

## Introduction to Information Retrieval on the Web.

*Presented by*

**Mike Grehan**  
**Global VP Content**  
**incisivemedia – New York**

The world's largest repository of digital marketing information.



[www.searchenginewatch.com](http://www.searchenginewatch.com)



[www.searchenginestrategies.com](http://www.searchenginestrategies.com)



[www.clickz.com](http://www.clickz.com)

[www.incisivemedia.com](http://www.incisivemedia.com)

- Data retrieval provides a solution to a user of a database system.
- It does not solve the problem of retrieving INFORMATION about a subject or topic.
- An information retrieval system must interpret the document contents and rank them according to relevance.
- The notion of relevance is at the centre of information retrieval.
- Formerly, information retrieval was of interest only to librarians and information experts.
- Modern information retrieval, in particular information retrieval on the web, is a new form of science establishing its roots.

“Mendel's concept of the laws of genetics was lost to the world for a generation . . . this sort of catastrophe is undoubtedly repeated all about us, as truly significant attainments become lost in the mass of the inconsequential.”

**Vannevar Bush, 1945**



**Bush argued that as humans turned from war, scientific efforts should shift from increasing physical abilities to making all previous collected human knowledge more accessible.**

“Google's mission is to organize the world's information and make it universally accessible and useful.”

