

Google's PageRank and Beyond:

The Science of Search Engine Rankings

Amy Langville

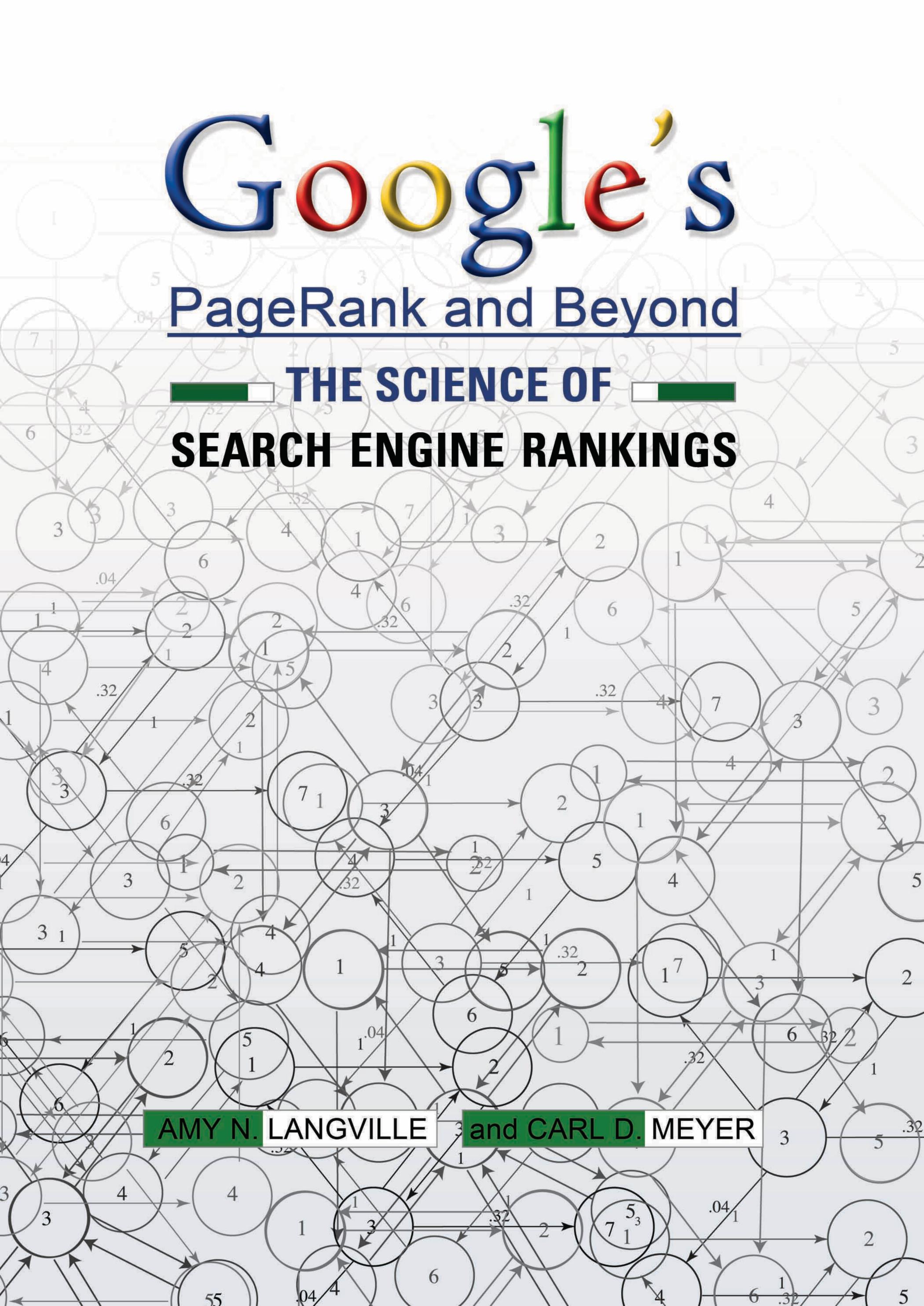
langvillea@cofc.edu

Department of Mathematics

College of Charleston

Charleston, SC

UIUC 1/25/07



Google's

PageRank and Beyond

THE SCIENCE OF
SEARCH ENGINE RANKINGS

AMY N. LANGVILLE

and CARL D. MEYER

Outline

- Introduction to Information Retrieval
- Elements of a Search Engine
- Link Analysis
- Current Issues in Web Search
- PageRank and You

Short History of IR

IR = search within doc. coll. for particular info. need (query)

B. C.	cave paintings
7-8th cent. A.D.	Beowulf
12th cent. A.D.	invention of paper, monks in scriptoria
1450	Gutenberg's printing press
1700s	Franklin's public libraries
1872	Dewey's decimal system Card catalog
1940s-1950s	Computer
1960s	Salton's SMART system
1989	Berner-Lee's WWW

the pre-1998 Web

Yahoo

- hierarchies of sites
- organized by humans

Best Search Techniques

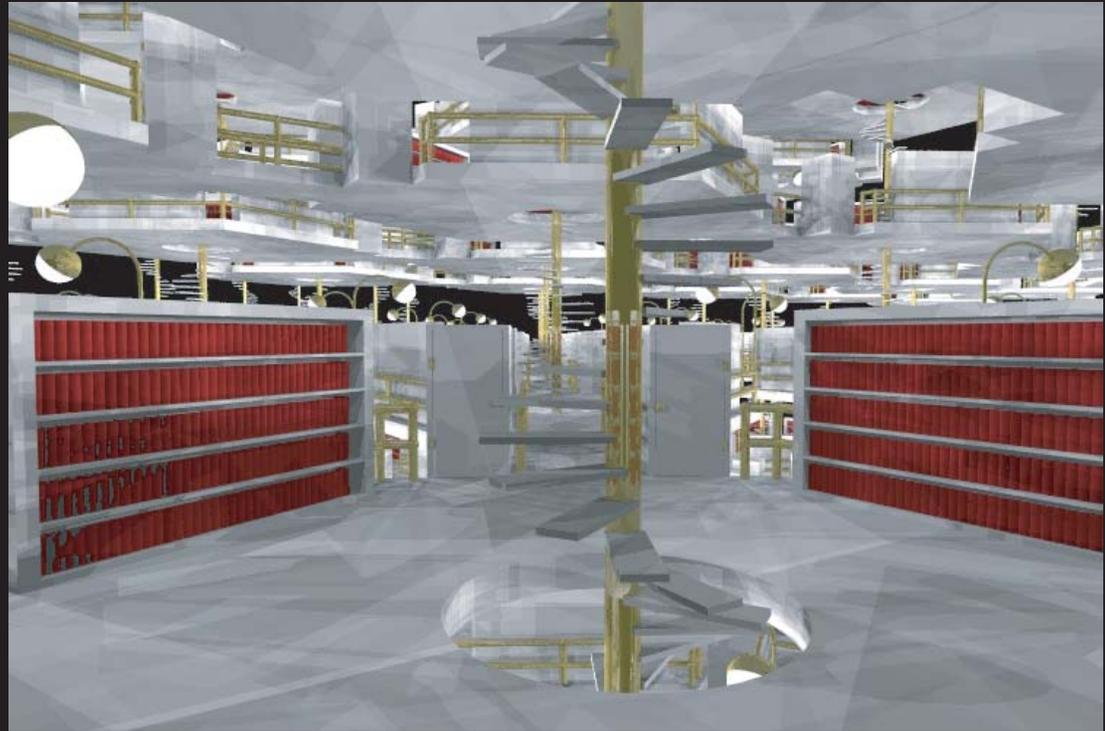
- word of mouth
- expert advice

Overall Feeling of Users

- Jorge Luis Borges' 1941 short story, *The Library of Babel*

When it was proclaimed that the Library contained all books, the first impression was one of extravagant happiness. All men felt themselves to be the masters of an intact and secret treasure. There was no personal or world problem whose eloquent solution did not exist in some hexagon.

... As was natural, this inordinate hope was followed by an excessive depression. The certitude that some shelf in some hexagon held precious books and that these precious books were inaccessible, seemed almost intolerable.



1998 ... enter Link Analysis

Change in User Attitudes about Web Search

Today

- “It’s not my homepage, but it might as well be. I use it to ego-surf. I use it to read the news. Anytime I want to find out anything, I use it.” - **Matt Groening, creator and executive producer, The Simpsons**
- “I can’t imagine life without Google News. Thousands of sources from around the world ensure anyone with an Internet connection can stay informed. The diversity of viewpoints available is staggering.” - **Michael Powell, chair, Federal Communications Commission**
- “Google is my rapid-response research assistant. On the run-up to a deadline, I may use it to check the spelling of a foreign name, to acquire an image of a particular piece of military hardware, to find the exact quote of a public figure, check a stat, translate a phrase, or research the background of a particular corporation. It’s the Swiss Army knife of information retrieval.” - **Garry Trudeau, cartoonist and creator, Doonesbury**

Web Information Retrieval

IR before the Web = traditional IR

IR on the Web = **web IR**

Web Information Retrieval

IR before the Web = traditional IR

IR on the Web = **web IR**

How is the Web different from other document collections?

Web Information Retrieval

IR before the Web = traditional IR

IR on the Web = **web IR**

How is the Web different from other document collections?

- It's huge.
 - over 10 billion pages, average page size of 500KB
 - 20 times size of Library of Congress print collection
 - Deep Web - 550 billion pages

Web Information Retrieval

IR before the Web = traditional IR

IR on the Web = **web IR**

How is the Web different from other document collections?

- It's huge.
 - over 10 billion pages, average page size of 500KB
 - 20 times size of Library of Congress print collection
 - Deep Web - 550 billion pages
- It's dynamic.
 - content changes: 40% of pages change in a week, 23% of .com change daily
 - size changes: billions of pages added each year

Web Information Retrieval

IR before the Web = traditional IR

IR on the Web = **web IR**

How is the Web different from other document collections?

- It's huge.
 - over 10 billion pages, average page size of 500KB
 - 20 times size of Library of Congress print collection
 - Deep Web - 550 billion pages
- It's dynamic.
 - content changes: 40% of pages change in a week, 23% of .com change daily
 - size changes: billions of pages added each year
- It's self-organized.
 - no standards, review process, formats
 - errors, falsehoods, link rot, and spammers!

Web Information Retrieval

IR before the Web = traditional IR

IR on the Web = **web IR**

How is the Web different from other document collections?

- It's huge.
 - over 10 billion pages, average page size of 500KB
 - 20 times size of Library of Congress print collection
 - Deep Web - 550 billion pages
- It's dynamic.
 - content changes: 40% of pages change in a week, 23% of .com change daily
 - size changes: billions of pages added each year
- It's self-organized.
 - no standards, review process, formats
 - errors, falsehoods, link rot, and spammers!

A Herculean Task!

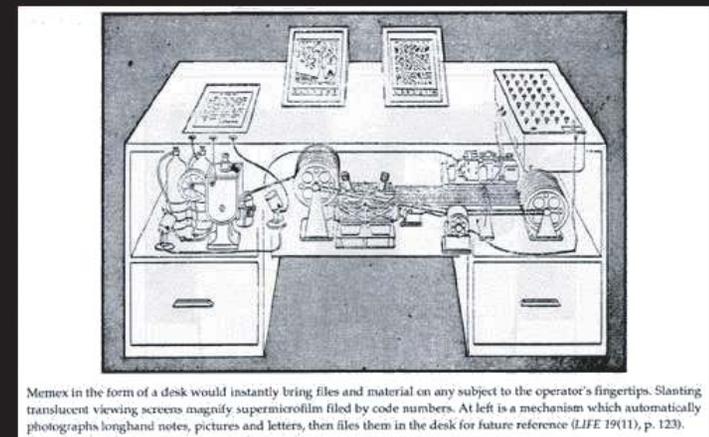
Web Information Retrieval

IR before the Web = traditional IR

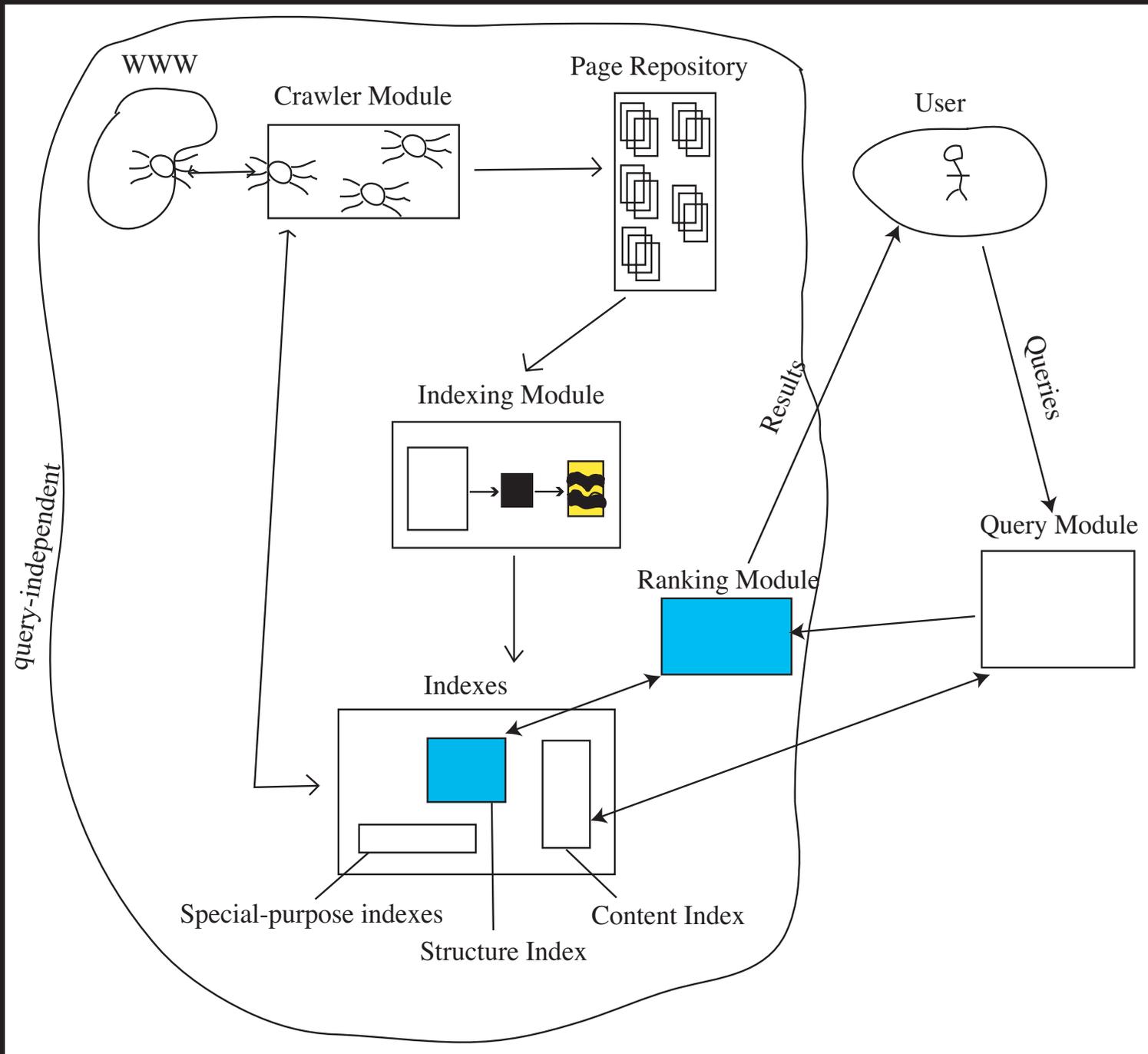
IR on the Web = **web IR**

How is the Web different from other document collections?

