

SEO for Web 2.0

Enterprise Search Summit West

September 2008



ascentium

Agenda

Search 0.0

Search 1.0

Web 2.0

Search 2.0

- Data the next “Intel inside”
- Harnessing the Collective Intelligence
- Rich User Experience
- Software Above the Device

SEO for Search 2.0

SEO 0.0

Human-mediated information seeking

- Ask a friend
- Ask Mom
- Look it up in the Encyclopedia
- Ask more friends
- Ask the librarian

Finding what you were looking for depended on how honest you were with the source



Search 1.0: The Birth of Relevance

Yahoo starts with a human-mediated directory of online resources personally reviewed by librarians

Technology mediated information finding follows

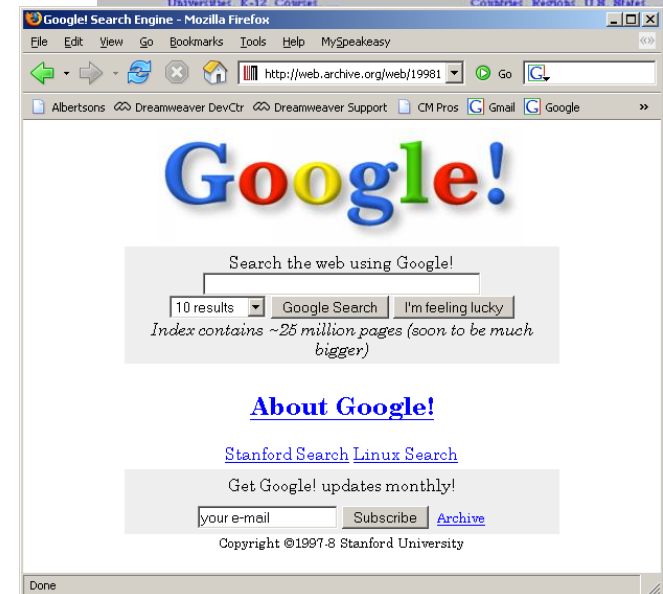
- Excite
- Altavista
- Northern Lights

Search engines for search engines come online [meta search engines]

- Dogpile
- Clusty

Google Beta Launch 1997

- Stripped down interface compared to competitors
- User interaction limited to query and # of results viewed
- “Did you mean? – spell check, synonyms



Web 2.0

Tim O'Reilly introduced 2005

Components

- Web as a platform
 - Everything happens in the browser
- Harnessing collective intelligence
 - Blogging and the wisdom of crowds
- Data is the next “Intel inside”
 - Mash-up technologies
- End of the software release cycle
 - Software as a Service
 - Linux and the open source movement
 - Agile, Scrum and the like
- Lightweight programming models
 - SOAP
 - REST
- Software above a single device
 - Mobile, mobile, mobile, mobile
- Rich User Experience
 - Ajax, Silverlight, Flex, AIR
 - VOIP

Search 2.0

Data is the next “Intel inside”

- Semantic Web & relationship modeling algorithms
 - Hilltop
 - HITS
 - Orion (Where is Ori Alon?)
- Latent Semantic Indexing (LSI)
- Google History

Harnessing the collective intelligence

- Folksonomy – tagging
- Online bookmarking – Del.icio.us
- Blogs
- Social networking

Rich User Experiences

- Multimedia search
- Personalized search results

Software above the device

- Mobile Search
- Local search

Data/Intel: Semantic Web: Hilltop

Developed by Bharat and Mihaila at U of Toronto and adopted by Google in 2001

Method

- Quality of links more important than quantity of links
- Segmentation of corpus into broad topics

Subset that is then extrapolated to Web as a whole

Selection of authority sources within these topic areas with authorities have lots of unaffiliated expert document on the same subject pointing to them

- Hubs are navigation pages that focus on authorities for a certain topic
Pages that point to many authority sources

Data/Intel: Semantic Web: Topic-Sensitive PageRank

Context sensitive relevance ranking based on a set of “vectors” and not just incoming links