**Google Mission Statement**

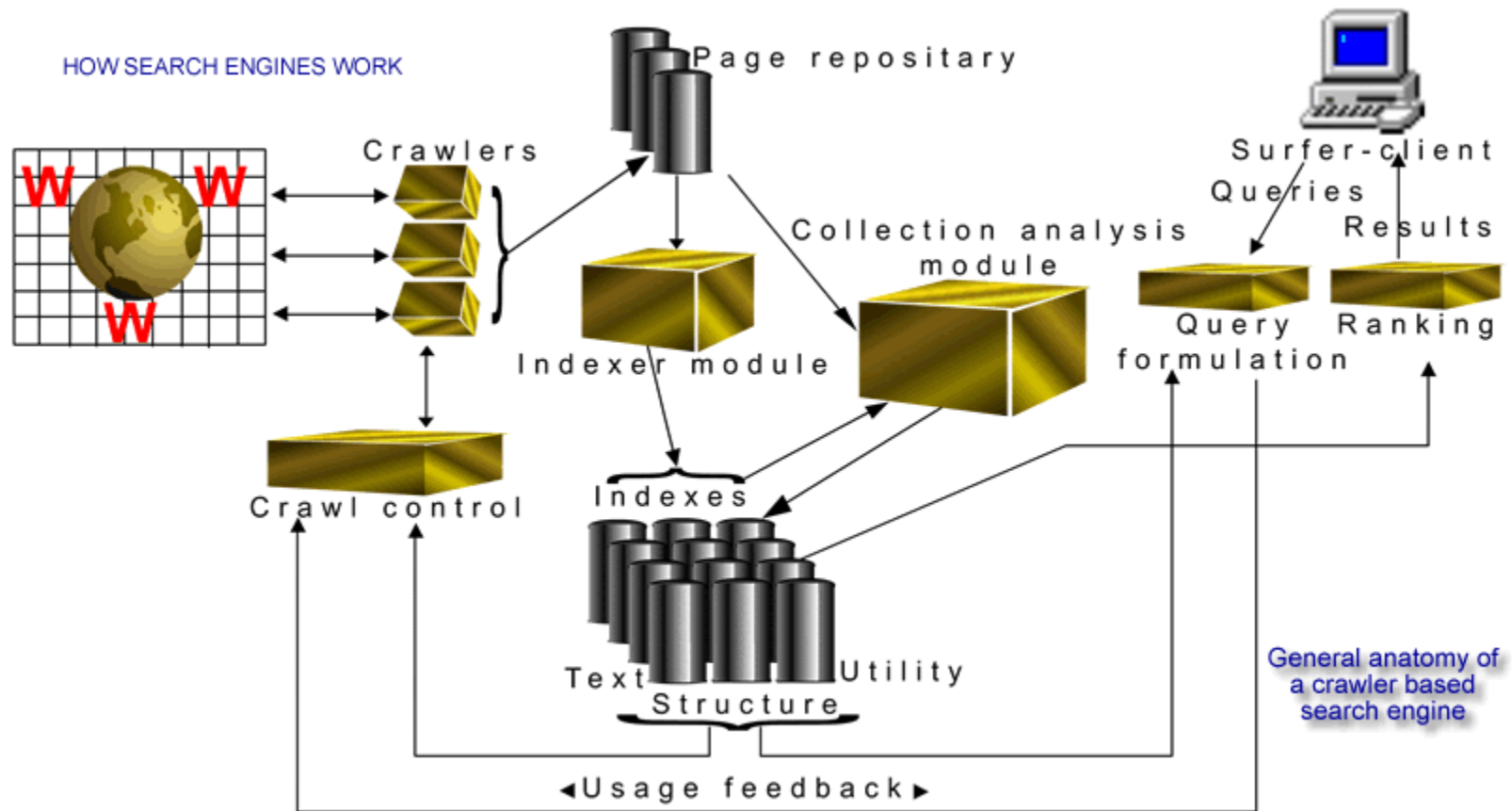




“I just had to take the hypertext idea and connect it to the Transmission Control Protocol and domain name system ideas and — ta-da! — the World Wide Web.”

**Sir Tim Berners Lee, 1989**

# How Search Engines Work



- **Crawlers follow links and collect text: period!**
- **They have only the TINIEST influence over ranking (and that's an off the page factor!)**
- **Stop obsessing over crawler activity!**
- **Would you believe it's possible to rank at a search engine without ever having a page crawled?**



Doc. 1	Doc. 2	Term.	Doc.
Imagine all the people living life in peace.	Yesterday all my troubles seemed so far away.	imagine	1
		all	1
		the	1
		people	1
		living	1
		life	1
		in	1
		peace	1
		yesterday	2
		all	2
		my	2
		troubles	2
		seemed	2
		so	2
		far	2
		away	2

Term.	Doc.		Term.	Doc.
imagine	1	}	all	1
all	1		all	2
the	1		away	2
people	1		far	2
living	1		imagine	1
life	1		in	1
in	1		life	1
peace	1		living	1
yesterday	2		my	2
all	2		peace	1
my	2		people	1
troubles	2		seemed	2
seemed	2		so	2
so	2		the	1
far	2		troubles	2
away	2		yesterday	2

Doc's	Term.	tf in doc 1	tf in doc 2	idf	tfidf doc1	tfidf doc2
Doc 1 Love is joy and joy love.	love	0.33333	0	1	0.33333	0
	is	0.16666	0.125	0	0	0
	joy	0.33	0.125	0	0	0
	and	0.16666	0	1	0.16666	0
Doc 2 A wonderful time is sharing a joy together.	a	0	0.25	1	0	0.25
	wonderful	0	0.125	1	0	0.125
	time	0	0.125	1	0	0.125
	sharing	0	0.125	1	0	0.125
	together	0	0.125	1	0	0.125

**Keywords in title.**

**Title**

**Keywords in headline.**

**Headline.**

**Keywords in body copy.**

Body copy with keywords  
Lorem ipsum. Lorem ipsum  
Lorem ipsum Lorem ipsum  
Lorem ipsum.

**Keywords in alt tags.**

Blah, blah  
blah. Blah  
blah blah.  
Blah blah  
blah



**Text links and/or link to site map.**

Lorem ipsum Lorem ipsum  
Lorem ipsum. Lorem ipsum  
Lorem ipsum Lorem ipsum.  
Lorem ipsum Lorem ipsum  
Lorem ipsum.

**Link : Link : Link : Link**

**A page written about Beethoven's Fifth symphony by a music student may have the same keyword density (and relevancy) as a page on the subject written by famous orchestra conductor Andre Previn...**

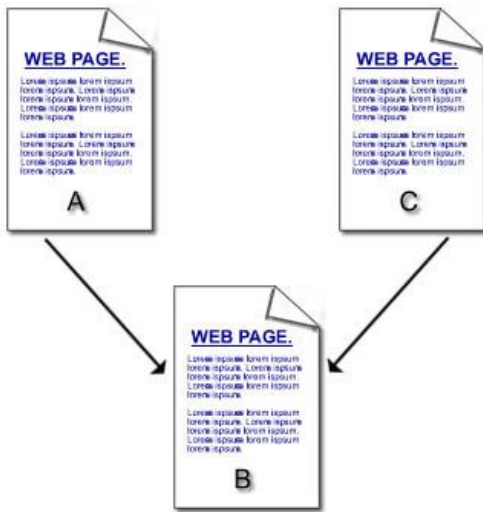
**Who is likely to write the most authoritative piece?**