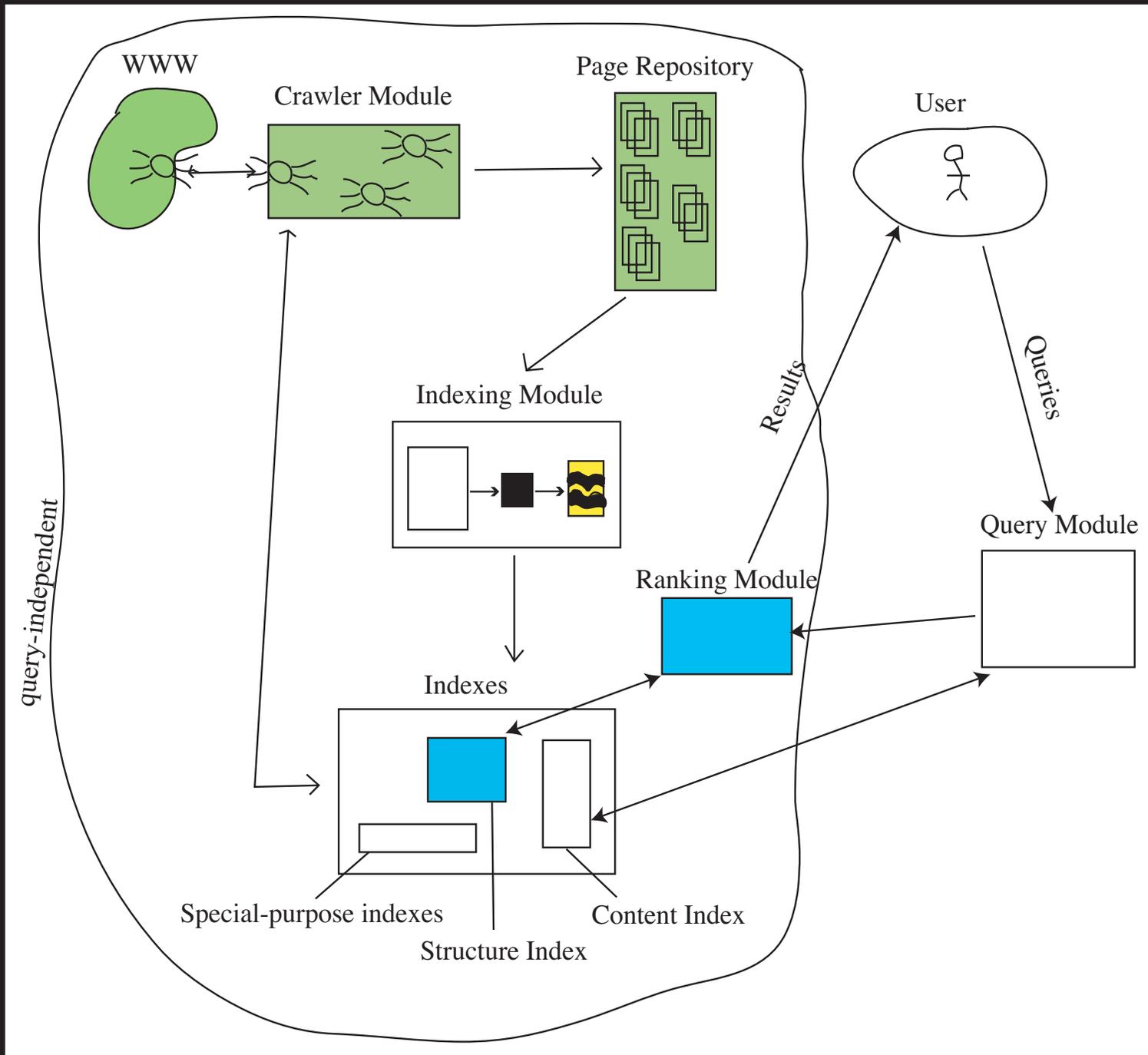
- It's huge.
 - over 10 billion pages, each about 500KB
 - 20 times size of Library of Congress print collection
 - Deep Web - 550 billion pages
- It's dynamic.
 - content changes: 40% of pages change in a week, 23% of .com change daily
 - size changes: billions of pages added each year
- It's self-organized.
 - no standards, review process, formats
 - errors, falsehoods, link rot, and spammers!
- Ah, but it's hyperlinked !
 - Vannevar Bush's 1945 memex



Elements of a Web Search Engine



Submitting your Site to a Search Engine





Add your URL to Google

[Home](#)

[About Google](#)

[Advertising Programs](#)

[Business Solutions](#)

[Webmaster Info](#)

▶ [Submit Your Site](#)

Find on this site:

Share your place on the net with us.

We add and update new sites to our index each time we crawl the web, and we invite you to submit your URL here. We do not add all submitted URLs to our index, and we cannot make any predictions or guarantees about when or if they will appear.

Please enter your full URL, including the `http://` prefix. For example: `http://www.google.com/`. You may also add comments or keywords that describe the content of your page. These are used only for our information and do not affect how your page is indexed or used by Google.

Please note: Only the top-level page from a host is necessary; you do not need to submit each individual page. Our crawler, Googlebot, will be able to find the rest. Google updates its index on a regular basis, so updated or outdated link submissions are not necessary. Dead links will 'fade out' of our index on our next crawl when we update our entire index.

URL:

Comments:

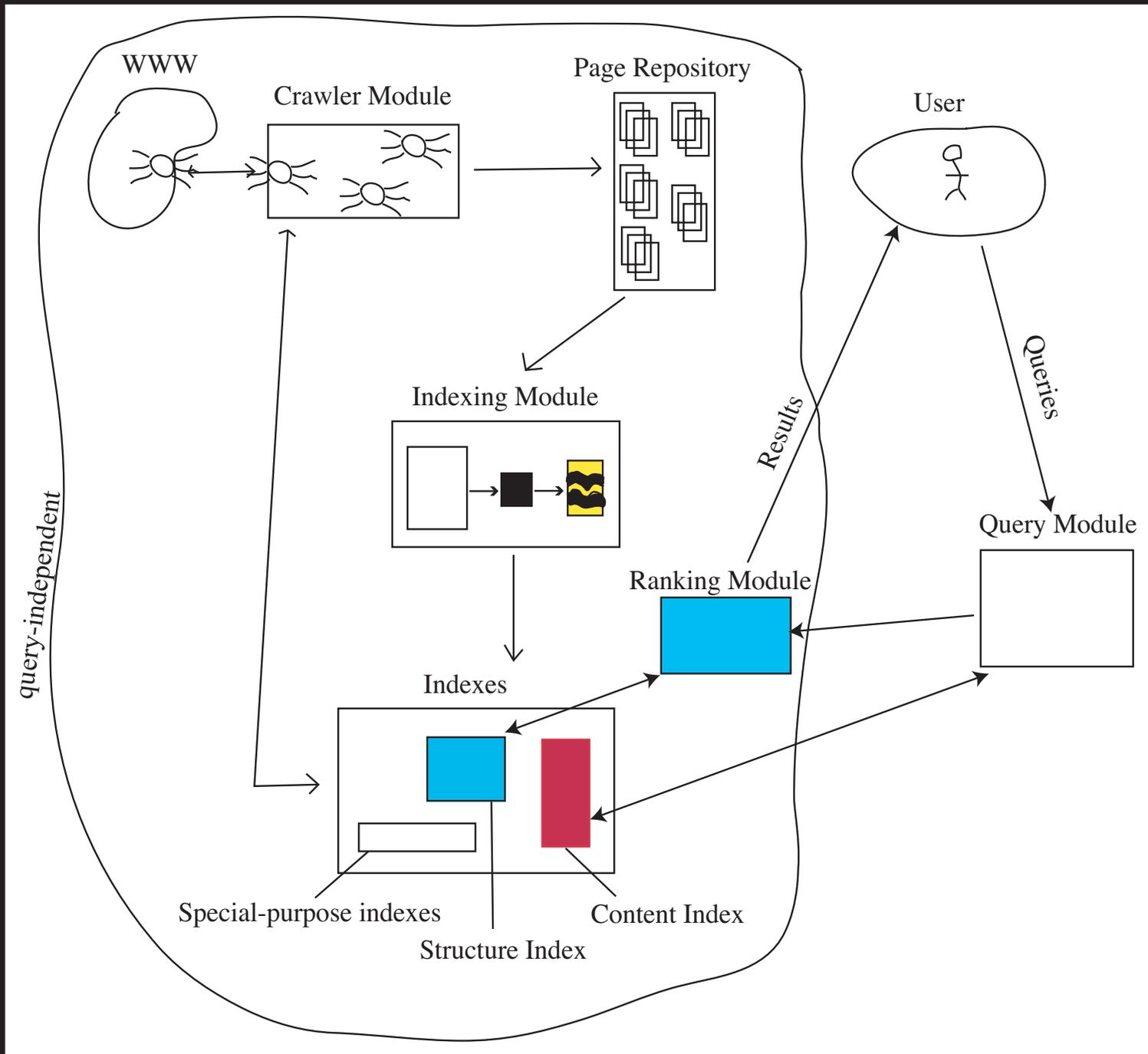
Optional: To help us distinguish between sites submitted by individuals and those automatically entered by software robots, please type the squiggly letters shown here into the box below.



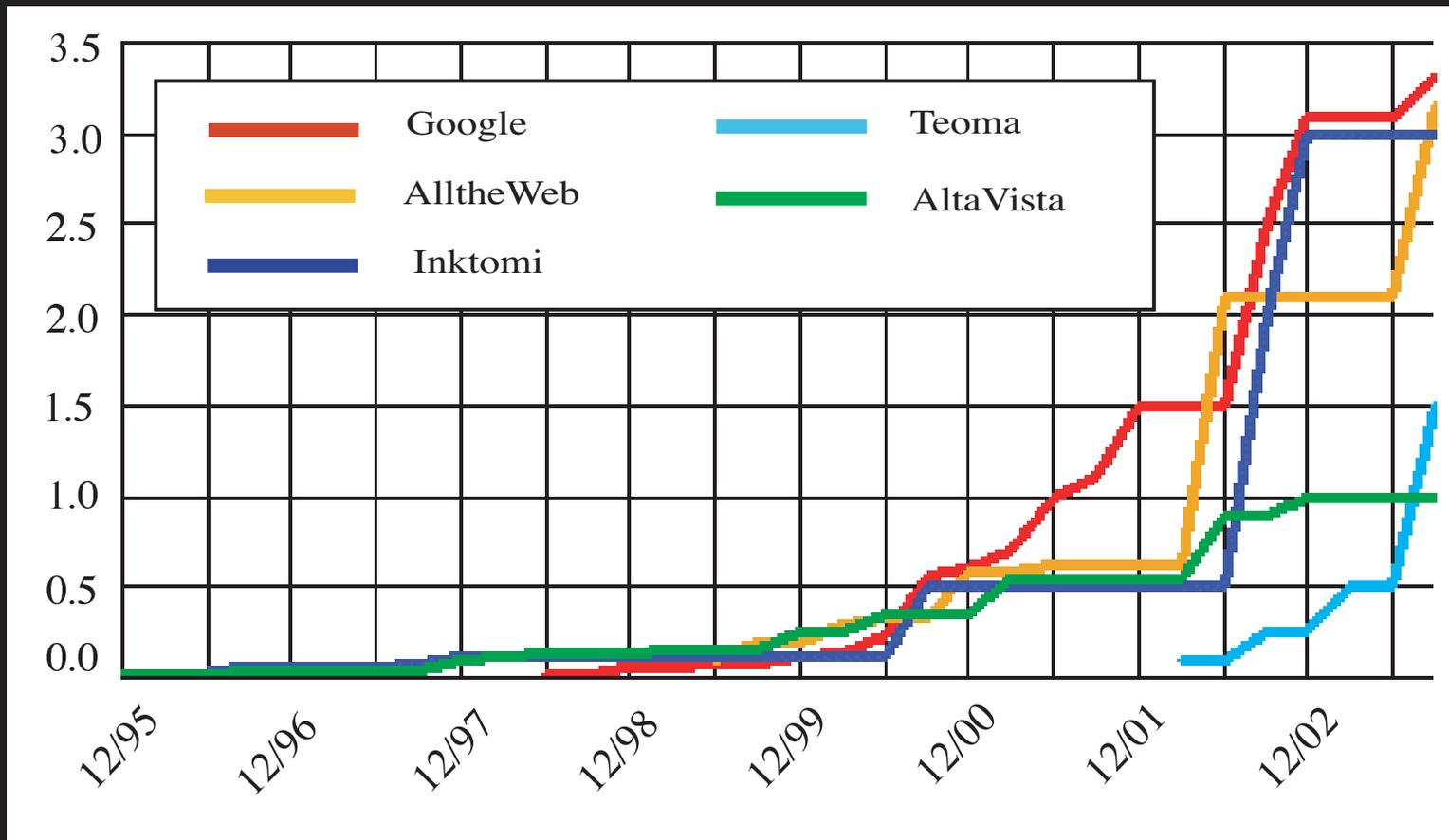
Need to remove a site from Google? For more information, [click here](#).



