Cynical Generation by Peter West, see http://tinyurl.com/24dady (2005)
[20] Generation Y: They've arrived at work with a new attitude, by Stephanie Armour, USA Today, see http://tinyurl.com/8pmyh (2005)

Section 5: Social-assisted scoring

[21] What is Web 2.0? by Tim O'Reilly, see http://tinyurl.com/743r5 (2005)

Section 6: Sharing mobile search results

[22] B. Smyth and C. Cotter, 'The Plight of the Navigator: Solving the Navigation Problem for Wireless Portals', in *Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'02)*, (2002).

Section 7: Human-assisted results

[23] The Fat Tail Will Be Human, The Medium Tail Social, The Long Tail Algorithmic, by Chris Anderson, see http://tinyurl.com/22pmr5 (2008)

Section 8: Crawling and indexing social info

[24] I've Been Kicked Off of Facebook, by Robert Scoble, see http://tinyurl.com/26g2y2 (2008)

Section 9: Taptu: a roadmap for making mobile search social

[25] MsearchGroove, the industry source of news, analysis and commentary on mobile search & social media, see www.msearchgroove.com