### Hyperlink analysis for the web.

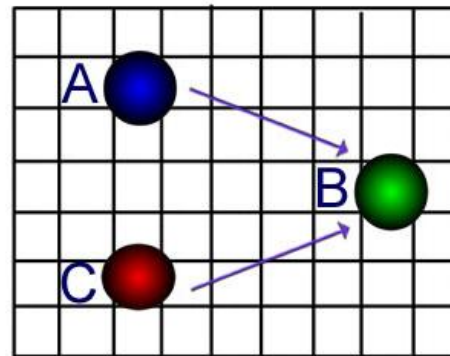
- PageRank and HITS by Page/Brin – Kleinberg
- Pinski & Narin – Citation Analysis (Uses influence weights)
- Hubs, authorities and communities
- Not all links are equal – some links are more equal than others...  
and some are infinitely more equal.

Hyperlink analysis algorithms make either one or both of these simple assumptions:

- **Assumption 1** - A hyperlink from page A to page B is a recommendation of page B by the author of page A.
- **Assumption 2** – If page A and page B are connected by a hyperlink, then they might be on the same topic.



Set of linked pages.

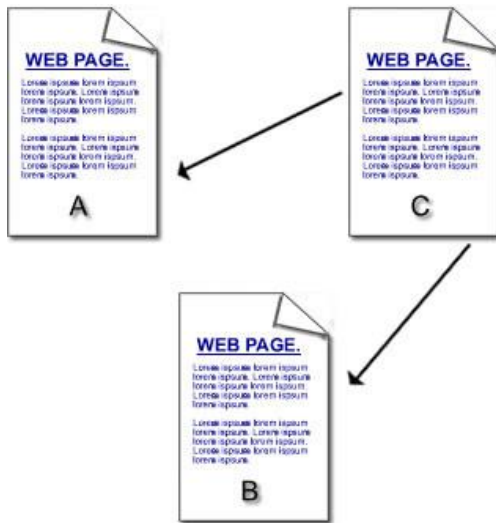


Link graph.

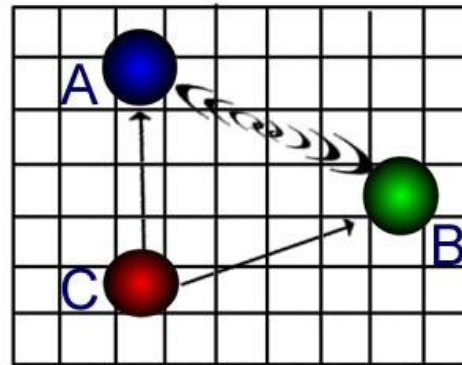
Web pages linked together are nodes in the web graph. When web page A links to web page B, this is a 'directed edge.'

Search engines see much more in linkage data.

- If page C cites pages A and B, then A and B are said to be co-cited by C.
- Pages A and B being co-cited by many other pages is evidence that A and B are somehow related to each other.



Set of linked pages.