Elements of a Web Search Engine



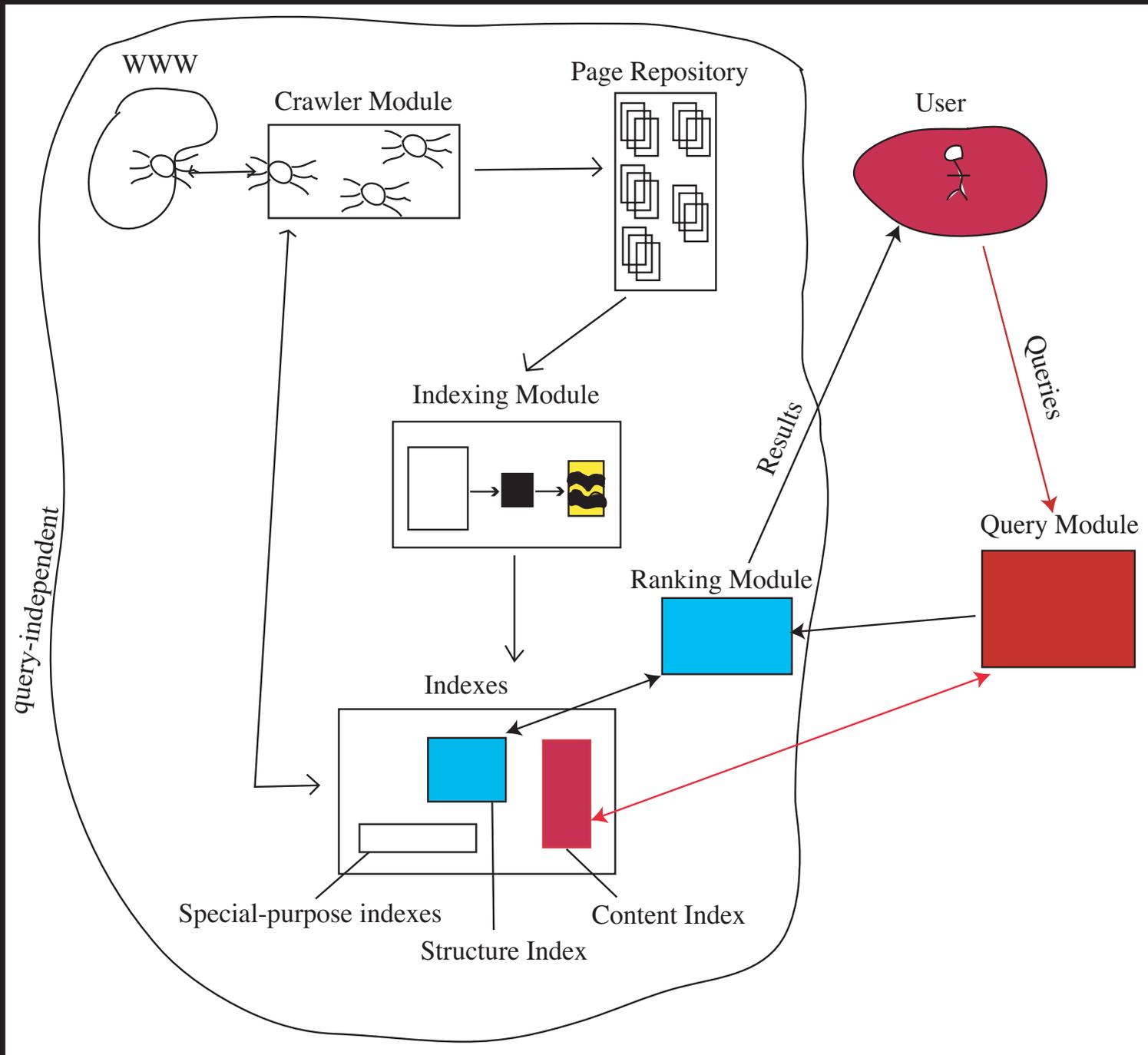
Indexing Wars



Actual Index King =

Internet Archive - <http://web.archive.org>

Elements of a Web Search Engine



Ranking on the Web

the pre-1998 Web

⋮

- border patrol: 4; 567; 809; 1103;

⋮

- hezbollah: 9; 12; 339; 942; 15158;

⋮

- global warming: 178; 12980; 445532;

Index

- k -step transition matrix, 179
- a** vector, 37, 38, 75, 80
- A9, 142
- absolute error, 104
- absorbing Markov chains, 185
- absorbing states, 185
- accuracy, 79–80
- adaptive PageRank method, 89–90
- Adar, Eytan, 146
- adjacency list, 77
- adjacency matrix, 33, 76, 116, 132, 169
- advertising, 45
- aggregated chain, 197
- aggregated chains, 195
- aggregated transition matrix, 105
- aggregated transition probability, 197
- aggregation, 94–97
 - approximate, 102–104
 - exact, 104–105
 - exact vs. approximate, 105–107
 - iterative, 107–109
 - partition, 109–112
- aggregation in Markov chains, 197
- aggregation theorem, 105
- Aitken extrapolation, 91
- Alexa traffic ranking, 138
- algebraic multiplicity, 157
- algorithm
 - PageRank, 40
 - Aitken extrapolation, 92
 - dangling node PageRank, 82, 83
 - HITS, 116
 - iterative aggregation updating, 108
 - personalized PageRank power method, 49
 - quadratic extrapolation, 93
 - query-independent HITS, 124
- α parameter, 37, 38, 41, 47–48
- Amazon’s traffic rank, 142
- anchor text, 48, 54, 201
- Ando, Albert, 110
- aperiodic, 36, 133
- aperiodic Markov chain, 176
- Application Programming Interface (API), 65, 73, 97
- approximate aggregation, 102–104
- arc, 201
- Arrow, Kenneth, 136
- asymptotic convergence rate, 165
- asymptotic rate of convergence, 41, 47, 101, 119, 125
- Atlas of Cyberspace*, 27
- authority, 29, 201
- authority Markov chain, 132
- authority matrix, 117, 201
- authority score, 115, 201
- authority vector, 201
- Babbage, Charles, 75
- back button, 84–86
- BadRank, 141
- Barabasi, Albert-Laszlo, 30
- Berry, Michael, 7
- bibliometrics, 32, 123
- bipartite undirected graph, 131
- BlockRank, 94–97, 102
- blog, 55, 144–146, 201
- Boldi, Paolo, 79
- Boolean model, 5–6, 201
- bounce back, 84–86
- bowtie structure, 134
- Brezinski, Claude, 92
- Brin, Sergey, 25, 205
- Browne, Murray, 7
- Bush, Vannevar, 3, 10
- Campbell, Lord John, 23
- canonical form, reducible matrix, 182
- censored chain, 104
- censored chains, 194
- censored distribution, 104, 195
- censored Markov chain, 194
- ensorship, 146–147
- Cesàro sequence, 162
- Cesàro summability, stochastic matrix, 182
- characteristic polynomial, 120, 156
- Chebyshev extrapolation, 92
- Chien, Steve, 102
- cloaking, 44
- clustering search results, 142–143
- co-citation, 123, 201
- co-reference, 123, 201
- Collatz–Wielandt formula, 168, 172
- complex networks, 30
- compressed matrix storage, 76
- condition number, 59, 71, 155
- Condorcet, 136
- connected components, 127, 133

Ranking on the Web

the pre-1998 Web

⋮

- border patrol: 4; 567; 809; 1103; . . . (8,700,000 in total)

⋮

- hezbollah: 9; 12; 339; 942; 15158; . . . (15,100,000 in total)

⋮

- global warming: 178; 12980; 445532; . . . (33,200,000 in total)

Ranking on the Web

the pre-1998 Web

⋮

- border patrol: 4; 567; 809; 1103; . . . (8,700,000 in total)

⋮

- hezbollah: 9; 12; 339; 942; 15158; . . . (15,100,000 in total)

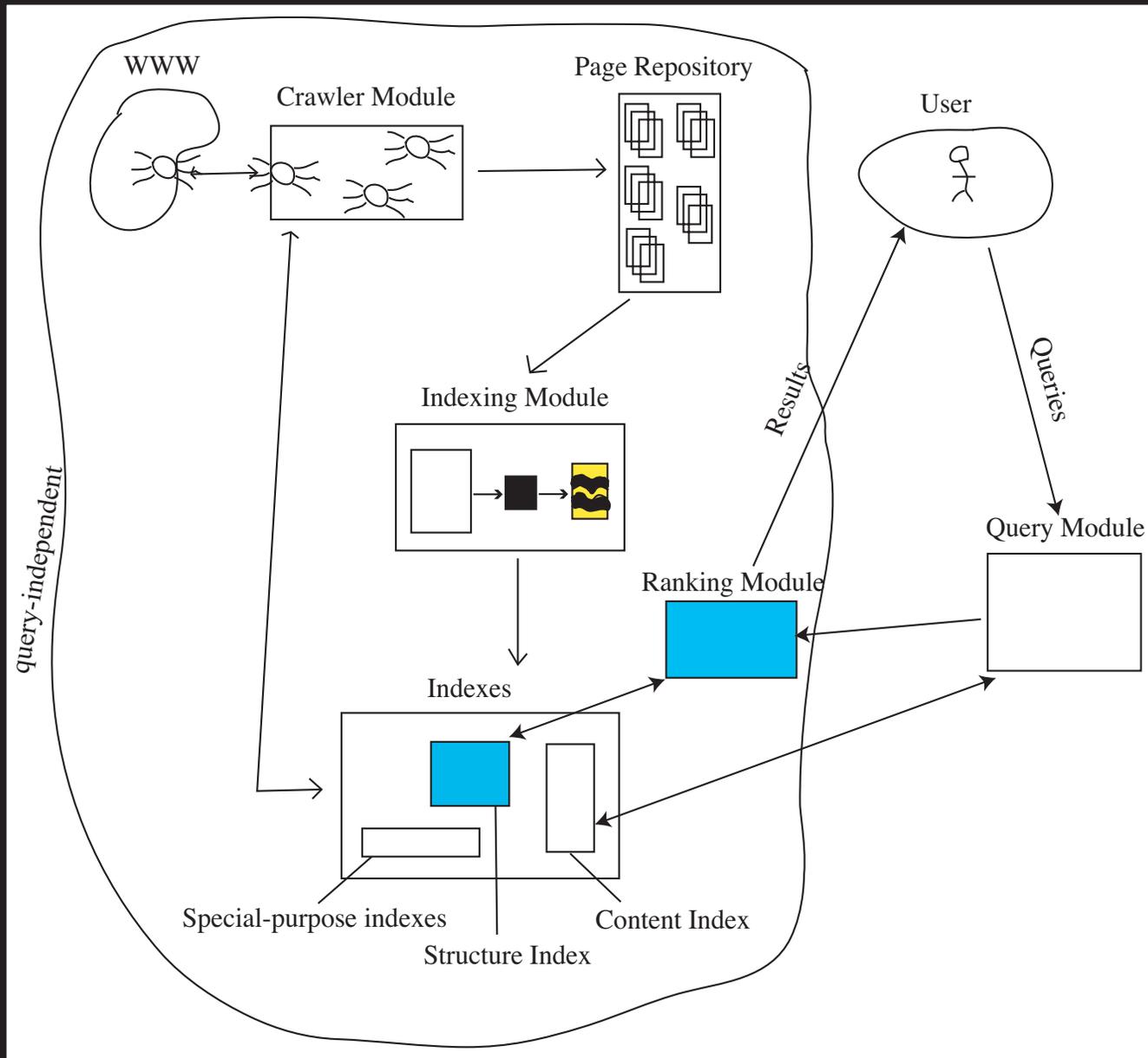
⋮

- global warming: 178; 12980; 445532; . . . (33,200,000 in total)

too many results per search term
easily spammed

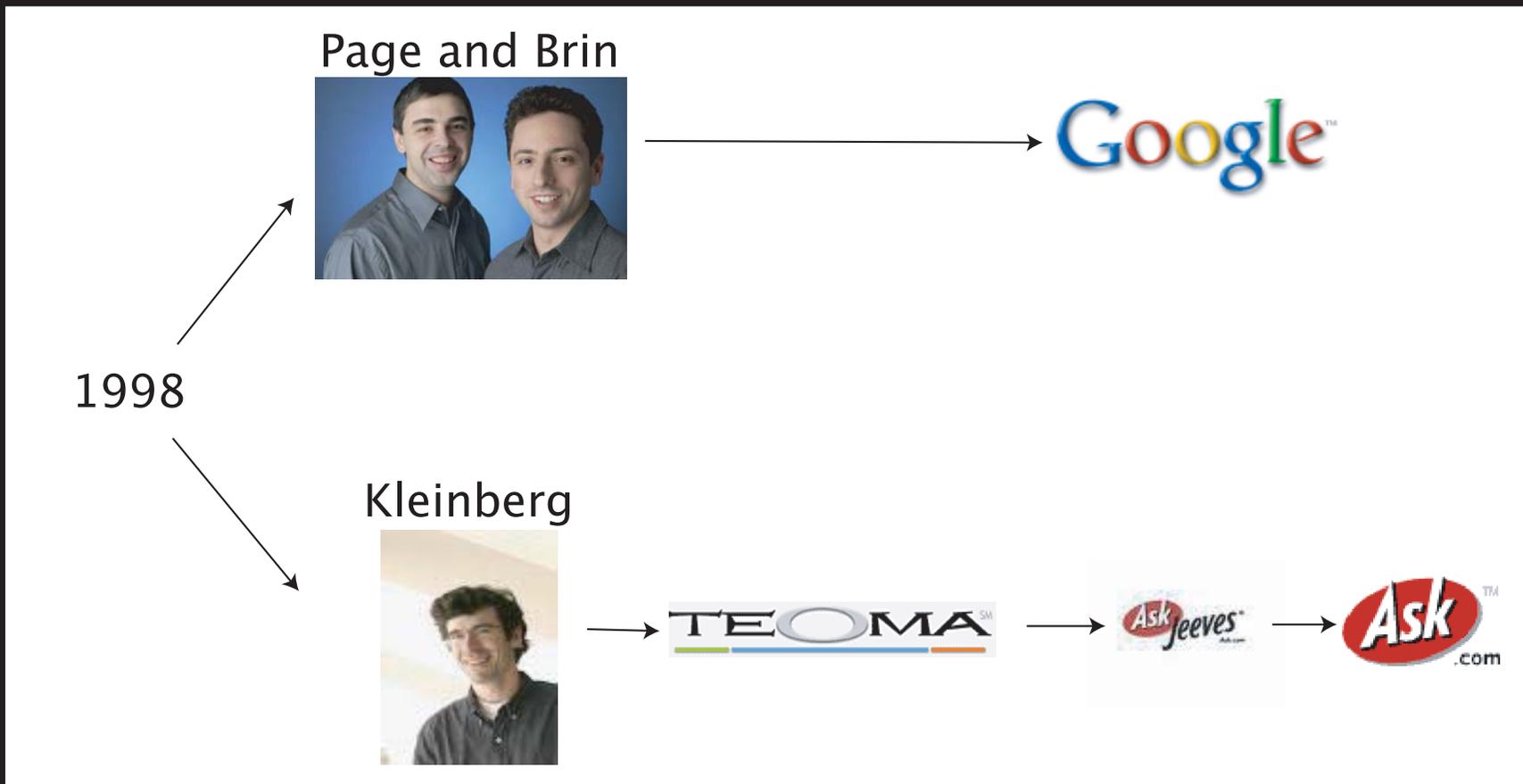
1998: enter Link Analysis

- uses hyperlink structure to focus the relevant set



1998: enter Link Analysis

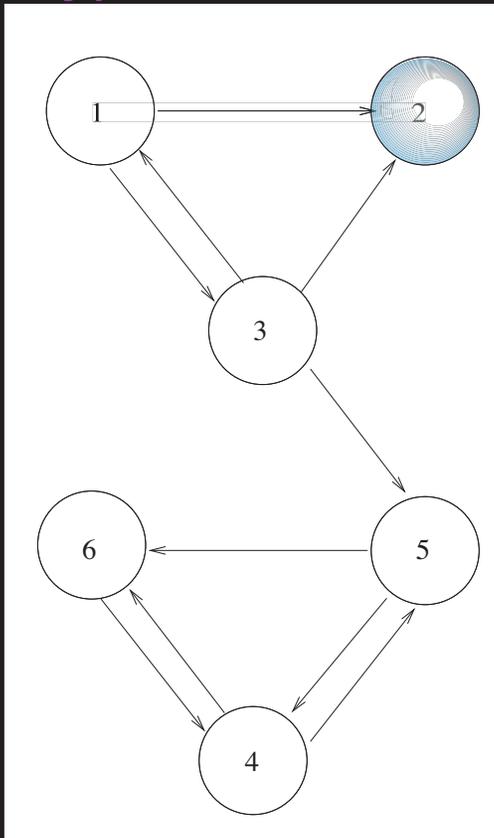
- uses hyperlink structure to focus the relevant set
- combine IR score with popularity score