Pre-query calculation of factors based on subset of corpus

- Context of term use in document
- Context of term use in history of queries
- Context of term use by user submitting query

Creator now a Senior Engineer at Google

Search 2.0

ascentium

Search 2.0: Automatic Query Expansion with LSI

Latent Semantic Indexing [LSI] is a weighting not retrieval algorithm

Documents with keywords in common are perceived as semantically related

- Pattern recognition, not sentience
- Recognizes terms that often appear together
- Co-occurrence

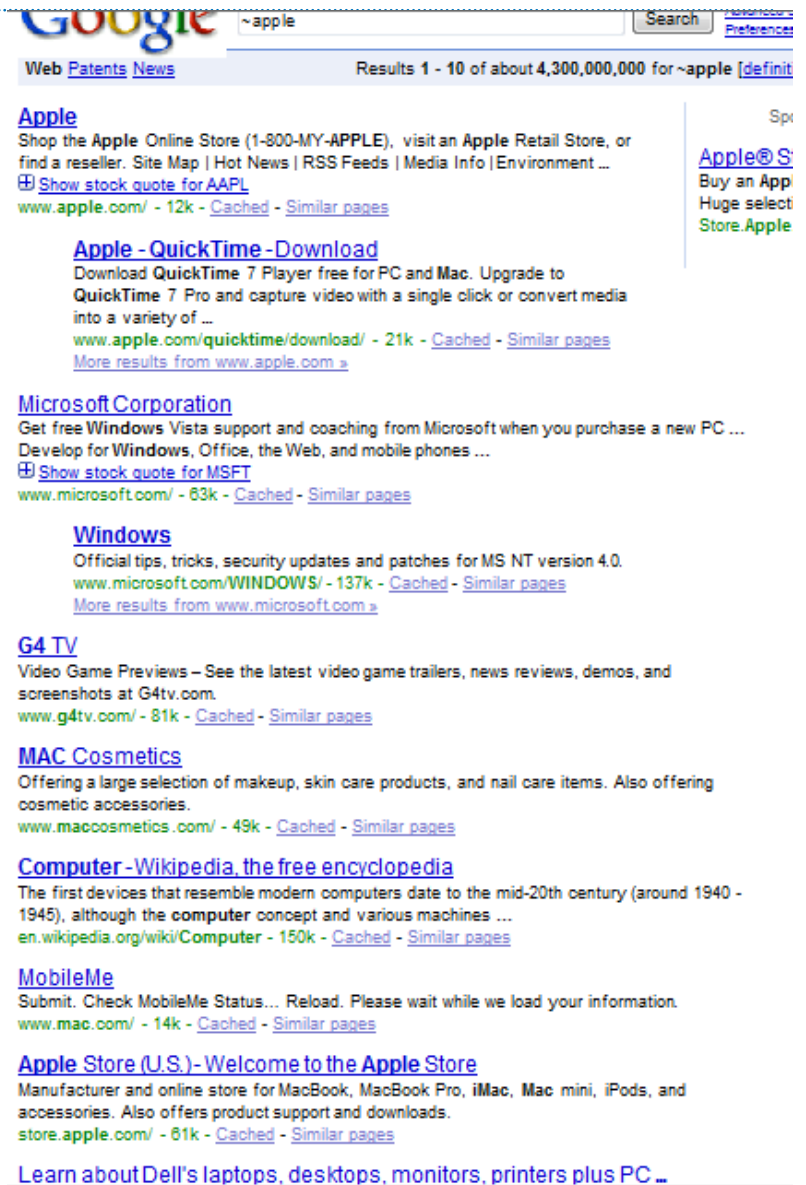
Search engine learns through repeated indexing of the document corpus

LSI on Google

Using a ~<search term> will initiate Google's LSI and produce a list of results that contains your original term as well as documents that the search engine determines are relevant to your query.