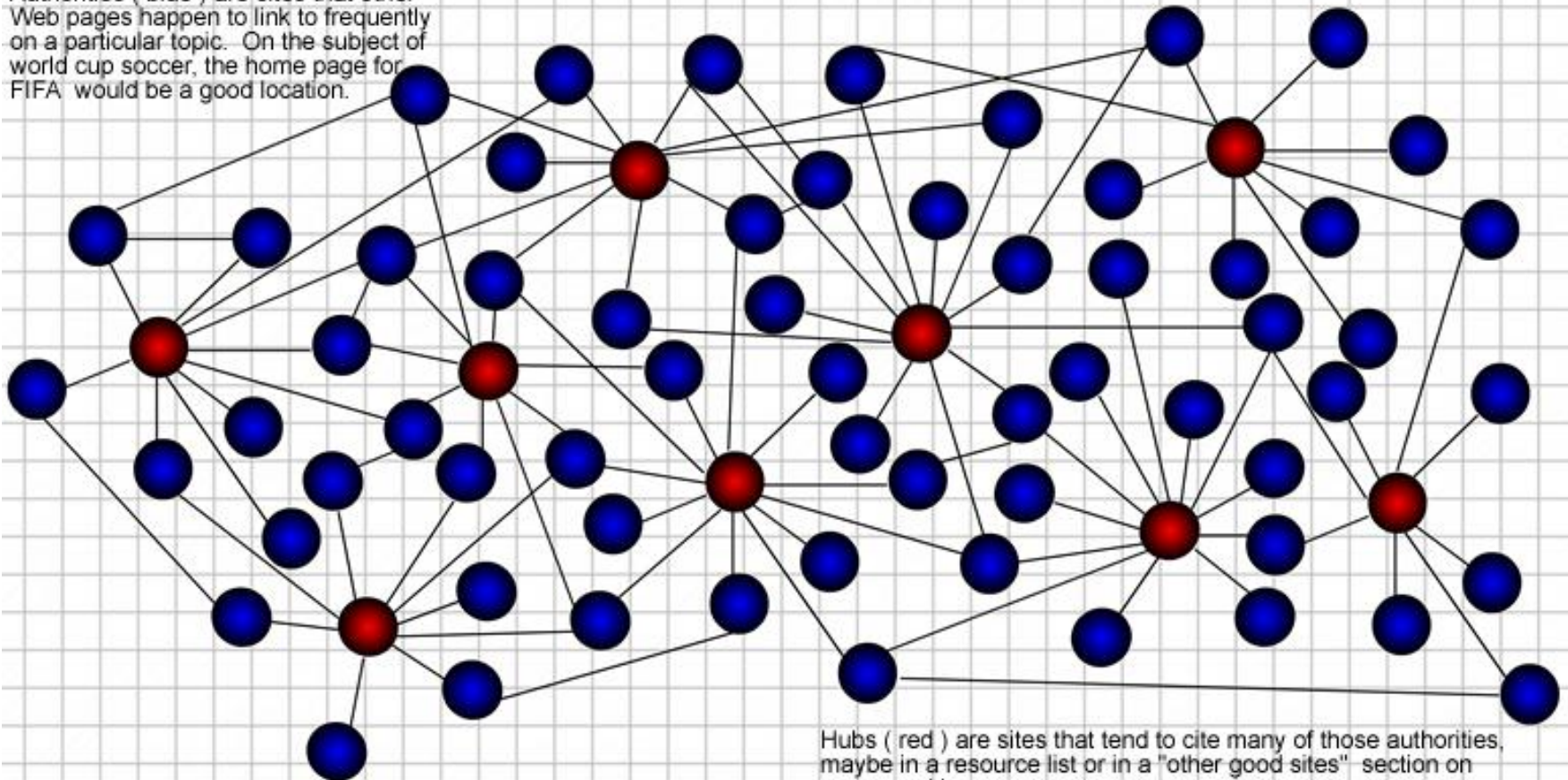
Link graph.

If page C links to both A and B, then A and B are connected by an undirected edge in the graph and are viewed as being co-cited by C.



