

How Search engines Work

An overview of the basics and some detail on the major engines.



Computer Productivity Consulting
Richard G. Myers

Search vs. Other Marketing

- Customers use search engines to research products 41% of the time, TV 9%, print ads 10%
- B2C companies surveyed indicated website leads were 62% more profitable than other media
- The Internet is the first choice of media for adults 18 to 54—more than TV.
- Case study: Using an integrated online mix of SEM, email, and affiliate marketing, the campaign drove site traffic where 88% were new visitors, and 85% of them purchased on their first visit
- Direct mail's cost-per-lead is \$10 while search averages just 29 cents.



Computer Productivity Consulting
Richard G. Myers

Search Engine Shoppers

- **Market share of all U.S. visits to Shopping and Classifieds sites reached their highest weekly levels during the week ending Dec. 11, 2004, 9.73%.**
 - Google: 4.26%
 - Yahoo! Search: 2.24%
 - MSN Search: 0.54%
- **Search engines are stronger in sending their users to varying retail segments:**
 - Google: Books, Sports and Fitness, and Music.
 - Yahoo! Search: Video and Games, Automotive and Classifieds.
 - MSN Search: Apparel & Accessories, House & Garden and Appliances & Electronics.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Search Engines Drive Traffic to Shopping Sites, By Ina Steiner,
AuctionBytes.com, December 15, 2004,
<http://www.auctionbytes.com/cab/abn/y04/m12/i15/s03>

Basic Operations

- **Gather data from the Internet.**
- **Analyze, index & file information about site.**
- **Receive queries from users.**
- **Search indexed information files.**
- **Apply ranking algorithms to sort.**



Computer Productivity Consulting
Richard G. Myers

Adding Pages to the Index

- All engines crawl a different segment of the Internet.
- Some pages are eliminated for “spamming”.
- Complaints from users can eliminate pages.



Computer Productivity Consulting
Richard G. Myers

Keyword Relevance

- **Keywords searched for appear in index.**
- **HTML Title tag indicates relevance.**
- **Placement near the beginning shows importance.**
- **Frequency compared to other words.**



Computer Productivity Consulting
Richard G. Myers

Off the Page Factors

- **Links from other sites.**
 - “Artificial” links screened out.
- **Clickthrough measurement.**
 - High ranking pages may not attract traffic.



Computer Productivity Consulting
Richard G. Myers

Google Ranking Factors

- **Location & frequency of keywords.**
- **Text size gets a boost.**
- **Number, quality & context of links.**
 - Links from your site count.
- **Sometimes clickthrough is measured.**
- **Results are clustered.**
- **Penalties assessed for spamming.**
 - Use “link: ...” to see if you are banned.



Getting into Google

- **Submit new pages.**
 - Home page plus several inside pages.
 - No automatic submission (spam!)
- **Links from other sites.**
- **File types indexed include:**
 - Microsoft Word (doc), Excel (xls), PowerPoint (ppt), Rich Text Format (rtf), PostScript (ps).
- **Current Index Size:**
 - By August 2001, over 1.3 billion web pages
 - In early 2004, exceeded 4 billion pages.
 - 2005 Google - Searching 8,058,044,651 web pages



Computer Productivity Consulting
Richard G. Myers

Getting into Inktomi (Yahoo)

- **Link crawling.**
 - Links from external sources preferred.
 - Inktomi maintains a “webmap”.
 - More likely to index “shallow” documents.
- **Paid inclusion.**
 - Search Submit: smaller web sites.
 - Index Connect: large web sites who wish to list 1,000 or more pages.



Inktomi Ranking Factors

- **Keyword frequency in copy.**
- **Keywords in Title tag.**
- **Meta keyword & description will boost ranking.**
- **Link analysis.**
- **Clickthrough measurements.**
- **Internal editorial staff tweaking.**



Computer Productivity Consulting
Richard G. Myers

MSN

- **Paid listings from Overture.**
- **Editorial listings from Looksmart.**
 - MSN applies its own ranking.
- **Uses Inktomi engine at present.**
- **Highest rank goes to:**
 - MSN network, then advertisers, & then editorial preferences.



Computer Productivity Consulting
Richard G. Myers

MSN Search Characteristics

- Like Google, MSN's spider finds new sites by following links directed to those sites.
- MSN revisits sites very frequently as well.
- Over the past year, MSN has compiled a 5-Billion site database.
- MSN looks at the number of links directed to that site.
- It is assumed by most that anchor text plays a major role in the ranking algorithm.
- Sites with great text and clear internal link-paths are ranking very well with MSN.
- Strong, keyword enriched titles and body text continue to provide strong placements.