In example:

- There are no listings for “apple” the fruit – the search engine has been trained to associate “apple” singular with the computer company
- The #2 result is Microsoft – the search engine has been trained to associate the term Apple with Microsoft
- The #4 result is for Mac cosmetics because the search engine does not know the difference between Mac computers and Mac mascara.



The screenshot shows a Google search results page for the query '~apple'. The search bar at the top contains '~apple' and the search button is labeled 'Search'. Below the search bar, the text reads 'Results 1 - 10 of about 4,300,000,000 for ~apple'. The results are listed as follows:

- Apple**
Shop the **Apple** Online Store (1-800-MY-APPLE), visit an **Apple** Retail Store, or find a reseller. Site Map | Hot News | RSS Feeds | Media Info | Environment ...
[Show stock quote for AAPL](#)
www.apple.com/ - 12k - [Cached](#) - [Similar pages](#)
- Apple - QuickTime - Download**
Download **QuickTime** 7 Player free for PC and Mac. Upgrade to **QuickTime** 7 Pro and capture video with a single click or convert media into a variety of ...
www.apple.com/quicktime/download/ - 21k - [Cached](#) - [Similar pages](#)
[More results from www.apple.com »](#)
- Microsoft Corporation**
Get free **Windows** Vista support and coaching from Microsoft when you purchase a new PC ... Develop for **Windows**, Office, the Web, and mobile phones ...
[Show stock quote for MSFT](#)
www.microsoft.com/ - 63k - [Cached](#) - [Similar pages](#)
- Windows**
Official tips, tricks, security updates and patches for MS NT version 4.0.
www.microsoft.com/WINDOWS/ - 137k - [Cached](#) - [Similar pages](#)
[More results from www.microsoft.com »](#)
- G4 TV**
Video Game Previews – See the latest video game trailers, news reviews, demos, and screenshots at G4tv.com.
www.g4tv.com/ - 81k - [Cached](#) - [Similar pages](#)
- MAC Cosmetics**
Offering a large selection of makeup, skin care products, and nail care items. Also offering cosmetic accessories.
www.macosmetics.com/ - 49k - [Cached](#) - [Similar pages](#)
- Computer - Wikipedia, the free encyclopedia**
The first devices that resemble modern computers date to the mid-20th century (around 1940 - 1945), although the **computer** concept and various machines ...
en.wikipedia.org/wiki/Computer - 150k - [Cached](#) - [Similar pages](#)
- MobileMe**
Submit. Check MobileMe Status... Reload. Please wait while we load your information.
www.mac.com/ - 14k - [Cached](#) - [Similar pages](#)
- Apple Store (U.S.) - Welcome to the Apple Store**
Manufacturer and online store for MacBook, MacBook Pro, iMac, Mac mini, iPods, and accessories. Also offers product support and downloads.
store.apple.com/ - 61k - [Cached](#) - [Similar pages](#)

At the bottom of the page, there is a link: [Learn about Dell's laptops, desktops, monitors, printers plus PC ...](#)

Search 2.0: Orion Algorithm

Automatic Query Expansion

Purchased by Google in April 2006 for A LOT of money

- Developed by Ori Alon, an Israeli computer science doctoral fellow who has not been heard from since going to Goolge-opolis

Results include expanded text extracts from the websites

Integrates results from related concepts into query results

Rumored to be part of the 2006 Big Daddy Update

Relational Content Modeling

Search 2.0: Google History

Search engine factors in additional data to determine relevance to the query

Works off of Google Toolbar and Igoogle

- How closely did you read that Terms of Agreement?