### Kleinberg's 'hubs & authorities'.

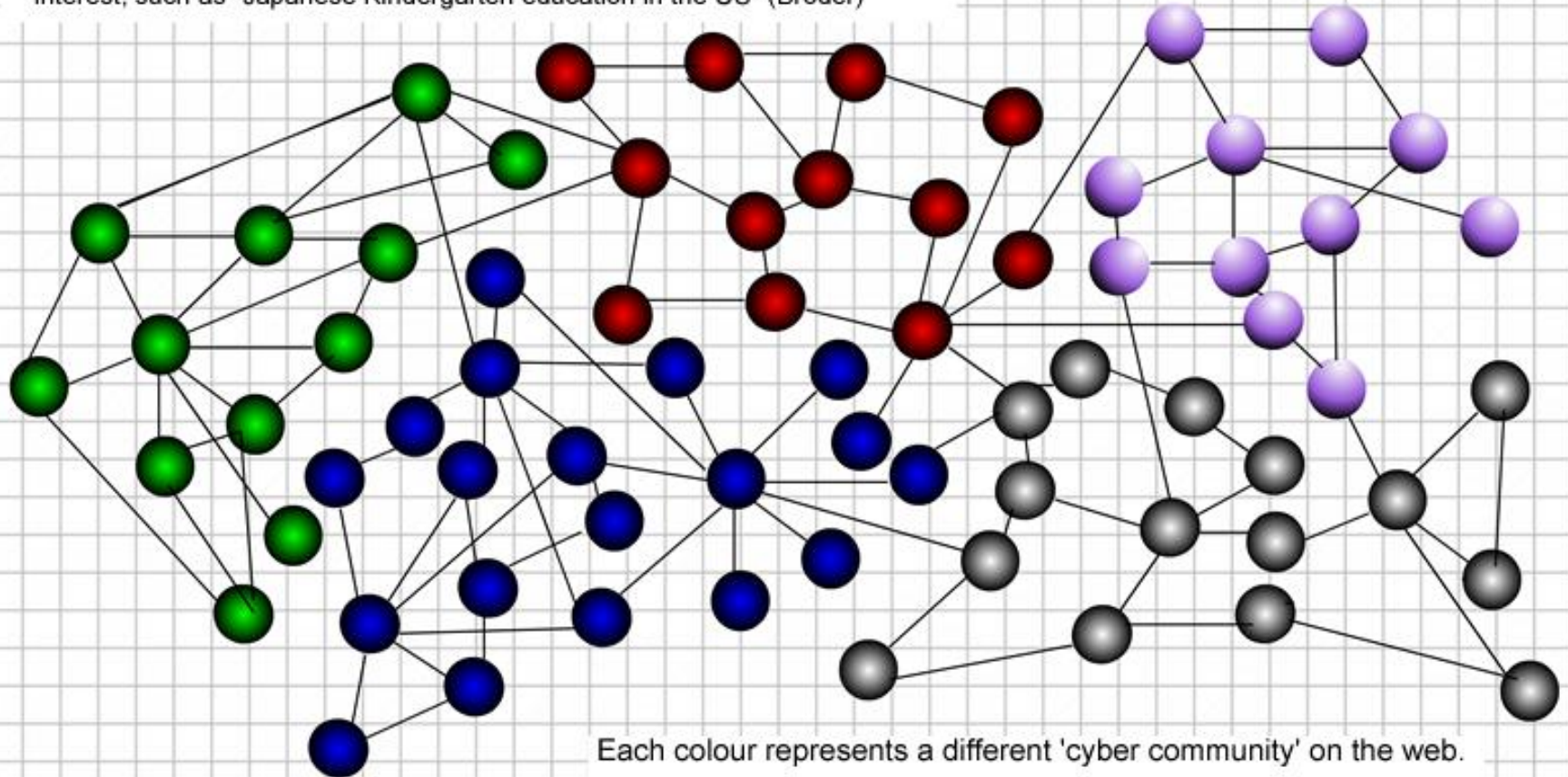
Authorities ( blue ) are sites that other Web pages happen to link to frequently on a particular topic. On the subject of world cup soccer, the home page for FIFA would be a good location.



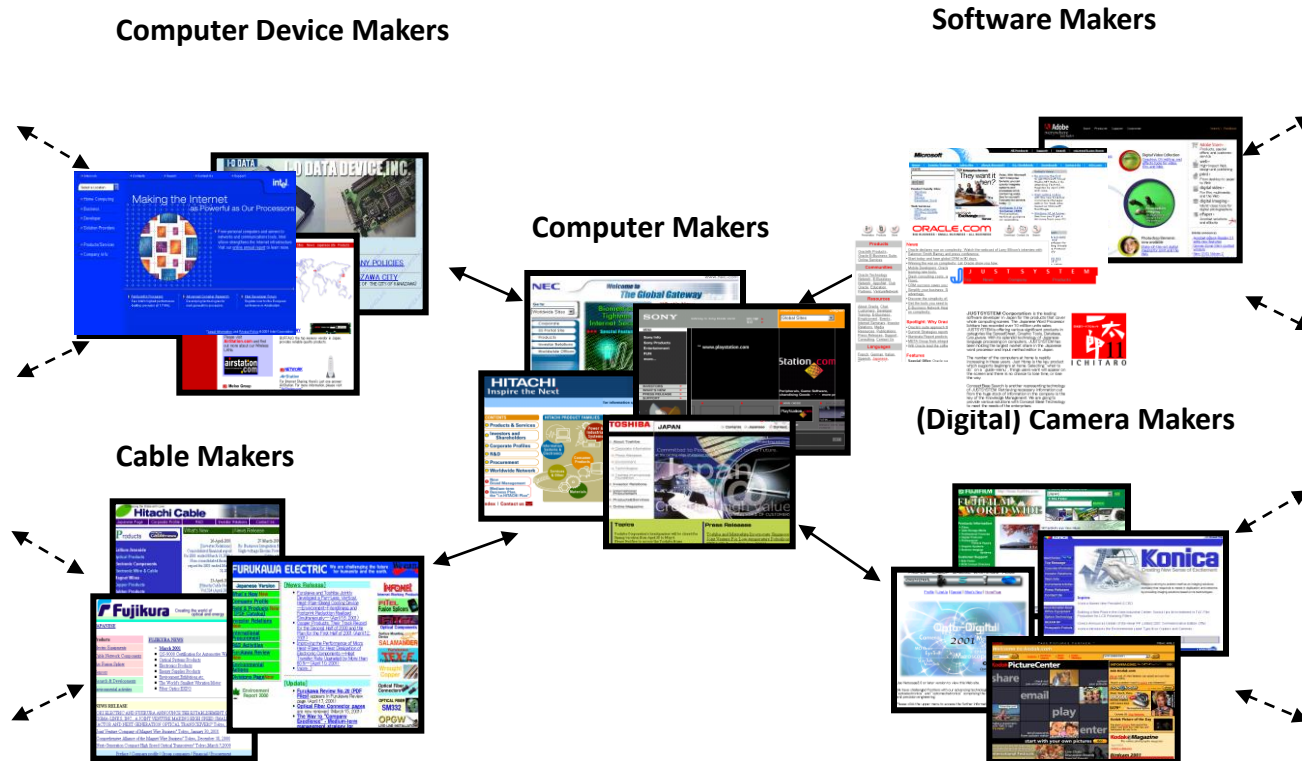
Hubs ( red ) are sites that tend to cite many of those authorities, maybe in a resource list or in a "other good sites" section on a personal home page.