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

Hyperlink as vote

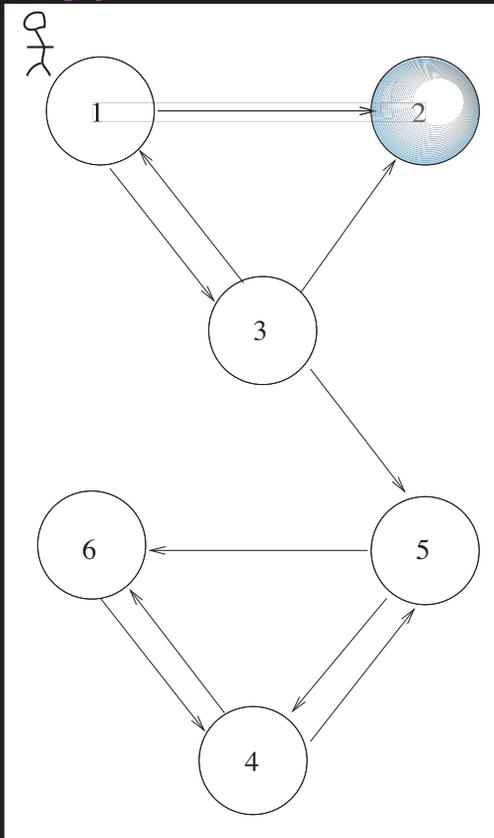


Markov chain

Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

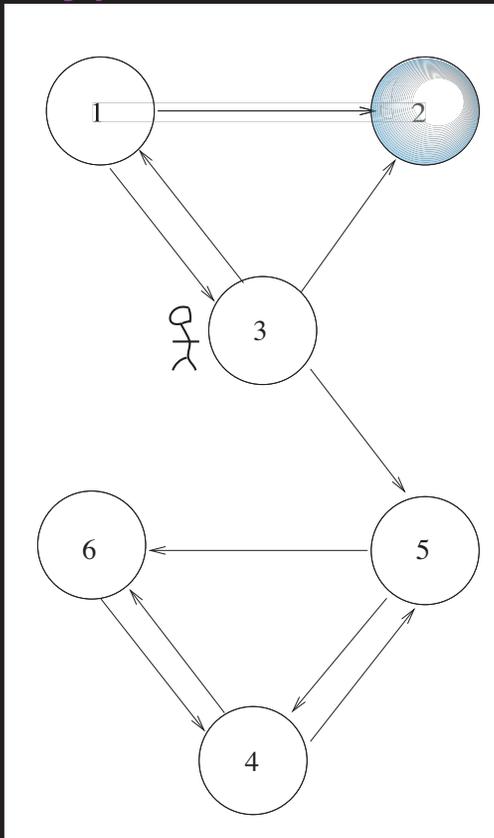
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

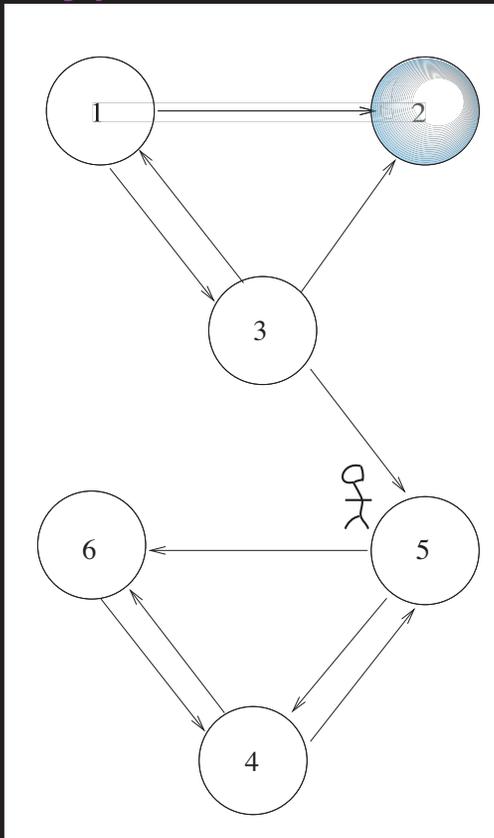
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

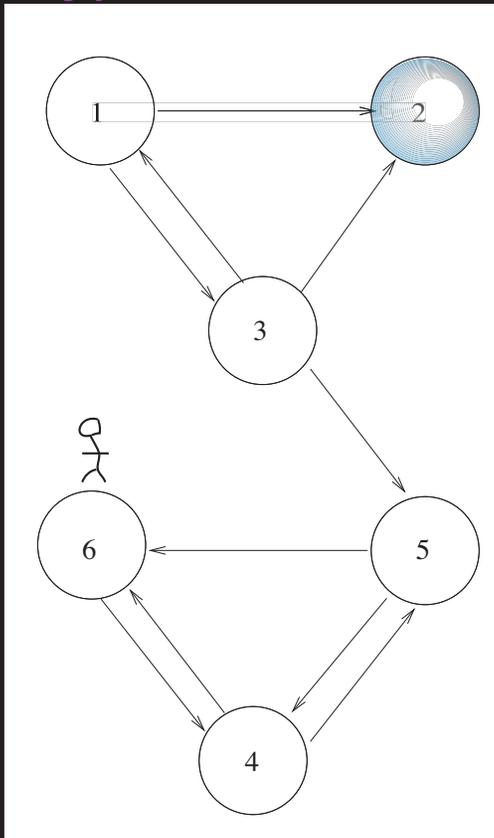
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

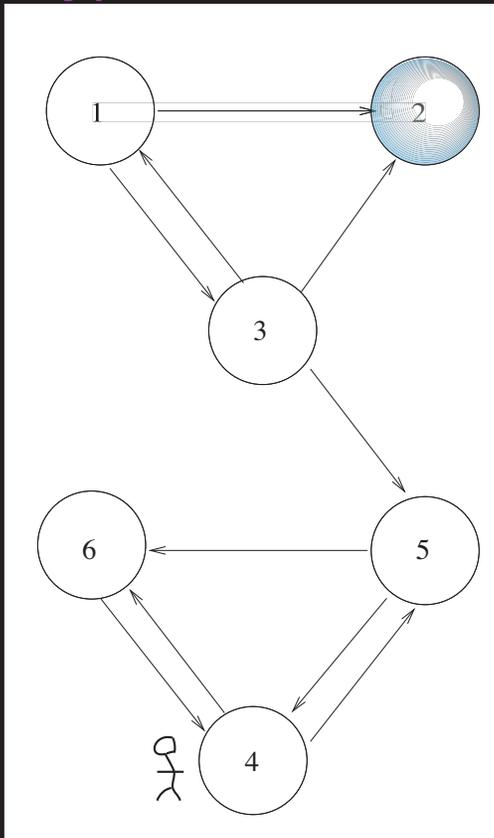
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

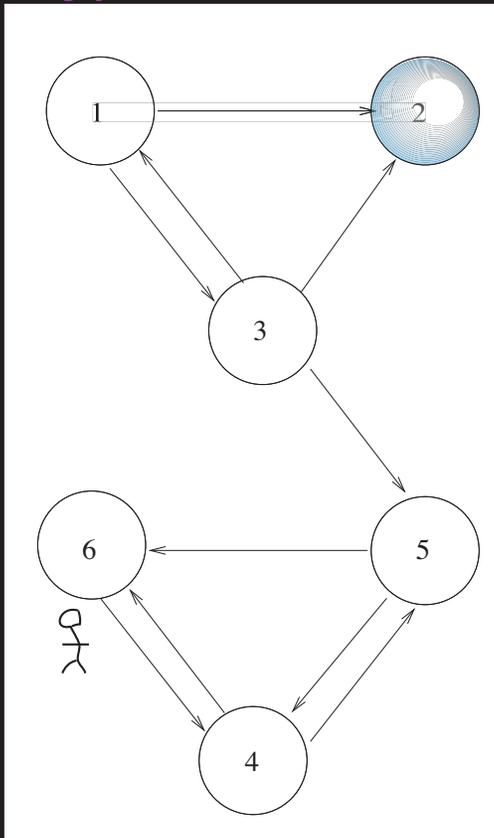
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

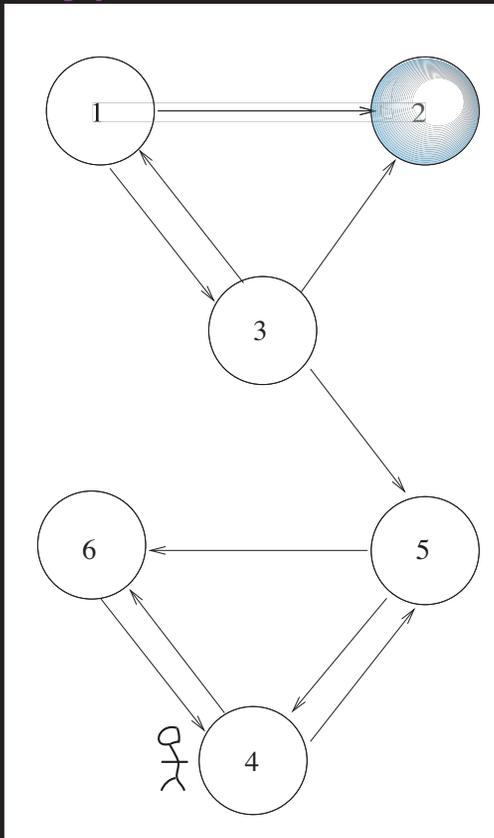
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

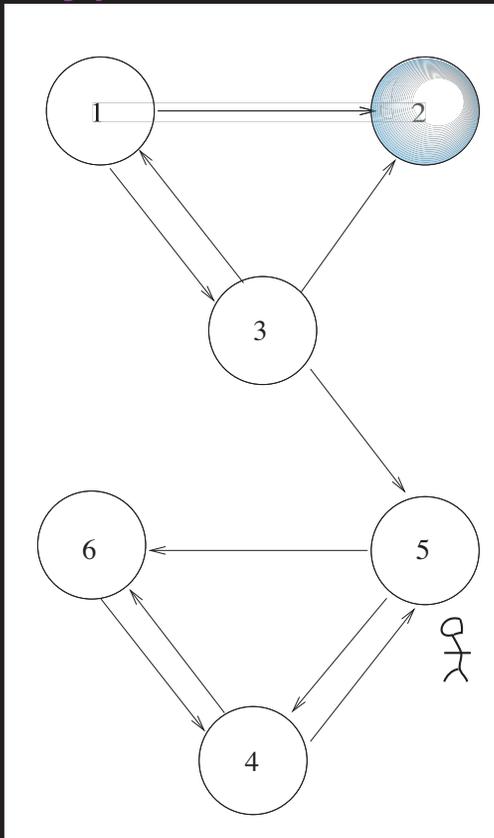
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

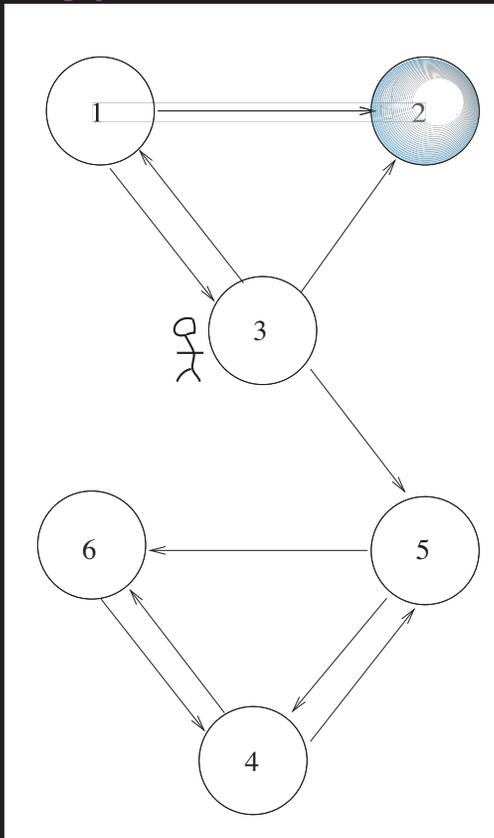
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

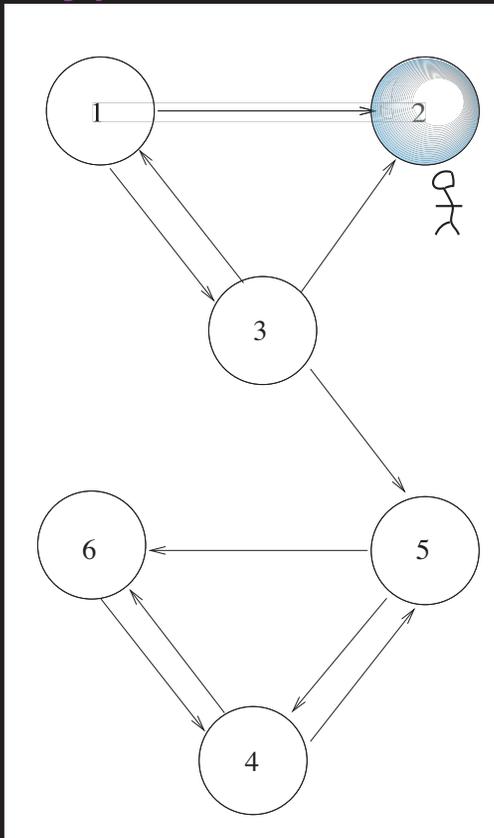
Hyperlink as vote



Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

Hyperlink as vote

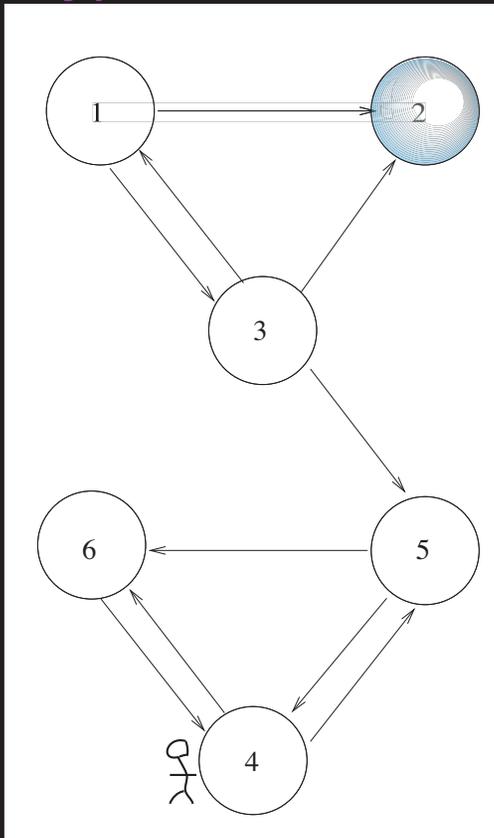


page 2 is a dangling node

Ranking with a Random Surfer

- Rank each page corresponding to a search term by number and *quality* of votes cast for that page.