History/behavior collection includes:

- Document – topicality, created date, content changes, time between changes
- Link – trustworthiness of links, Authority/Hub weighting, # of in links, # of out links, anchor text
- Previous behavior – how often is the document selected from results, how long user stays on page
- Query analysis – how often document is included in results set, movement up and/or down in placement
- Traffic patterns – topicality of query, obviously “the Olympics” is a specialized query
- Domain – age, address of named servers, etc



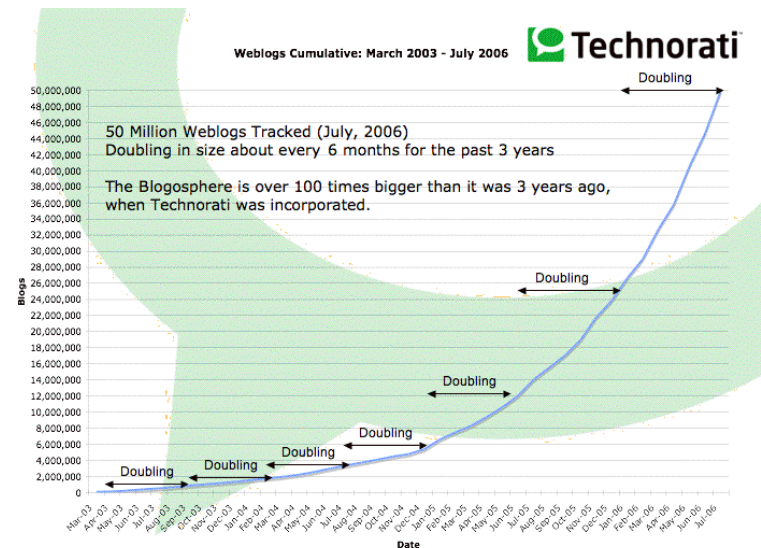
Search 2.0: Social Search

A significant portion of colleagues use social media

- 8% of Americans are deep users of the participatory Web and mobile applications
- Another 23% are heavy, pragmatic tech adopters – they use gadgets to keep up with social networks or be productive at work
- 10% rely on mobile devices for voice, texting, or entertainment

Blogs are a common means of sharing thoughts and ideas among friends, colleagues and strangers who are interested in same subject

- Number of blogs doubles every 6 months
- Customers talking about you and or your products and service



<http://www.sifry.com/alerts/archives/000436.html>

Search 2.0: Search Independent from a Single Device: Mobile Search

Google Mobile Search Study

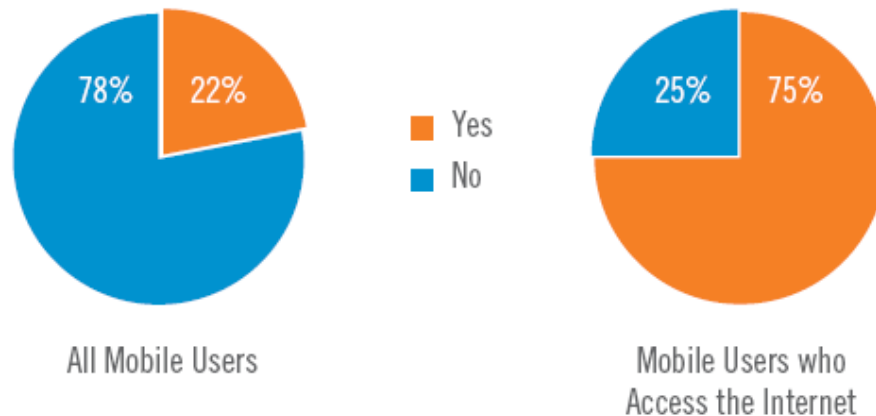
- 40% of mobile customers in US use search engines to find information with their devices*
- 12 million active US users in May 2008*

Yahoo Fire Eagle: location aware search

MOBILE SEARCH ACTIVITY AMONG MOBILE USERS AND MOBILE USERS WHO ACCESS THE INTERNET ON THEIR MOBILE DEVICE

Question: Do you search the Internet on your mobile device? (yes, no)

Base: U.S. mobile users ages 16+ (n=1,001) / Mobile users who access the Internet on their mobile device (n=300)