## The New Science Of Networks – Cyber Communities

- Link analysis can discover 'cyber communities', even pages with a very narrow interest, such as "Japanese Kindergarten education in the US" (Broder)



Masashi Toyoda - Kentarou Fukuchi





[Advanced Search](#)  
[Preferences](#)

Web

Results 1 - 10 of about 2,400,000 for [history](#) of [cookies](#). (0.35 seconds)

Tip: Save time by hitting the return key instead of clicking on "search"

## [History of Cookies](#)

Earlier names for **cookies** such as Snickerdoodles and Cry Babies originated with the New England states. Even with its early **history**, **cookies** did not become ...

[whatscookingamerica.net/History/CookieHistory.htm](http://whatscookingamerica.net/History/CookieHistory.htm) - 89k - [Cached](#) - [Similar pages](#)

## [The History of Cookies](#)

A little **cookie history**: The first **cookies** were created by accident. Cooks used a small amount of cake batter to test their oven temperature before baking a ...

[www.kitchenproject.com/history/cookies.htm](http://www.kitchenproject.com/history/cookies.htm) - 20k - [Cached](#) - [Similar pages](#)

## [History Of Cookies](#)

The **history of cookies**. THE NIBBLE, Great Food Finds, is an online gourmet food magazine with 1000+ reviews of gourmet food and beverages, including gourmet ...

[www.thenibble.com/reviews/main/cookies/cookies2/cookie-history.asp](http://www.thenibble.com/reviews/main/cookies/cookies2/cookie-history.asp) - 55k -

[Cached](#) - [Similar pages](#)

## [The Food Timeline: history notes--cookies, crackers & biscuits](#)

According to the food **history** reference books, "Ammonia" **cookies** are not one ..... Food historians trace the **history** of these **cookies** and cakes to Medieval ...

[www.foodtimeline.org/foodcookies.html](http://www.foodtimeline.org/foodcookies.html) - 209k - [Cached](#) - [Similar pages](#)

## [The History of Cookies - Associated Content](#)

Oct 27, 2006 ... Everyone loves them, but where did **cookies** come from?

[www.associatedcontent.com/article/75995/the\\_history\\_of\\_cookies.html](http://www.associatedcontent.com/article/75995/the_history_of_cookies.html) - 51k -

[Cached](#) - [Similar pages](#)

# Changes In End User Behavior

[Web](#) [News](#) [Images](#) [Video](#) [Blogs](#) [Music](#)

Results 1 - 10 of about 69,400,000 for **beyonce**. (0.07 seconds)

## News results for **beyonce**



[Beyonce Sings For Obamas](#) - 7 hours ago  
Barack and Michelle Obama had their first dance since the inauguration as **Beyonce** Knowles sung last night in Washington. The former Destiny's Child star ...  
[MTV UK](#) - [2119 related articles »](#)  
[Beyonce to sing 'At Last' for the Obamas](#) - [The Associated Press](#) - [628 related articles »](#)  
[Bono, Bruce, Beyonce to play inaugural concert](#) - [MSNBC](#) - [562 related articles »](#)

## [Home](#) | [The Official Beyonce Site](#)

The official site of the singer offers a biography, hit list, videos, photographs, news and events calendar.