Hyperlink as vote



surfer “teleports”

Ranking with a Random Surfer

- If a page is “important,” it gets lots of votes from other important pages, which means the random surfer visits it often.
- Simply count the number of times, or *proportion of time*, the surfer spends on each page to create ranking of webpages.

Ranking with a Random Surfer

- If a page is “important,” it gets lots of votes from other important pages, which means the random surfer visits it often.
- Simply count the number of times, or *proportion of time*, the surfer spends on each page to create ranking of webpages.

Proportion of Time

Page 1 = .04

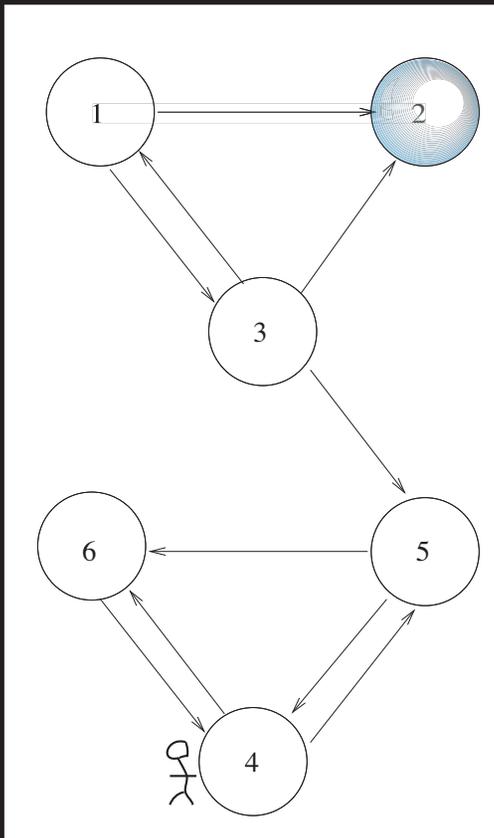
Page 2 = .05

Page 3 = .04

Page 4 = .38

Page 5 = .20

Page 6 = .29



Ranked List of Pages

Page 4

Page 6

Page 5

Page 2

Page 1

Page 3

Ranking with a Random Surfer

- If a page is “important,” it gets lots of votes from other important pages, which means the random surfer visits it often.
- Simply count the number of times, or *proportion of time*, the surfer spends on each page to create ranking of webpages.

Proportion of Time

Page 1 = .04

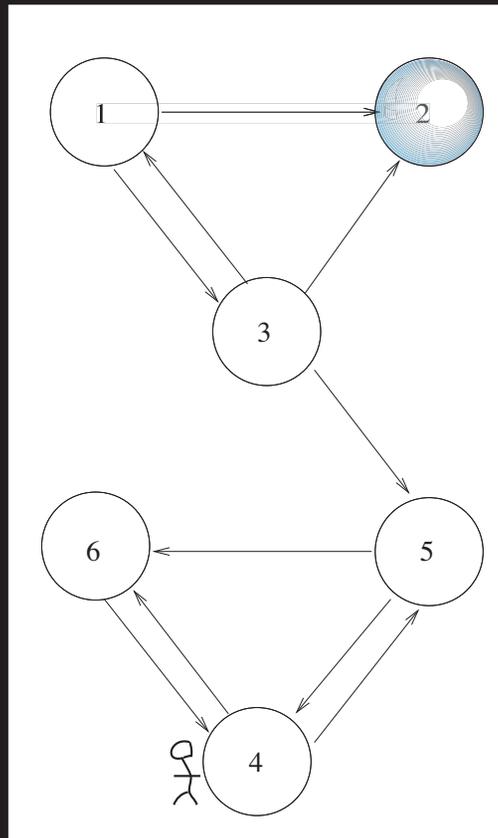
Page 2 = .05

Page 3 = .04

Page 4 = .38

Page 5 = .20

Page 6 = .29



Ranked List of Pages

Page 4

Page 6

Page 5

Page 2

Page 1

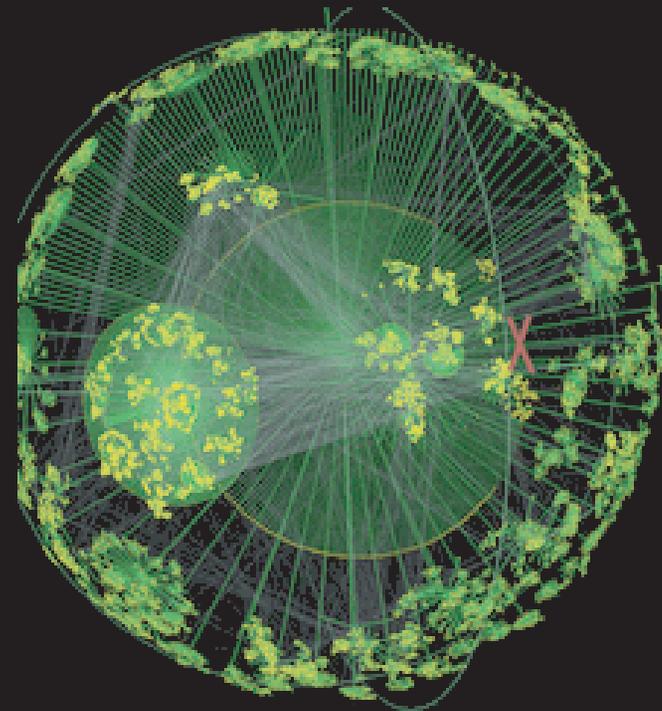
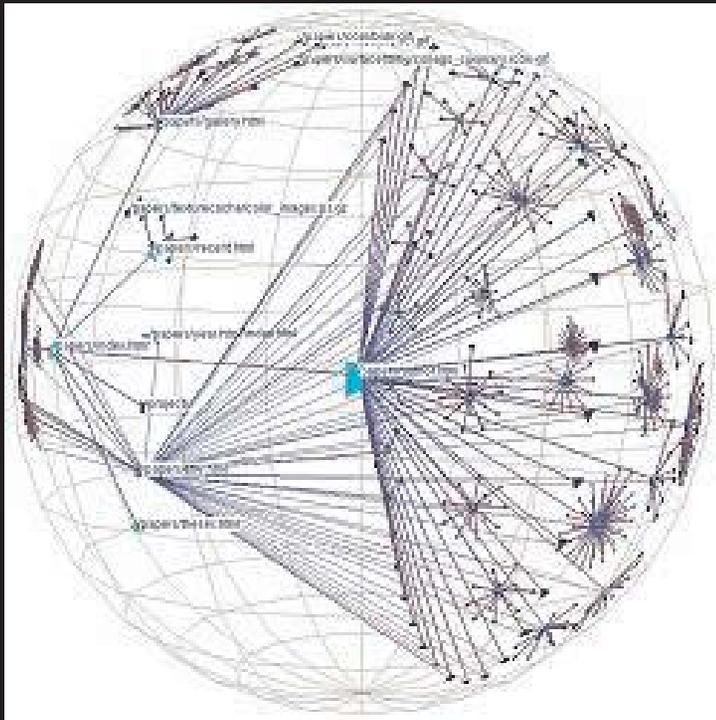
Page 3

query-independent

Web Graphs

CSC and MATH problems here:

- store adjacency matrix
- update adjacency matrix
- visualize web graph
- locate clusters in graph





Our Search: Google Technology

[Home](#)

[All About Google](#)

[Help Central](#)

[Google Features](#)

[Services & Tools](#)

Our Technology

▶ [Why Use Google](#)
[Benefits of Google](#)

Find on this site:

Google searches more sites more quickly, delivering the most relevant results.

Introduction

Google runs on a unique combination of advanced hardware and software. The speed you experience can be attributed in part to the efficiency of our search algorithm and partly to the thousands of low cost PC's we've networked together to create a superfast search engine.

The heart of our software is PageRank™, a system for ranking web pages developed by our founders [Larry Page](#) and [Sergey Brin](#) at Stanford University. And while we have dozens of engineers working to improve every aspect of Google on a daily basis, PageRank continues to provide the basis for all of our web search tools.

PageRank Explained

PageRank relies on the uniquely democratic nature of the web by using its

Google's PageRank is an eigenvector of a matrix of order 2.7 billion.

One of the reasons why Google is such an effective search engine is the PageRank™ algorithm, developed by Google's founders, Larry Page and Sergey Brin, when they were graduate students at Stanford University. PageRank is determined entirely by the link structure of the Web. It is recomputed about once a month and does not involve any of the actual content of Web pages or of any individual query. Then, for any particular query, Google finds the pages on the Web that match that query and lists those pages in the order of their PageRank.

Imagine surfing the Web, going from page to page by randomly choosing an outgoing link from one page to get to the next. This can lead to dead ends at pages with no outgoing links, or cycles around cliques of interconnected pages. So, a certain fraction of the time, simply choose a random page from anywhere on the Web. This theoretical random walk of the Web is a *Markov chain* or *Markov process*. The limiting probability that a dedicated random surfer visits any particular page is its PageRank. A page has high rank if it has links to and from other pages with high rank.

Let W be the set of Web pages that can be reached by following a chain of hyperlinks starting from a page at Google and let n be the number of pages in W . The set W actually varies with time, but in May 2002, n was about 2.7 billion. Let G be the n -by- n connectivity matrix of

BY CLEVE MOLER

It tells us that the largest eigenvalue of A is equal to one and that the corresponding eigenvector, which satisfies the equation

$$x = Ax,$$

exists and is unique to within a scaling factor. When this scaling factor is chosen so that

$$\sum_i x_i = 1$$

then x is the state vector of the Markov chain. The elements of x are Google's PageRank.

If the matrix were small enough to fit in MATLAB, one way to compute the eigenvector x would be to start with a good approximate solution, such as the PageRanks from the previous month, and simply repeat the assignment statement

$$x = Ax$$

until successive vectors agree to within specified tolerance. This is known as the power method and is about the only possible approach for very large n . I'm not sure how Google actually computes PageRank, but one step of the power method would require one pass over a database of Web pages, updating weighted reference counts generated by the hyperlinks between pages.

Search Issues

Spamming

- Link Farms

THE WALL STREET JOURNAL.

© 2003 Dow Jones & Company. All Rights Reserved

WEDNESDAY, FEBRUARY 26, 2003 - VOL. CCXLI NO. 39 - ★★★ \$1.00

WSJ.com

What's News—

Business and Finance

NEWSPAPER CORP. and Liberty are no longer working together on a joint offer to take control of Hughes, with News Corp. proceeding on its own and Liberty considering an independent bid. The move threatens to cloud the process of finding a new owner for the GM unit.

(Article on Page A3)

The SEC signaled it may file civil charges against Morgan Stanley, alleging it doled out IPO shares based partly on investors' commitments to buy more stock.

(Article on Page C1)

Ahold's problems deepened as U.S. authorities opened inquiries into accounting at the Dutch company's U.S. Foodservice unit.

Fleming said the SEC upgraded to a formal investigation an inquiry into the food wholesaler's trade practices with suppliers.

(Articles on Page A2)

Consumer confidence fell to its lowest level since 1993, hurt by energy costs, the terrorism threat and a stagnant job market.

(Article on Page A3)

The industrials rebounded on rumors of a peaceful solution to

World-Wide

BUSH IS PREPARING to present Congress a huge bill for Iraq costs.

The total could run to \$95 billion depending on the length of the possible war and occupation. As horse-trading began at the U.N. to win support for a war resolution, the president again made clear he intends to act with or without the world body's imprimatur. Arms inspectors said Baghdad provided new data, including a report of a possible biological bomb. Gen. Franks assumed command of the war-operations center in Qatar. Allied warplanes are aggressively taking out missile sites that could threaten the allied troop buildup. (Column 4 and Pages A4 and A6)

Turkey's parliament debated legislation to let the U.S. deploy 62,000 to open a northern front. Kurdish soldiers lined roads in a show of force as U.S. officials traveled into Iraq's north for an opposition conference.

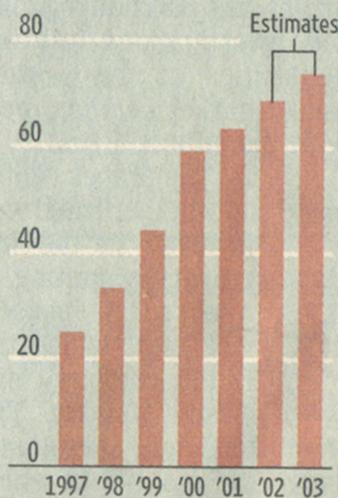
Powell said North Korea hasn't restarted a reactor and plutonium-processing facility at Yongbyon, hinting such forbearance might constitute an overture. But saber rattling continued a day after a missile test timed for the inauguration in Seoul. Pyongyang accused U.S. spy planes of violating its airspace and told its army to prepare for U.S. attack. (Page A14)

The FBI came under withering bipartisan criticism in a Senate Judiciary report in which Sen. Specter

Web Master

As the Web spreads...

Total Internet users, by household, in millions



Sources: Forrester Research; Nielsen NetRatings

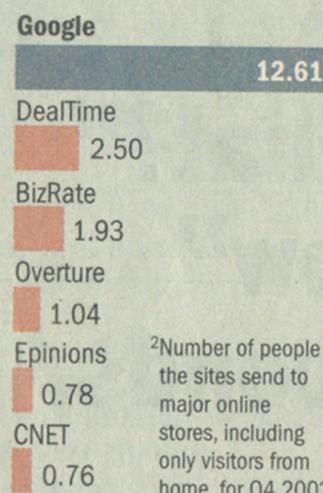
Google's U.S. presence expands

Top search engines, in millions of unique visitors¹



¹Including visitors from home and work, in January 2003

Top shopping-referral sites, in millions of referrals²



²Number of people the sites send to major online stores, including only visitors from home, for Q4 2002

Bush to Seek up to \$95 Billion To Cover Costs of War on Iraq

By GREG JAFFE
And JOHN D. MCKINNON

WASHINGTON—The Bush administration is preparing supplemental spending requests totaling as much as \$95 billion for a war with Iraq, its aftermath and new expenses to fight terrorism, officials said.

The total could be as low as \$60 billion because Pentagon budget planners don't know how long a military conflict will last, whether U.S. allies will contribute more than token sums to the effort and what damage Saddam Hussein might do

to his own country to retaliate against conquering forces.

Budget planners also are awaiting the outcome of an intense internal debate over whether to include \$13 billion in the requests to Congress that the Pentagon says it needs to fund the broader war on terrorism, as well as for stepped up homeland security. The White House Office of Management and Budget argues that the money might not be necessary. President Bush, Defense Secretary Donald Rumsfeld and budget director Mitchell Daniels Jr. met yesterday to discuss the matter but didn't reach a final agreement. Mr. Rumsfeld plans to continue pressing his

Cat and Mouse

As Google Becomes Web's Gatekeeper, Sites Fight to Get In

Search Engine Punishes Firms That Try to Game System; Outlawing the 'Link Farms'

Exoticleatherwear Gets Cut Off

By MICHAEL TOTTY
And MYLENE MANGALINDAN

Joy Holman sells provocative leather clothing on the Web. She wants what nearly everyone doing business online wants: more exposure on Google.