* Nielsen Research: Critical Mass of the Mobile Web: July 2008

Search 2.0: Rich User Experience

Applications

- Flash
- Silverlight
- AJAX

Search Engines

- www.cuil.com
- <http://www.searchme.com/>

Video

Images

The screenshot shows the cuil search engine interface. At the top, there is a search bar with the text 'enterprise search summit west' and a 'Search' button. To the right of the search bar are links for 'Preferences' and the cuil logo. Below the search bar, it indicates '703 results for enterprise search summit west'. The results are displayed in a grid format. The first result is 'Enterprise Search Summit West 2008' with a brief description and a link to 'www.enterprisesearchsummit.com/west...'. The second result is 'ContentBlogger: KM World 2007 Adds Enterprise Search ...' with a link to 'www.shore.com/commentary/weblogs/20...'. The third result is 'Taxonomy Boot Camp 2008' with a link to 'www.taxonomybootcamp.com/'. The fourth result is 'KMWorld & Intranets 2007/Enterprise Search Summit ...' with a link to 'www.kmworld.com/...'. On the right side of the results, there is a 'Explore by Category' dropdown menu with options for 'Internet Search Engines' and 'Software Companies'. Below the menu, there are several sponsored links, including 'Ravi Mynampaty' and 'Taxonomy Boot Camp 2008'.

The screenshot shows the searchme search engine interface. At the top, there is a search bar with the text 'rise search summit west'. To the left of the search bar are tabs for 'web', 'video', and 'images'. To the right of the search bar are tabs for 'advertising & marketing', 'companies', and 'blogs'. Below the search bar, it indicates 'search all'. The results are displayed in a grid format. The first result is 'Enterprise Search Summit West 2008' with a brief description and a link to 'www.enterprisesearchsummit.com/west...'. The second result is 'ContentBlogger: KM World 2007 Adds Enterprise Search ...' with a link to 'www.shore.com/commentary/weblogs/20...'. The third result is 'Taxonomy Boot Camp 2008' with a link to 'www.taxonomybootcamp.com/'. The fourth result is 'KMWorld & Intranets 2007/Enterprise Search Summit ...' with a link to 'www.kmworld.com/...'. On the right side of the results, there is a 'Explore by Category' dropdown menu with options for 'Internet Search Engines' and 'Software Companies'. Below the menu, there are several sponsored links, including 'Ravi Mynampaty' and 'Taxonomy Boot Camp 2008'.

SEO 2.0

ascentium

SEO 2.0: Data is the next “Intel inside”

Customers don't know what they don't know

- Facets to narrow results set
- Allow searcher to be part of relevance ranking

Know your customer

- Search logs
- Web analytics
- Personas

Federated search

Search engines like content

- More content = more representation in index = more opportunity to appear in search results
- Deep subject-specific content is what makes Authority resources
- Newspaper model

Know how search engines “see your website”

Titles

- Descriptions
- Keywords

SEO 2.0: Harnessing the Collective Intelligence

Folksonomy – tagging

- Relevance is subjective – what means something to you means something else to your customer
- Tags reveal priorities and understanding

Make it easy to use online bookmarking (Del.icio.us) and online sharing sites (Technorati, Digg, Spinn, etc)

- Pareto's 80/20 rule
- Bookmarks reveal the 20% that resonates with your customers

Blogs

- Talking about you and or your products and services
- Participate in the conversation
- Blog search engines
- Reputation Management software

Sharing Is Good.



SEO 2.0: Social Networking

Know what your customers are saying

- Visit social networking site and use your product or service as search term
 - Facebook
 - Linked in
 - Technorati
- Visit online bookmarking sites (Del.icio.us) and use your domain as a search term. See what customers are bookmarking and what user tags are associated with it
- Visit online bookmarking sites (del.icio.us) and use terms that you believe best describe your product/services and see what sites users have bookmarked

SEO 2.0: Search Independent from a Single Device: Mobile Search

Optimize sections of website for mobile search

Get mobile sub-domain (i.e. mobile.ascentium.com)