[www.beyonceonline.com/](#) - 67k - [Cached](#) - [Similar pages](#) - 

## [Beyonce on MySpace Music](#) - [Free Streaming MP3s, Pictures & Music ...](#)

SASHA FIERCE is **Beyoncé's** first new studio collection since her ... **Beyoncé's** SASHA FIERCE performances include "Single Ladies (Put A Ring On It)," "Radio," ...

[www.myspace.com/beyonce](#) - 181k - [Cached](#) - [Similar pages](#) - 

## [Beyoncé Knowles](#) - [Wikipedia, the free encyclopedia](#)

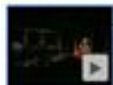
In support of the album, Knowles embarked on her lengthy The **Beyoncé** Experience concert tour, visiting over ninety venues worldwide. ...

[en.wikipedia.org/wiki/Beyoncé\\_Knowles](#) - 282k - [Cached](#) - [Similar pages](#) - 

## [Image results for beyonce](#) - [Report images](#)



## Video results for **beyonce**



[YouTube](#) - [Listen - Beyonce](#)  
3 min 48 sec  
[www.youtube.com](#)



[Beyoncé - If I Were A Boy](#)  
5 min  
[www.youtube.com](#)

## [Blog posts about beyonce](#)

[Beyonce's Obama Joy: Offers To Sing For Him](#) - [Latest News](#) - Nov 6, 2008

[EURweb.com - BEYONCE VIDEOS DOMINATE YOUTUBE: Singer's music clips ...](#) -

[Black Entertainment : Black News : Urban News ...](#) - Jan 5, 2009

[Beyonce Sings 'Fall On Me' On Inaugural Ball As Obamas](#)

### Informational.

**“This applies to the surfer who is really looking for factual information on the web. So they make a query like say...‘low haemoglobin’ for instance. This is a medical condition. They are looking for specific information about this condition. That’s very close to classical information retrieval.”**

### **Navigational.**

**“Navigational is when a surfer really wants to reach a particular web site. If they do a query like, say, United Airlines, for instance. Probably what they really want is to go directly to the web site of United Airlines – like [www.ua.com](http://www.ua.com) just like if someone typed BBC, it’s most likely they want the web site of the BBC - and not the history of the BBC and broadcasting. They probably want to just go directly to the web site. We all do a lot of these types of searches.”**

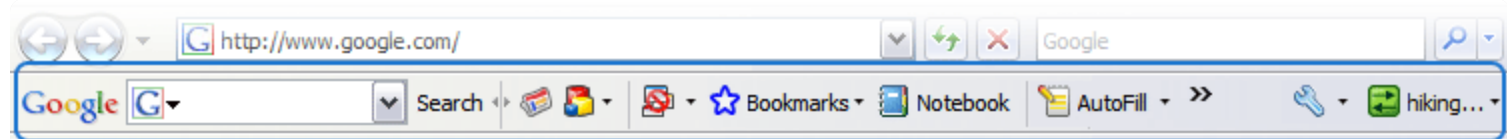
### Transactional.

**“Transactional means that ultimately the surfer wants to do something on the web, through the web. Shopping is a good example. You really want to buy stuff. Or you want to download a file, or find a service like, say, yellow pages. What you really want to do is get involved in a transaction of information or services. Take a shopping query, these are transactional queries where people want to buy stuff and so on. So, they are wanting a return which satisfies this need.”**



- Query Chains
- User Trails

- **Text** on an HTML page
- **Linkage** data and link anchor text
- **Social Media** – Tagging, bookmarking, rating etc.



# User Intent

night stands - Google Search - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.google.com/#hl=en&source=hp&sq=night+stands&aq=f&aqi=s&aql=g10&aq=sfp=e8d6ef47431c6a4e

night stands - Google Search

Web Images Videos Maps News Shopping Gmail more ▾

Google night stands Search Advanced Search

Web [Show options...](#) Results 1 - 10 of about 26,900,000 for night stands [definition] (0.29 seconds)

**Night Stand Superstore** Sponsored Links  
[NightStands-Store.com/Night\\_Stands](#) ALL Top Night Stands - Save 10%-30% 400+ Night Stands in Stock, Buy Now [Google Checkout](#)  
[Show products from this advertiser](#)






**Decorative Night Stands**  
[www.HomeDecorators.com/Nightstands](#) Stunning Array Of Bedside Tables & Decorative Nightstands/Top Values  
[Show products from Home Decorators Collection for night stands](#)

**Night Stands**  
[www.Target.com/furniture](#) Night Stands Online. Shop at Target.  
[Show products from Target for night stands](#)