So from the time she launched exoticleatherwear.com last May, she tried all sorts of tricks to get her site to show up among the first listings when a user of Google Inc.'s popular search engine typed in "women's leatherwear" or "leather apparel." She buried hidden words in her Web pages intended to fool Google's computers. She signed up with a service that promised to have hundreds of sites link to her online store—thereby boosting a crucial measure in Google's system of ranking sites.

The techniques worked for a



Web Sites Fight for Prime Real Estate on Google

Continued From First Page
advertising that tried to capitalize on Google's formula for ranking sites. In effect, SearchKing was offering its clients a chance to boost their own Google rankings by buying ads on more-popular sites. SearchKing filed suit against the search company in federal court in Oklahoma, claiming that Google "purposefully devalued" SearchKing and its customers, damaging its reputation and hurting its advertising sales.

Google won't comment on the case. In court filings, the company said SearchKing "engaged in behavior that would lower the quality of Google search results" and alter the company's ranking system.

Google, a closely held company founded by Stanford University graduate students Sergey Brin and Larry Page, says Web companies that want to rank high should concentrate on improving their Web pages rather than gaming its system. "When people try to take scoring into their own hands, that turns into a worse experience for users," says Matt Cutts, a Google software engineer.

Coding Trickery

Efforts to outfox the search engines have been around since search engines first became popular in the early 1990s. Early tricks included stuffing thousands of widely used search terms in hidden coding, called "metatags." The coding fools a search engine into identifying a site with popular words and phrases that may not actually appear on the site.

Another gimmick was hiding words or terms against a same-color background. The hidden coding deceived search engines that relied heavily on the number of times a word or phrase appeared in ranking a site. But Google's system, based on links, wasn't fooled.

Mr. Brin, 29, one of Google's two founders and now its president of technology, boasted to a San Francisco search-engine conference in 2000 that Google wasn't worried about having its results clogged with irrelevant results because its search methods couldn't be manipulated.

That didn't stop search optimizers from finding other ways to outfox the system. Attempts to manipulate Google's results even became a sport, called Google-bombing. Bombsters would try to

creating Web sites that were nothing more than collections of links to the clients' site, called "link farms." Since Google ranks a site largely by how many links or "votes" it gets, the link farms could boost a site's popularity.

In a similar technique, called a link exchange, a group of unrelated sites would agree to all link to each other, thereby fooling Google into thinking the sites have a multitude of votes. Many sites also found they could buy links to themselves to boost their rankings.

Ms. Holman, the leatherwear retailer, discovered the consequences of trying to fool Google. The 42-year-old hospital laboratory technician, who learned computer skills by troubleshooting her hospital's

'The big search engines determine the laws of how commerce runs,' says Mr. Massa.

equipment, operates her online apparel store as a side business that she hopes can someday replace her day job.

When she launched her Exotic Leather Wear store from her home in Mesa, Ariz., she quickly learned the importance of appearing near the top of search-engine results, especially on Google. She boned up on search techniques, visiting online discussion groups dedicated to search engines and reading what material she could find on the Web.

At first, Ms. Holman limited herself to modest changes, such as loading her page with hidden metatag coding that would help steer a search toward her site when a user entered words such as "haltertops" or "leather miniskirts." Since Google doesn't give much weight to metatags in determining its rankings, the efforts had little effect on her search results.

She then received an e-mail advertisement from AutomatedLinks.com, a Wirral, England, company that promised to send traffic "through the roof" by linking more than 2,000 Web sites to hers. Aside from attracting customers, the links were designed to improve her site's search engine rankings by taking

In theory, when Google encounters the AutomatedLinks code, it treats it as a legitimate referral to the other sites and counts them in toting up the sites' popularity.

Shortly after Ms. Holman signed up with AutomatedLinks in July, she read on an online discussion group that Google objected to such link arrangements. She says she immediately stripped the code from her Web pages. For a while her site gradually worked its way up in Google search results, and business steadily improved because links to her site still remained on the sites of other AutomatedLinks customers. Then, sometime in November, her site was suddenly no longer appearing among the top results. Her orders plunged as much as 80%.

Ms. Holman, who e-mailed Google and AutomatedLinks, says she has been unable to get answers. But in the last few months, other AutomatedLinks customers say they have seen their sites apparently penalized by Google. Graham McLeay, who runs a small chauffeur service north of London, saw revenue cut in half during the two months he believes his site was penalized by Google.

The high-stakes fight between Google and the optimizers can leave some Web-site owners confused. "I don't know how people are supposed to judge what is right and wrong," says Mr. McLeay.

AutomatedLinks didn't respond to requests for comment. Google declined to comment on the case. But Mr. Cutts, the Google engineer, warns that the rules are clear and that it's better to follow them rather than try to get a problem fixed after a site has been penalized. "We want to return the most relevant pages we can," Mr. Cutts says. "The best way for a site owner to do that is follow our guidelines."

Crackdown

Google has been stepping up its enforcement since 2001. It warned Webmasters that using trickery could get their sites kicked out of the Google index and it provided a list of forbidden activities, including hiding text and "link schemes," such as the link farms. Google also warned against "cloaking"—showing a search engine a page that's designed to score well while giving visitors a different, more attractive page—or creating multiple Web addresses that take visitors to a single site.

To stay one step ahead of the Web

homa City-based SearchKing, an online directory for hundreds of small, specialty Web sites. SearchKing also sells advertising links designed both to deliver traffic to an advertiser and boost its rankings in Google and other search results.

Bob Massa, SearchKing's chief executive, last August launched the PR Ad Network as a way to capitalize on Google's page-ranking system, known as PageRank. PageRank rates Web sites on a scale of one to 10 based on their popularity, and the rankings can be viewed by Web users if they install special Google software. PR Ad Network sells ads that are priced according to a site's PageRank, with higher-ranked sites commanding higher prices. When a site buys an advertising link on a highly ranked site, the ad buyer could see its ratings improve because of the greater weight Google gives to that link.

Shortly after publicizing the ad network, Mr. Massa discovered that his site suddenly dropped in Google's rankings. What's more, sites that participated in the separate SearchKing directory also had their Google rankings lowered. He filed a lawsuit in Oklahoma City federal court, claiming Google was punishing him for trying to profit from the company's page-ranking system.

A Google spokesman won't comment on the case. In its court filings, Google said it demoted pages on the SearchKing site because of SearchKing's attempts to manipulate search results. The company has asked for the suit to be dismissed, arguing that the PageRank represents its opinion of the value of a Web site and as such is protected by the First Amendment.

"The big search engines determine the laws of how commerce runs," says Mr. Massa, who is persisting with the lawsuit even though the sites have had their page rankings partly restored. "Someone needs to demand accountability."

Google is taking steps that many say could satisfy businesses trying to boost their rankings. Google has long sold sponsored links that show up on the top of many search-results pages, separate from the main listings. Last year, the company expanded its paid-listings program, so that there are now more slots where sites can pay for a prominent place in the results. Many sites now are turning to advertising instead of tactics to optimize their rankings.

Home Depot E Amid First Qu

By CHAD TERHUNE

ATLANTA—Home Depot Inc. reported fiscal fourth-quarter earnings down 3.4% on disappointing sales.

Speaking to investors and industry analysts, the company's chairman and chief executive, Bob Nardelli, said Home Depot is prepared to win back dissatisfied customers and answer a competitive challenge from its chief rival with remodeled stores, increased inventory and improved customer service.

The nation's largest home-improvement retailer said net income for the quarter ended Feb. 2 decreased to \$686 million or 30 cents a share, from \$710 million or 30 cents a share, a year earlier. Sales rose 2% to \$13.21 billion from \$13.49 billion, but first quarterly sales decline in the company's 24-year history. Home Depot net income the latest quarter was a week shorter than a year earlier. Using comparable 13-week periods, the company said quarterly sales increased 5% and net income rose 8%.

Same-store sales, or sales at stores open at least a year, declined 6% in the quarter. Home Depot said stronger sales last month offset a disastrous December and helped the retailer avoid its earnings estimate that same-store sales could fall as much as 10%. In 4 p.m. New York Stock Exchange composite trading, Home Depot shares rose 66 cents to \$22.84.

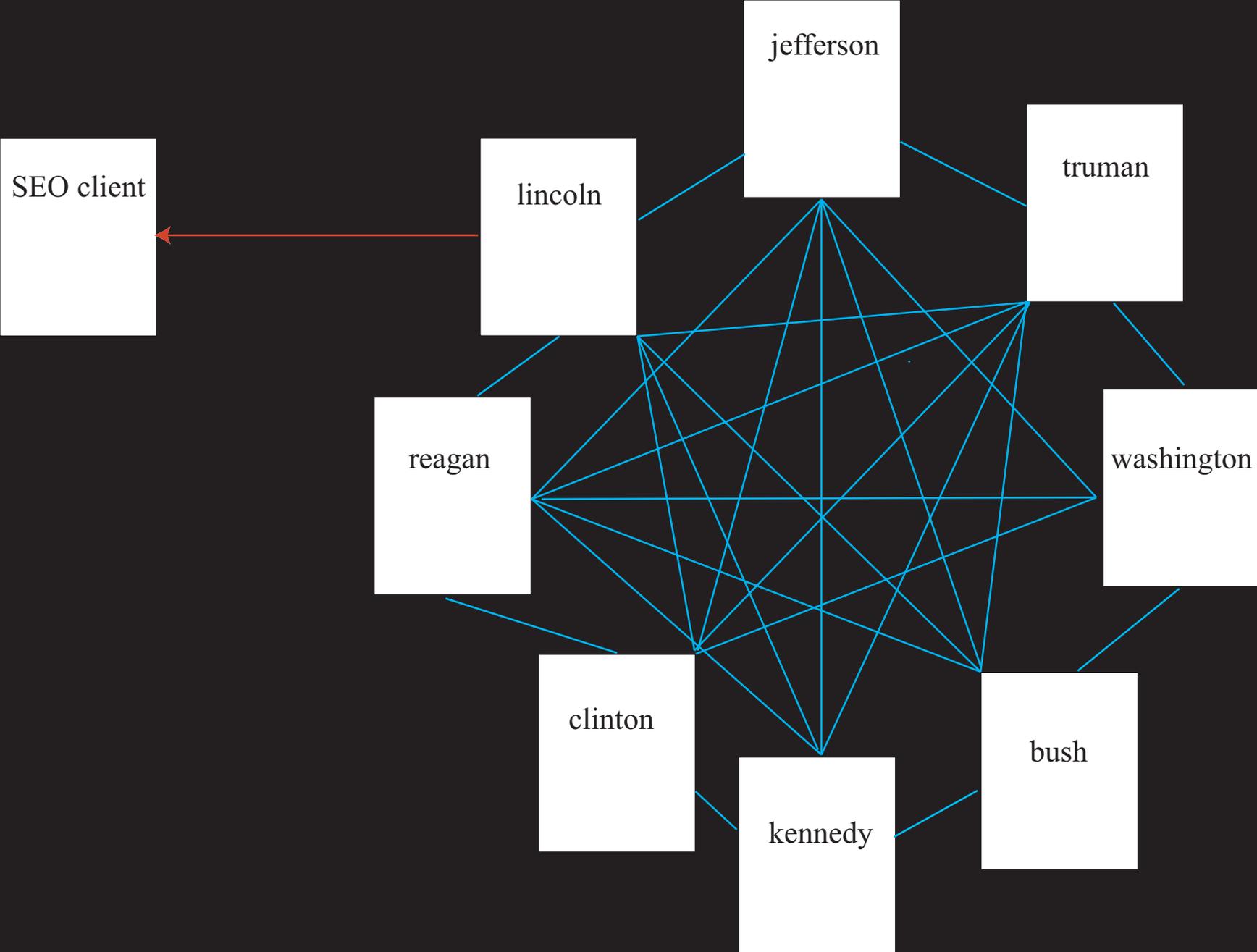
Fiat Patriarch Is Set to Becom

By ALESSANDRA GALLONI

ROME—Umberto Agnelli is due to be named Fiat SpA chairman on Friday, stepping into the driver's seat as the Italian glomeration works on an 11th-hour refinancing of its unprofitable car unit.

Mr. Agnelli, the 68-year-old brother of Fiat patriarch Gianni Agnelli, who last month, was widely expected to be replaced by current chairman, Sergio Marchionne, later this year. But Mr. Agnelli, who has served as chairman since

Link Farms



Search Issues

Spamming

- Link Farms
- Google Bombs



SEARCH

Low Graphics version | [Change edition](#)

[Feedback](#) | [Help](#)



News Front Page



- [Africa](#)
- [Americas](#)
- [Asia-Pacific](#)
- [Europe](#)
- [Middle East](#)
- [South Asia](#)
- [UK](#)
- [Business](#)
- [Health](#)
- [Science/Nature](#)
- [Technology](#)
- [Entertainment](#)

[Have Your Say](#)

[Country Profiles](#)
[In Depth](#)

[Programmes](#)

RELATED SITES

- [BBC SPORT](#)
- [BBC WEATHER](#)
- [BBC ON THIS DAY](#)

LANGUAGES

- [ESPAÑOL](#)
- [BRASIL](#)
- [CARIBBEAN](#)

Last Updated: Sunday, 7 December, 2003, 15:04 GMT

[E-mail this to a friend](#) [Printable version](#)

'Miserable failure' links to Bush

George W Bush has been Google bombed.

Web users entering the words "miserable failure" into the popular search engine are directed to the biography of the president on the White House website.

The trick is possible because Google searches more than just the contents of web pages - it also counts how often a site is linked to, and with what words.

Thus, members of an online community can affect the results of Google searches - called "Google bombing" - by linking their sites to a chosen one.

Weblogger Adam Mathes is credited with inventing the practice in 2001, when he used it to link the phrase "talentless hack" to a friend's website.

The search engine can be manipulated by a fairly small group of users, one report suggested.

Newsday newspaper says as few as 32 web pages with the words "miserable failure" link to the Bush biography.

The Bush administration has been on the receiving end of pointed Google bombs before.

In the run-up to the Iraq war, internet users manipulated Google so the phrase "weapons of mass destruction" led to a joke page saying "These Weapons of Mass Destruction cannot be displayed."

The site suggests "clicking the regime change button", or "If you are George Bush and typed the country's name in the address bar, make sure that it is spelled correctly (IRAQ)".