MSN Search Characteristics (cont.)

- **Size matters with MSN**
 - larger sites with long-term content appear to be doing very well under more generic keyword searches.
- **A website that has a large number of incoming links will get noticed and spidered a number of times.**
- **Google is taking a very refined approach to contextual-quality**
 - Yahoo and MSN seem more interested in the number of links.
- **Titles make a big difference at all three .**
 - MSN also seems to be able to read text found in drop-down menus such as the ones on the right hand side of the Field Turf index page.
- **MSN states on its "How MSN Search Works" page that pages that are active will be spidered more frequently and achieve stronger rankings.**



Online Shoppers Survey - 1

- **More than 72 percent of online shoppers use general search engines nine or more times per month,**
 - growing portion of those searches are shopping-related or local in nature,
- **74 percent of respondents perform local searches,**
 - 37 percent of online consumers are very familiar with shopping search sites.
- **search continues to be the dominant way people find things on the Internet**
- **Respondents said 27 percent of their total search behavior was for local information.**
 - When the data is extrapolated to the general Internet population, it shows that 20 percent of all searches are local, Sterling said.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Local and Shopping Search Grow, ClickZ News, Kevin Newcomb,
10/29/2004, <http://www.clickz.com/news/article.php/3428741>

Online Shoppers Survey - 2

- **Nearly half of respondents, 43 percent, said that they preferred to research and shop online,**
 - roughly 28 percent displayed some combination of offline/online behavior.
- **"As consumers use the Internet more to find local information, specifically business information, that will impact advertising revenue,"**
 - as consumer activity moves online, advertisers will surely follow.
 - "That's the \$14 billion question," he said, referring to the amount spent on yellow pages advertising by small businesses.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Local and Shopping Search Grow, ClickZ News, Kevin Newcomb,
10/29/2004, <http://www.clickz.com/news/article.php/3428741>

Online Shoppers Survey - 3

- **currently between 35 and 40 percent of small businesses with Web sites,**
 - only about 24 percent of small businesses advertise online.
- **Reaching those advertisers has been a challenge for Internet publishers and search engines.**
 - yellow pages, newspapers and radio each receive more than three-quarters of ad revenue from local advertising,
 - online advertising generates less than a third of its revenue from local ads.
 - That is not expected to change soon, the report found, indicating that short-term growth in local search online will be driven by national advertisers targeting locally.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Local and Shopping Search Grow, ClickZ News, Kevin Newcomb,
10/29/2004, <http://www.clickz.com/news/article.php/3428741>

Online Shoppers Survey - 4

- **local online advertising will grow 28 percent in 2004,**
 - from \$1.9 billion in 2003 to \$2.4 billion this year.
 - It is expected to reach \$4.9 billion in 2009.
 - Local advertising's share of total online ad spend is not expected to rise much beyond its current 29 percent mark.
- **Search companies are trying to reach local businesses by pairing with established small business service providers,**
 - Google is with BellSouth,
 - FindWhat with Canada's Yellow Pages Group.
 - Similarly, Web hosting providers Affinity Internet and Interland, teaming with yellow pages publisher Dex Media.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Local and Shopping Search Grow, ClickZ News, Kevin Newcomb,
10/29/2004, <http://www.clickz.com/news/article.php/3428741>

Online Shoppers Survey - 5

- **among the 3,887 online shoppers in the panel, Internet mapping sites rate as the top search activity,**
 - based on familiarity, frequency of use, and loyalty, followed by general search engines.
 - Only 35 percent indicated loyalty to one general search engine.
- **After mapping sites and general search,**
 - Internet yellow pages, online classifieds, shopping search sites and entertainment information sites round out the top five online search behaviors,
 - followed by online travel sites, local destination sites and vertical directories.



Computer Productivity Consulting
Richard G. Myers

Inside the Searcher's Mind - 1

- **Most marketers mistakenly envision a linear search process**
 - prospects conduct a search, get their results, view a site, and immediately convert.
 - actual user experience is dramatically different.
- **"search funnel."**
 - Searching starts with a generic, inclusive keyword,
 - the search becomes increasingly specific.
 - "Over 70 percent of participants start with a generic keyphrase and narrow it down from there,



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

Inside the Searcher's Mind - 2

- **actual search scenario**
 - searcher began her search using the word "cruise"
 - reviewing the initial results, the user refined her search for the phrase "Caribbean cruise."
- **awareness" phase.**
 - Characterized by searchers researching their options
 - narrowing their scope of interest,
 - chances of converting during this phase is highly unlikely (around one to two percent).