Use mobile sitemap.xml to ensure indexing

Use search engine mobile-specific submit service

Separate content from display with separate CSS for mobile



SEO 2.0: Search Independent from a Single Device: Local Search

Put locations on page as well as in footer and on Contact Us

List in local directories

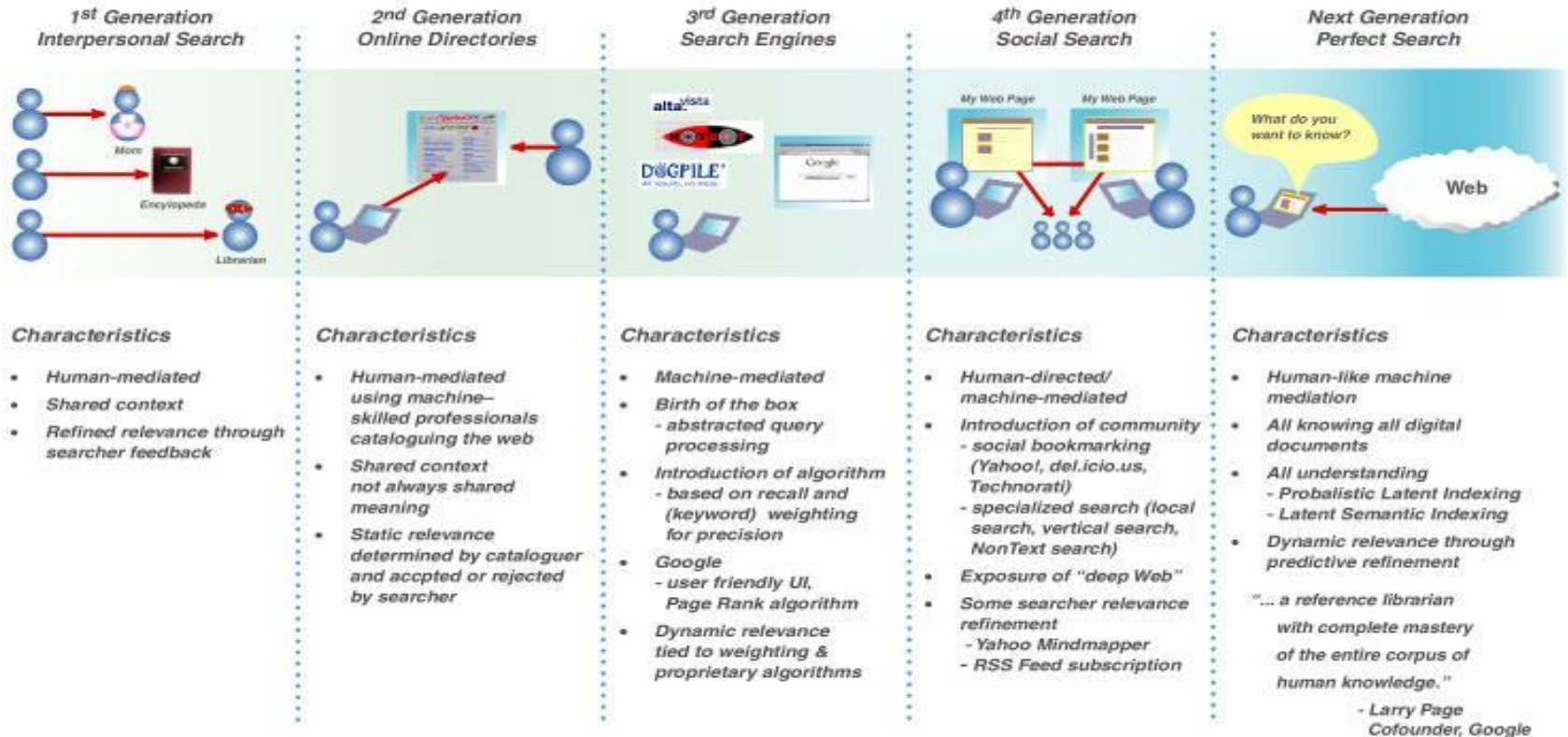
- Yellow pages
- Seattle 24/7

Fill out your Google Local Business Center profile

Get links from local resources

Allow user generated content as they often mention “location-specific” terms/phrases

Where It Was – Where It Is Going



Design: Arlene Rubin