Bush has been the target of similar pranks before

“ If you are George Bush and typed the country's name in the address bar, make sure that it is spelled correctly (IRAQ) ”

Prank website

SEE ALSO:

- [WMD spoof is internet hit](#)
04 Jul 03 | [West Midlands](#)
- [Google hit by link bombers](#)
13 Mar 02 | [Science/Nature](#)

RELATED INTERNET LINKS:

- [White House](#)
- [Google bombing](#)

The BBC is not responsible for the content of external internet sites

TOP AMERICAS STORIES NOW

- [US army battles to keep soldiers](#)
- [Report backs US Catholic bishops](#)
- [Envoys bid to ease BSE fears](#)
- [Protests widen over sky marshals](#)

[E-mail this to a friend](#) [Printable version](#)

LINKS TO MORE AMERICAS STORIES

Select

[E-mail services](#) | [Desktop ticker](#) | [Mobiles/PDAs](#) |

© BBC MMIV

[Back to top](#) ^^

[News Front Page](#) | [Africa](#) | [Americas](#) | [Asia-Pacific](#) | [Europe](#) | [Middle East](#) | [South Asia](#)
[UK](#) | [Business](#) | [Entertainment](#) | [Science/Nature](#) | [Technology](#) | [Health](#)
[Have Your Say](#) | [Country Profiles](#) | [In Depth](#) | [Programmes](#)

[BBCi Homepage >>](#) | [BBC Sport >>](#) | [BBC Weather >>](#) | [BBC World Service >>](#)

[ABOUT BBC NEWS](#) | [Help](#) | [Feedback](#) | [News sources](#) | [Privacy](#) | [About the BBC](#)

BLAH3.COM

Dusty & Yellowing - [The Blah3 Archives](#)
Complaints, compliments, arguments? [Email me](#)



[<< "Happy Ramadan, y'all..."](#) [\[Main Index\]](#) [\[>> "His heart just isn't in it..."\]](#)

10/27/2003 Archived Entry: "I'm taking part in a new web project..."

I'm taking part in a new web project...

From this day forth, I will refer to George W. Bush as a [Miserable Failure](#) at least once a day. Why, you ask? Well, someone came up with this great idea to link George W. Bush and [Miserable Failure](#) in popular search engines. **If you have a blog or web site, help raise the link between George W. Bush and the phrase 'miserable failure' by copying this link and placing somewhere on your site or blog.**

Thank you very much for your participation.

Replies: 16 people speak up

Great idea!

Posted by [rlr](#) @ 10/27/2003 10:06 PM NY

That is genius. I could add a few other keywords, like "pathetic". I will post it on my blog now...

Posted by [Political Pulpit](#) @ 10/28/2003 02:32 PM NY

Miserable Failure? I'm down with that....

Stay tuned...

Posted by [Drewcifer](#) @ 10/28/2003 02:35 PM NY

Done!

Posted by [Maru](#) @ 10/28/2003 08:46 PM NY

that's great, another thing I think might be good to use: tax cuts for the wealthy....welfare for the wealthy. just my 2 cents.

Posted by [doodaa](#) @ 10/29/2003 03:01 AM NY

Call me a liberal lemming, I guess. :) I'm in.

Posted by [BJ](#) @ 10/29/2003 09:28 AM NY

The key is stating it in connection with terms that will be widely searched. It does no good to simply say "George Bush is a miserable failure" because no one will ever search for that. It might be fun at a parties to show how often the two are in the same sentence in a Google search, but otherwise it does little to advance the theme.

What will work is connecting it to frequent search times, such as "Iraq policy". For instance "George Bush's Iraq Policy is a miserable failure."

The plan shouldn't be to link Miserable Failure to George Bush, but to link Miserable Failure to George Bush and two or three choice, frequently searched phrases.

Overture.com has a keyword suggestion tool that shows how many times certain terms are coming up in searches. Using that tool, I can determine that in September the search for "bush george iraq saddam" gets about 12 times more queries than "george bush iraq speech". "george bush biography" gets a huge amounts of hits compared to something like "george bush policy".

So someone needs to write about three complete sentences using these terms based on verifiable search results and including the "miserable failure" phrase and then advocate for that exact usage.

According to Overture, the phrases "george Bush miserable failure" were not queried even once in their sample during the month just passed.

Posted by [Joe Briefcase](#) @ 10/29/2003 10:51 AM NY

how about drunken, illiterate, mendacious, runt-like miserable failure?

Posted by [tim](#) @ 10/29/2003 11:58 AM NY

Hahaha, that's very productive. This is why everyone knows that liberals are stupid. They do stupid things.

Posted by [Reek Stankleberry](#) @ 10/29/2003 12:04 PM NY

how about, instead of calling it lies--anyone can lie--how about calling it HORSEFEATHERS AND CODSWALLOP! Pin that on him too.

['Den of Thieves'](#)
[The ? Campaign, 2002](#)
['Fair & Balanced' Day](#)



Blahroll

- [A Level Gaze](#)
- [A Skeptical Blog](#)
- [Ain't No Bad Dude](#)
- [Angry Bear](#)
- [Ann Slanders](#)
- [Apathy, Inc.](#)
- [Army of Fun](#)
- [Atrios](#)
- [Attorney At Arms](#)
- [Avedon's Sideshow](#)
- [Bag Times](#)
- [BartCop E!](#)
- [BartCop!](#)
- [Bellum Americanum](#)
- [Big Picnic](#)
- [Bitter Obscurity](#)
- [Booknotes](#)
- [Bunsen](#)
- [Burgblog](#)
- [Bush Is A Moron](#)
- [BushFlash](#)
- [BushLiar](#)
- [BusyBusyBusy](#)
- [Byrd's Brain](#)
- [Certain Shade of Green](#)
- [Chimes at Midnight](#)
- [Chris Nelson](#)
- [Circumspect](#)
- [CNN Lies](#)
- [Conniption](#)
- [Counterspin](#)
- [Cursor](#)
- [Daily Brew](#)
- [Daily Cynic](#)
- [Daily Kos](#)
- [Daily Outrage](#)
- [Daily War News](#)
- [Damfacrats](#)
- [Deckie Holmes](#)
- [Democratic Veteran](#)
- [Dodona](#)
- [dratfink](#)
- [Duckwing](#)
- [E Pluribus Unum](#)
- [Estimated Prophet](#)
- [Ethe!](#)
- [Federal Examiner](#)
- [Fengi](#)
- [For Freedom Century](#)
- [Frog'N'Blog](#)
- [Ge. JC Christian](#)
- [GeekPol](#)
- [Geoan Sailor](#)
- [GeoDog](#)
- [Get Donkey!](#)


[Advanced Search](#) [Preferences](#) [Language Tools](#) [Search Tips](#)

miserable failure

[Web](#) [Images](#) [Groups](#) [Directory](#) [News](#)
Searched the web for **miserable failure**. Results **1 - 10** of about **257,000**. Search took **0.08** seconds.

Tip: In most browsers you can just hit the return key instead of clicking on the search button.

[Michael Moore.com](#)

Wednesday, January 14th, 2004 I'll Be Voting For Wesley Clark /

Good-Bye Mr. Bush — by Michael Moore. Many of you have written ...

Description: Official site of the gadfly of corporations, creator of the film Roger and Me and the television show...

Category: Arts > Celebrities > M > Moore, Michael

www.michaelmoore.com/ - 43k - [Cached](#) - [Similar pages](#)

[Biography of President George W. Bush](#)

Home > President > Biography President George W. Bush En Español.

George W. Bush is the 43rd President of the United States. He ...

Description: Biography of the president from the official White House web site.

Category: Kids and Teens > School Time > ... > Bush, George Walker

www.whitehouse.gov/president/gwbbio.html - 29k - [Cached](#) - [Similar pages](#)

[Biography of Jimmy Carter](#)

Home > History & Tours > Past Presidents > Jimmy Carter. Jimmy Carter.

Jimmy Carter aspired to make Government "competent and compassionate ...

Description: Short biography from the official White House site.

Category: Society > History > ... > Presidents > Carter, James Earl

www.whitehouse.gov/history/presidents/jc39.html - 36k - [Cached](#) - [Similar pages](#)

[Senator Hillary Rodham Clinton: Online Office Welcome Page](#)

Dear Friend,. Thank you for visiting my on-line office! I appreciate

your interest in the issues before the United States Senate. ...

Description: Official US Senate web site of Senator Hillary Rodham Clinton (D - NY).

Category: Society > History > ... > First Ladies > Clinton, Hillary

clinton.senate.gov/ - 9k - [Cached](#) - [Similar pages](#)

[BBC NEWS | Americas | 'Miserable failure' links to Bush](#)

'Miserable failure' links to Bush. ... Prank website. Newsday newspaper says as few as 32 web pages with the words "miserable failure" link to the Bush biography. ...

news.bbc.co.uk/2/hi/americas/3298443.stm - 31k - [Cached](#) - [Similar pages](#)

[Atlantic Unbound | Politics & Prose | 2003.09.24](#)

... Atlantic Unbound | September 24, 2003 Politics & Prose | by Jack Beatty

"A Miserable Failure" Will Bush be re-elected? Only if voters ...

www.theatlantic.com/unbound/polipro/pp2003-09-24.htm - 22k - [Cached](#) - [Similar pages](#)

[miserable failure | Hillary Clinton | Hildebeest](#)

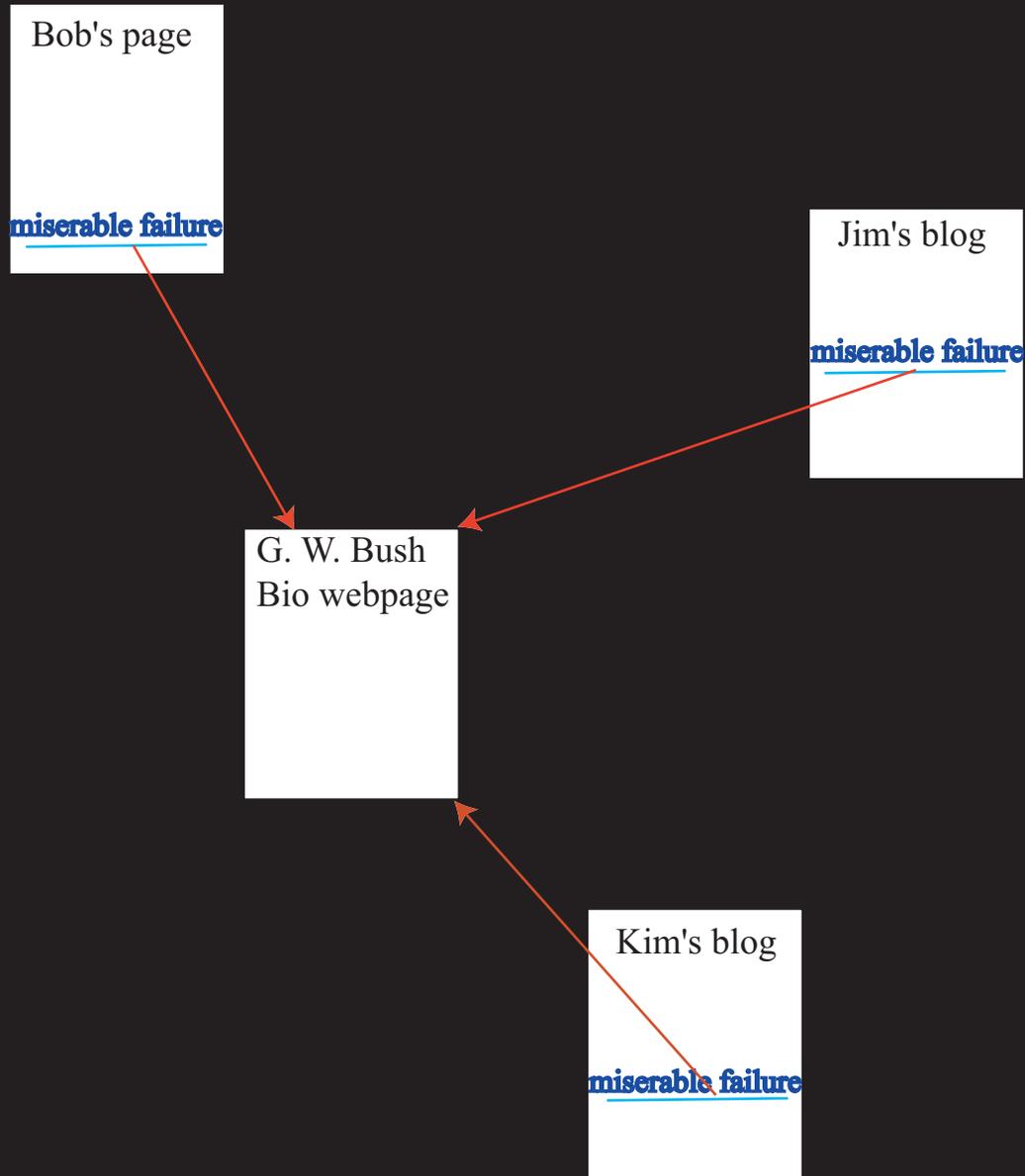
... Miserable Failure. Quotes for the History Books. ... You may also want to check out the Miserable Failure Project. and the cuckolded dyke Project. and the ...

miserable-failure.blogspot.com/ - 60k - [Cached](#) - [Similar pages](#)

[Dick Gephardt for President - Welcome](#)

... to preserve some large part of the Bush tax cut. I think retaining

Google Bomb



Search Issues

Spamming

- Link Farms
- Google Bombs

Personalization

- Google's psearch, A9, Kartoo

The screenshot displays the KartOO visual meta search engine interface. At the top, there is a search bar containing the text "langville" and a "Search" button. To the right of the search bar are buttons for "help", "english pages", "options", and "Products". The main area features a network map with nodes and connections. Nodes include website URLs like "www.cofc.edu", "www.krellinst.org", "www.informatik.uni-trier.de", "www.allbookstores.com", "www.pupress.princeton.edu", "www.langreiter.com", "www.realestate.com.au", "www.nni-news.com", "www.internetmathematics.org", "www.math.ncsu.edu", and "members.lycos.co.uk". Central nodes are labeled with terms like "book", "meyer", "pagerank", "carl", "publications", "abstract information retrieval", "google beyond description science", and "hard". The interface also includes a left sidebar with a "Saved map" section and a "Topics" list, a bottom-left "Preview by THUMBSHOT" logo with flags, and a bottom-right "next map" button and "18 700 Found results 25 - 40" status.

Search Issues

Spamming

- Link Farms
- Google Bombs

Personalization

- Google's psearch, A9, Kartoo

Privacy

- Great Firewall of China
- AOL Data Leak

HOME PAGE MY TIMES TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

TimesSelect Free 14-Day Trial Log In Register Now

The New York Times

Technology

Technology All NYT Search



WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS

CAMCORDERS CAMERAS CELLPHONES COMPUTERS HANDHELDS HOME VIDEO MUSIC PERIPHERALS WI-FI

A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2006

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.



Erik S. Lesser for The New York Times

Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

Multimedia

Graphic: [What Revealing Search Data Reveals](#)

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from “numb fingers” to “60 single men” to “dog that urinates on everything.”

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for “landscapers in Lilburn, Ga,” several people with the last name Arnold and “homes sold in shadow lake subdivision winnett county georgia.”

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. “Those are my searches,” she said, after a reporter read part of the list to her.

AOL removed the search data from its site over the weekend and apologized for its release, saying it was an unauthorized move by a team that had hoped it would benefit academic researchers.

But the detailed records of searches conducted by Ms. Arnold and 657,000 other Americans, copies of which continue to circulate online, underscore how much people unintentionally reveal about themselves when they use search engines — and how risky it can be for companies like AOL, [Google](#) and [Yahoo](#) to compile such data.