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

Inside the Searcher's Mind - 3

- **"awareness" phase poses a certain challenge for marketers.**
 - introducing a company's brand early in the search process often dictates the direction of a search, having an effect on the eventual outcome.
 - However, those "awareness" key terms may not realize stellar ROI, as actual conversion rates are low.
 - Some marketers may be tempted to slice those terms from their key phrase list, figuring that they aren't effective.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

Inside the Searcher's Mind - 4

- **Following the "awareness" phase is the "convert" phase,**
 - characterized by users researching options and
 - reviewing third-party testimonials.
 - Users are also reintroduced to brands during the "convert" phase.
- **During this phase, the user narrowed her search,**
 - reading third party reviews on Panama Cruises.
 - was reintroduced to the Princess Cruise Lines brand and their Panama cruises.
 - After determining that she was interested in a Princess cruise, she conducted her final search for "Princess Panama Cruise."



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

Inside the Searcher's Mind - 5

- This final targeted search has a 30-40 percent chance of conversion for Princess
- Additionally, if "Panama cruises" and the Princess brand weren't introduced early in the "awareness" stage, the searcher may never have refined her search in that direction.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

search engine behavior - 1

- 60.5 percent of respondents picked a natural search result as the one they found most relevant to a sample query.
- 60.8 percent of respondents who use Yahoo! picked a natural search result as "most relevant" to a sample query.
- 72.3 percent of respondents who use Google picked a natural search result as "most relevant."



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

search engine behavior - 2

- 22.6 percent of searchers will try another search after viewing the top listings.
- 18.6 percent will try another search after viewing the entire first page.
- 25.8 percent will abandon the query after checking the first two pages.
- 14.7 percent will relaunch a query after the first three pages.
- If a brand is not represented in the top three search results, that brand loses 80 percent of the online audience for that search query.
- "Rarely do searchers go past the third page,"



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

online branding

- **more than just relying on either PPC advertising or organic SEO.**
 - Even if 60.5 percent of respondents considered a natural search result more relevant, PPC is still important.
 - "Unequivocally, you have to do paid search as well," "Otherwise, you miss 30-40 percent of your audience."
- **The study also provided search marketers tips for targeting specific demographic groups.**
 - 43.1 percent of women chose a paid result as "most relevant" to a sample query.
 - If a marketer is targeting women, running paid search listings may result in search conversions.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

Which engine?

- **Google was perceived as the leader in terms of both customer experience and satisfaction.**
 - actual search results returned by Yahoo!, Google, Ask Jeeves, Lycos and MSN were not dramatically different,
 - Google users had a higher perceived rate of success and satisfaction.
- **"Search engines are at performance parity when success is measured objectively,"**
 - "Despite performance, people prefer Google."
- **The consumer preference, according to the report, could be due to how Google segregates sponsored from free listings.**
 - Google paid listings are clearly marked on the right side of the search results page.
 - However, Ask Jeeves and Lycos sponsored Web results appear before the free results.
 - "Users felt that they were 'trying to be fooled' to click on a sponsored link," said Mayer.



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Delving Deep Inside the Searcher's Mind, Heather Lloyd-Martin, 10/14/2004,
http://searchenginewatch.com/_subscribers/articles/article.php/3406901

Which engine? (2)

- **These results may tempt site owners to go gaga for Google and ignore other engines in their search marketing campaign.**
 - But beware this Google-centric technique.
 - 76 percent of users had a primary search engine, but users are not monogamous to just one engine.
 - When search expectations are not met, 47 percent of users turn to another search engine, and
 - up to 16 percent regularly use a different search engine for some searches.
 - Therefore, even engines with a lower market share (such as Ask Jeeves) still give sites good exposure.