Those risks have long pitted privacy advocates against online marketers and other Internet companies seeking to profit from the Internet's unique ability to track the comings and goings of users, allowing for more focused and therefore more lucrative advertising.

But the unintended consequences of all that data being compiled, stored and cross-linked are what Marc Rotenberg, the executive director of the Electronic Privacy Information Center, a privacy rights group in Washington, called “a ticking privacy time bomb.”

Mr. Rotenberg pointed to Google's own joust earlier this year with the Justice

[More Articles in Technology »](#)

Circuits E-Mail



Sign up for David Pogue's exclusive column, sent every Thursday.

Sign Up

[See Sample](#) | [Privacy Policy](#)

SIGN IN TO E-MAIL THIS

PRINT

SINGLE PAGE

REPRINTS

SAVE

ARTICLE TOOLS SPONSORED BY

HISTORY BOYS



WITH THE NEW HP PROLIANT ML150 G3 SERVER
featuring the Dual-Core Intel® Xeon® Processor



Smart buy priced at
\$1,299

» SAVE NOW



Search Issues

Spamming

- Link Farms
- Google Bombs

Personalization

- Google's psearch, A9, Kartoo

Privacy

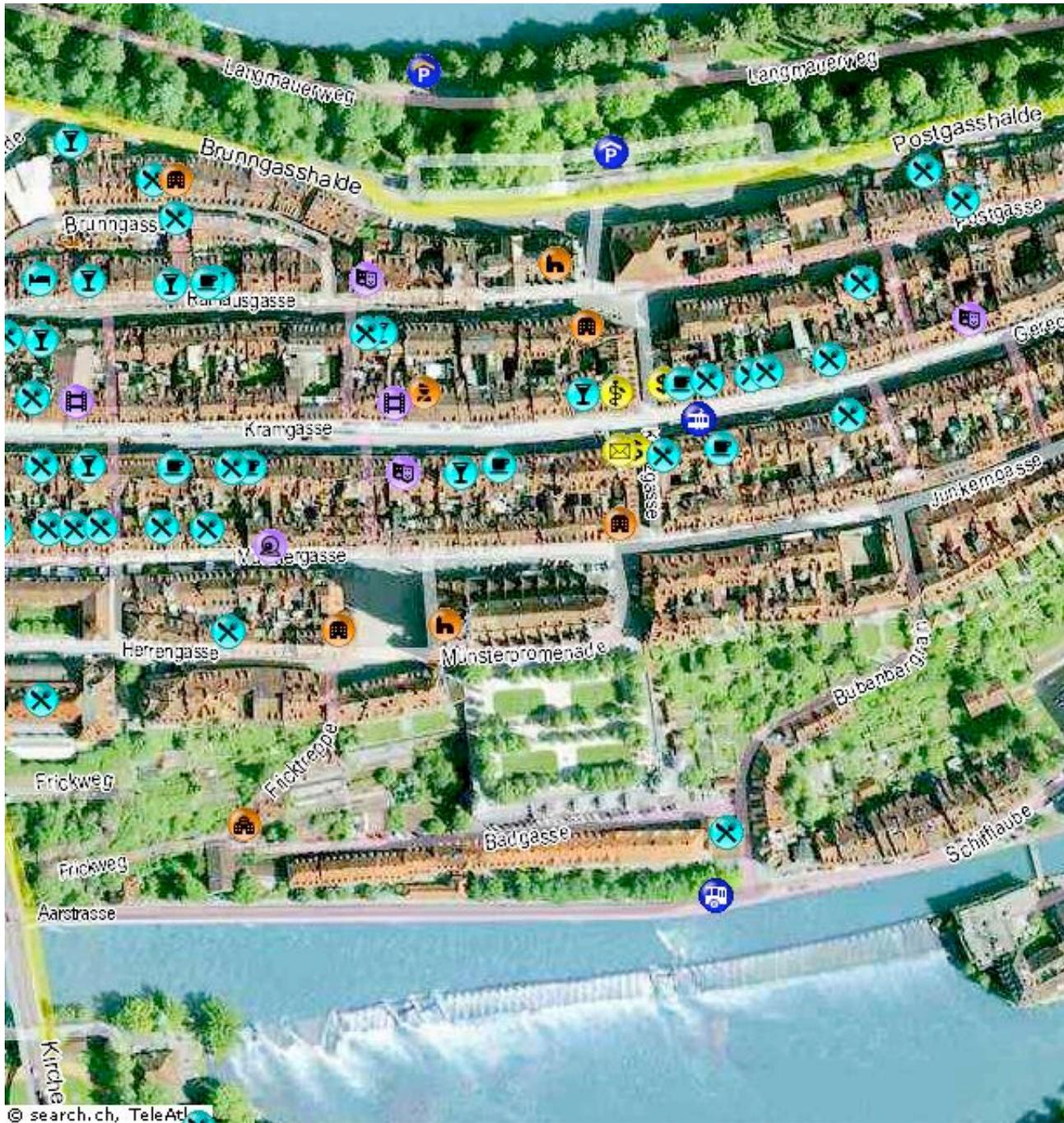
- Great Firewall of China
- AOL Data Leak

Data Fusion

- Search.ch

[map.search.ch]

Map: Bern



© search.ch, TeleAtl

50 m

PageRank and You

Get Indexed

- Submit your site to engines.



Add your URL to Google

[Home](#)

[About Google](#)

[Advertising Programs](#)

[Business Solutions](#)

[Webmaster Info](#)

▶ [Submit Your Site](#)

Find on this site:

Share your place on the net with us.

We add and update new sites to our index each time we crawl the web, and we invite you to submit your URL here. We do not add all submitted URLs to our index, and we cannot make any predictions or guarantees about when or if they will appear.

Please enter your full URL, including the `http://` prefix. For example: `http://www.google.com/`. You may also add comments or keywords that describe the content of your page. These are used only for our information and do not affect how your page is indexed or used by Google.

Please note: Only the top-level page from a host is necessary; you do not need to submit each individual page. Our crawler, Googlebot, will be able to find the rest. Google updates its index on a regular basis, so updated or outdated link submissions are not necessary. Dead links will 'fade out' of our index on our next crawl when we update our entire index.

URL:

Comments:

Optional: To help us distinguish between sites submitted by individuals and those automatically entered by software robots, please type the squiggly letters shown here into the box below.



Need to remove a site from Google? For more information, [click here](#).



PageRank and You

Get Indexed

- Submit your site to engines.

Content

- Use metatags properly.
- Select the right keywords for your market.



It's All About Results™

[Help](#) | [Contact Us](#)

Keyword Tool

The Keyword Tool generates potential [keywords for your ad campaign](#) and reports their Google statistics, including search performance and seasonal trends. Start your search by entering your own keyword phrases or a specific URL. You can then add new keywords to the green box at the right. [Learn more](#)

Important note: Please note that we cannot guarantee that these keywords will improve your campaign performance. We also reserve the right to disapprove any new keywords you add. Keep in mind that you alone are responsible for the keywords you select and for making sure that your use of the keywords does not violate any applicable laws, including any applicable trademark laws. For more details, please review our [Terms and Conditions](#).

Results are tailored to **English, United States** [Edit](#)

Keyword Variations
Site-Related Keywords

Enter one keyword or phrase per line:

pagerank

Use synonyms

Get More Keywords

Choose data to display:

Search Volume Trends

?

More specific keywords - sorted by relevance ?

Keywords	Avg Search Volume	Search Volume Trends (Jan - Dec 2006) ?	Highest Volume Occurred In	Match Type: ?
pagerank	[Bar chart]	[Bar chart]	Feb	Add »
page rank	[Bar chart]	[Bar chart]	Oct	Add »
search engine	[Bar chart]	[Bar chart]	Oct	Add »
link popularity	[Bar chart]	[Bar chart]	Feb	Add »
pop up blocker	[Bar chart]	[Bar chart]	Apr	Add »
search engines	[Bar chart]	[Bar chart]	Mar	Add »
backward links	[Bar chart]	[Bar chart]	Mar	Add »
page rank checker	[Bar chart]	[Bar chart]	Oct	Add »
future page				

Selected Keywords:

Click 'Sign up with these keywords' when you are finished building your keyword list.

No keywords added yet

[+ Add your own keywords](#)

Sign up with these keywords

future page rank		Oct	Add »
check page rank		Oct	Add »
page rank tool		Dec	Add »
my page rank		Oct	Add »
page rank calculator		Oct	Add »
live page rank	No data	Jan	Add »
increase page rank		Dec	Add »
page rank update		Feb	Add »
page rank prediction		Feb	Add »
page rank predictor	No data	Jan	Add »
yahoo page rank		Oct	Add »
powered by		Jun	Add »
web page		Mar	Add »
page rank algorithm	No data	Jan	Add »
page rank toolbar		Dec	Add »
what is my page rank	No data	Jan	Add »
find page rank		Oct	Add »
high page rank		Jun	Add »
web page rank		Jan	Add »
alexa page rank	No data	Jan	Add »
page rank 10	No data	Jan	Add »
improve page rank		Jun	Add »
how to increase page rank		Aug	Add »
page rank lookup		Nov	Add »
what is page rank		Oct	Add »
page rank finder	No data	Jan	Add »

PageRank and You

Get Indexed

- Submit your site to engines.

Content

- Use metatags properly.
- Select the right keywords for your market.

Links

- What is your current PageRank?
- Find who links to you.
- Be wary of link exchange programs.

Advanced Search

[Advanced Search Tips](#) | [About Google](#)

Find results with **all** of the words 10 results

with the **exact phrase** Google Search

with **at least one** of the words

without the words

Language Return pages written in any language

File Format Only return results of the file format any format

Date Return web pages updated in the anytime

Numeric Range Return web pages containing numbers between and

Occurrences Return results where my terms occur anywhere in the page

Domain Only return results from the site or domain
e.g. google.com, .org [More info](#)

Usage Rights Return results that are not filtered by license
[More info](#)

SafeSearch No filtering Filter using [SafeSearch](#)

Page-Specific Search

Similar Find pages similar to the page Search
e.g. www.google.com/help.html

Links Find pages that link to the page Search

Topic-Specific Searches

- [Google Book Search](#) - Search the full text of books
- [New! Google Code Search](#) - Search public source code
- [Google Scholar](#) - Search scholarly papers
- [Google News archive search](#) - Search historical news

- [Apple Macintosh](#) - Search for all things Mac
- [BSD Unix](#) - Search web pages about the BSD operating system
- [Linux](#) - Search all penguin-friendly pages
- [Microsoft](#) - Search Microsoft-related pages

- [U.S. Government](#) - Search all U.S. federal, state and local government sites
- [Universities](#) - Search a specific school's website

PageRank and You

Get Indexed

- Submit your site to engines.

Content

- Use metatags properly.
- Select the right keywords for your market.

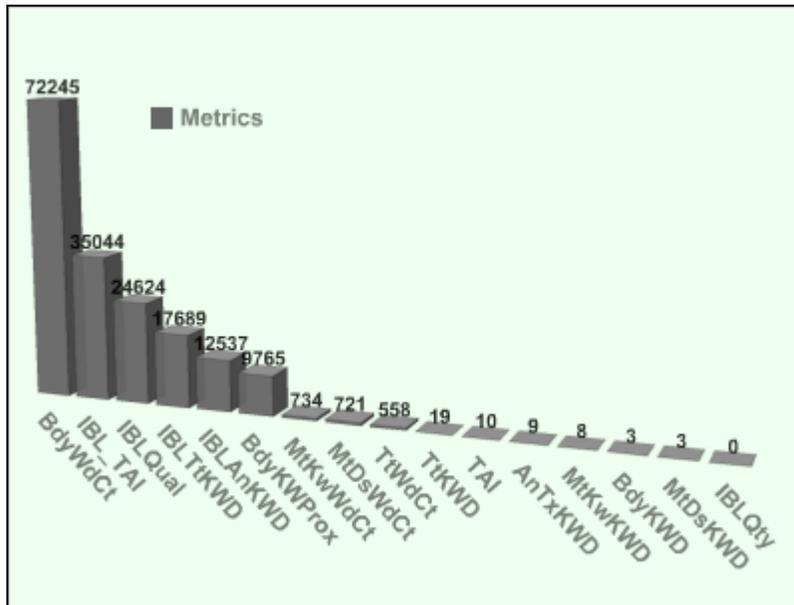
Links

- What is your current PageRank?
- Find who links to you.
- Be wary of link exchange programs.

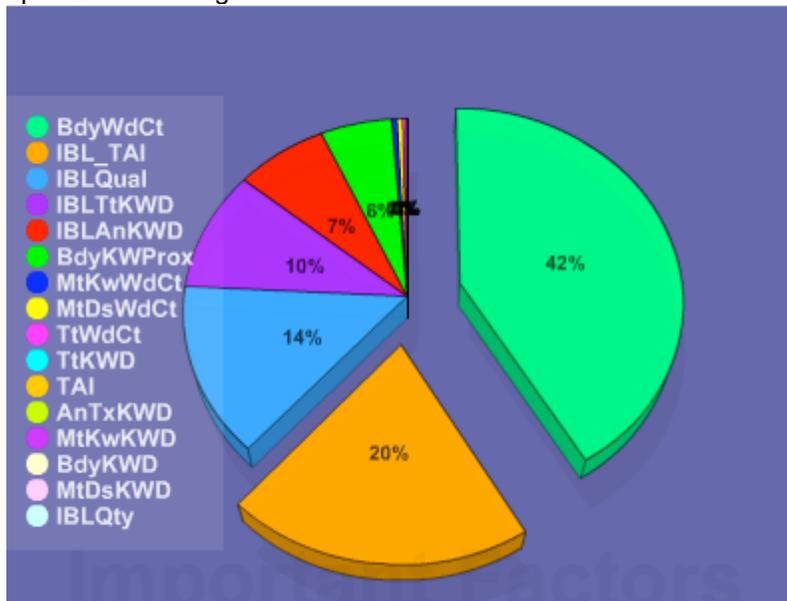
Paying Options

- Sponsored Links
- SEOs: Fortune Interactive

II. On/Off-Page Factors Relative Importance



Notice that the first metric, body word count, is the most significant metric according to the above graph. Since *off*-page factors have been shown again and again to be the most important metrics because they are harder to abuse than *on*-page factors, the above graph should not be taken to be declaring outright that body word count is the most important of all the metrics measures by SEMLogic. Rather, the body word count number is so “important” in the above graph because of the wordiness of the webpages in this particular competitive landscape. We recommend that you pay little attention to that metric, and focus in particular on the four remaining *off*-page metrics, Inbound Link TAI (Theme Relevancy), Inbound Link Quality, Inbound Link Title Keyword Density, and Inbound Link Anchor Keyword Density. A campaign focused primarily (although not exclusively) on those off-page metrics, with inbound link TAI in particular, should have a significantly positive impact on the quality of your web page for the keyword being optimized for Google.



The above graph also demonstrates the dominant sizes of Body Word Count (42%, but not important), IBL TAI (20%), and the remaining off-page factors.