Computer Productivity Consulting
Richard G. Myers

Search Engine Submissions

Search Engine	AllTheWeb	AltaVista	Google	Inktomi	Teoma
Listing tied to submit?	No	Yes	No, but may help	n/a	n/a
Submit limit	No limit	No limit	No limit	n/a	n/a
Pages appear	2-4 weeks	4-6 weeks	About 4 weeks	About 4 weeks	n/a
Overall freshness	1 day to 4 weeks	1 day to 6 weeks+	1 day to 4 weeks	1 day to 4 weeks	1 day to 4 weeks
Paid inclusion?	Yes	Yes	No	Yes	Yes
Inclusion/Refresh	3 days 2 days	2 days Daily	n/a	3 days 2 days	7 days 7 days
Cost per year	\$20 first, ~\$15 add	\$78 first, \$38-\$58 add	n/a	\$39 first; \$25 add	\$30 first; \$18 add
Cost for 1, 10, 100 url's	\$20 \$148 \$1,431	\$78 \$600 \$3,982	n/a	\$39 \$264 \$2,514	\$30 \$192 \$1,812
Self-Serve limit	1,000	500	n/a	1,000	1,000
Bulk program	Yes	Yes	n/a	Yes	Yes
Powers	See Who Powers Whom? Chart				
Updated	Sept. 15, 2003	Sept. 15, 2003	Sept. 15, 2003	Sept. 15, 2003	Sept. 15, 2003



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Crawler Submission Chart., Editor, September 15, 2003 ,
http://searchenginewatch.com/_subscribers/article.php/2149081



Directory	<u>Yahoo</u>	<u>Open Directory</u>	<u>LookSmart</u>
Title Length	60 characters	100 characters	65 characters
Description Length	25 words / 200 characters	25 - 30 words	170 characters
How Do I Submit?	"Suggest a Site" link at bottom of page	"add URL" link at top of page	"Get Listed" link on home page or "Submit a Site" link near top of categories
Required Info	Usual + Name, Email	Usual	Usual
Optional Info	Location, Comments	Email	Tracking URL, relevancy keywords
Maximum Categories	1 per site	1 per URL	1 per URL
Submission Fee?	\$299 per year for commercial categories; \$600 per year for adult categories \$299 optional one-time fee for non- commercial categories	None	\$29 + \$15 per month min in per click charges for commercial sites; Use Zeal.com for non-commercial sites
Turnaround Time (Paid)	7 days	n/a	5 days
Turnaround Time (Free)	2 weeks	3 weeks	1 week or less
Major Partners	None	AOL, Google, HotBot, Lycos, Netscape	AltaVista, MSN Search
Updated	Feb. 20, 2002	Feb. 20, 2002	April 29, 2003



Computer Productivity Consulting
Richard G. Myers

Directory Submission Chart, By Danny Sullivan, Editor, April 29, 2003
http://searchenginewatch.com/_subscribers/article.php/2149071

{How_SE_Work}

Who Powers Whom

Search Engine (Read Down)	Provider: Google	Provider: Yahoo/Overture	Notes	
Google	Main & Paid		Open Directory an option	25%+
Yahoo		Main & Paid	Yahoo Directory an option	
MSN		Main & Paid (12/05 & 6/05)		10%+
AOL	Main & Paid (est. 10/05+)		AOL-owned Open Directory an option	
Excite Network	Main & Paid (at iWon, MyWay, My Web Search)		Excite.com is InfoSpace-powered	1%+
Ask Jeeves	Paid (until 2007)		Main from Ask-owned Teoma.	
InfoSpace	Runs several meta search engines. Dogpile is most popular, representative of others. Google (2006), Yahoo (3/06) & many small providers have distribution deals.			



Computer Productivity Consulting
Richard G. Myers

Who Powers Whom? Search Providers Chart, By Danny Sullivan, Editor, July 23, 2004, <http://searchenginewatch.com/reports/article.php/2156401>

Who Powers Whom

Search Engine (Read Down)	Provider: Google	Provider: Yahoo/Overture	Notes
AltaVista		Main & Paid	Open Directory an option; owned by Yahoo
AllTheWeb		Main & Paid	Owned by Yahoo
HotBot	Paid (see note)	Main	Backup from Google & Ask; Owned by Lycos
Lycos	Paid (see note)	Backup (see note)	Main from LookSmart; Open Directory an option
Netscape	Main & Paid (est 10/05+)		Owned by AOL; Open Directory an option
Teoma	Paid (Sept 05)		Main from Teoma; owned by Ask; Paid can end as early as 9/04

<1%



Computer Productivity Consulting
Richard G. Myers

{How_SE_Work}

Who Powers Whom? Search Providers Chart, By Danny Sullivan, Editor, July 23, 2004, <http://searchenginewatch.com/reports/article.php/2156401>