**Amazon.com: Nightstands**  
Online shopping for Nightstands from a great selection of Furniture & Decor & more at everyday low prices.  
[www.amazon.com/Nightstands-Bedroom-Furniture.../b?ie... - Cached - Similar](#)


**Walmart.com: For the Home: Bedroom Furniture: Night Stands**  
Shop Low Prices on Bedroom Furniture: Night Stands  
[www.walmart.com / ... / Furniture / Bedroom Furniture - Cached - Similar](#)

Shopping results for night stands

 <a href="#">South Shore Furniture</a> 3127-062 \$76 new 36 stores	 <a href="#">Winsome 82115 Night Stand with</a> \$44 new 26 stores	 <a href="#">Winsome 82115 Natural Beechwood</a> \$38 new 26 stores	 <a href="#">Winsome 94918 Curved Night Stand</a> \$36 new 24 stores	 <a href="#">Nexera 4163 Pablo 3 Drawer Night Stand</a> \$160 new 26 stores
---	--	---	--	---

**Sponsored Links**

  
[Winsome 82115 Natural Beechwood End table](#)  
\$38.90 - Unbeatable Sale

  
[Nexera 5052 Renaissance 2 Drawer Night Stand](#)  
\$135.00 - Unbeatable Sale

**Sauder**  
Durable Wood & Laminate Furniture from Sauder! Order Online & Save  
[www.OfficeDepot.com](#)

**Nightstands Furniture**  
Over 450 Nightstands. We Beat ANY Advertised Prices. Free Shipping!  
[BedroomFurnitureMore.com](#)

**Bedroom Furniture**  
Shop for bedroom suites and more online at Horchow.com.  
[www.Horchow.com](#)

**Wood Night Stand**  
Save 30-60% On Our Selection Of Shop Today. Savings End Soon.  
[JCPenney.com](#)

Done

night stands - Google Search - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.google.com/?hl=en&source=hp&q=night+stands&aq=f&saql=s&aq=10&saql=s&fp=eb09ef47431c6a4e


night stands - Google Search


Web Images Videos Maps News Shopping Gmail more ▾


Google night stands Search Advanced Search


Web Show options... Results 1 - 10 of about 26,900,000 for night stands [definition] (0.29 seconds)


**Night Stand Superstore**  
NightStandStore.com/Night\_Stand ALL Top Night Stands - Save 10%-30% 400+ Night Stands in Stock. Buy Now! [Google Checkout .fr](#)  
[Hide products from this advertiser](#)


  
[Bossetto USA 20 Inch Secret Night](#)  
\$541.61

  
[Night Juniper 2 Drawer Night Stand](#)  
\$395.42


  
[Tommy Bahama Home Barbados Night Stand](#)  
\$999.00


  
[LifeStyle Solutions Knitch 1](#)  
\$309.00


  
[Nexera Nocce 2 Drawer Night Stand](#)  
\$153.40


  
[Home Styles Furniture Bedford](#)  
\$133.11


**Decorative Night Stands**  
[www.HomeDecorators.com/Nightstands](#) Stunning Array Of Bedside Tables & Decorative Nightstands/Top Values  
[Hide products from Home Decorators Collection for night stands](#)


  
[Hawthorne One-Drawer Anywhere](#)  
\$129.00

  
[Hawthorne One-Drawer Anywhere](#)  
\$129.00

  
[Hawthorne Open Anywhere Table](#)  
\$99.00


  
[Hawthorne Open Anywhere Table](#)  
\$99.00


  
[Hawthorne Open Anywhere Table](#)  
\$99.00

  
[Hawthorne Open Anywhere Table/Telephone](#)  
\$139.00

**Night Stands**  
[www.Target.com/furniture](#) Night Stands Online. Shop at Target.  
[Hide products from Target for night stands](#)

**Sponsored Links**

  
[Winsome B2115 Natural Beechwood End table](#)  
\$38.90 - Unbeatable Sale

  
[Nexera 5052 Renaissance 2 Drawer Night Stand](#)  
\$135.00 - Unbeatable Sale

**Sauder**  
Durable Wood & Laminate Furniture from Sauder! Order Online & Save [www.OfficeDepot.com](#)

**Nightstands Furniture**  
Over 450 Nightstands We Beat ANY Advertised Prices. Free Shipping! [BedroomFurnitureMore.com](#)

**Bedroom Furniture**  
Shop for bedroom suites and more online at [Horchow.com](#) [www.Horchow.com](#)

**Wood Night Stand**  
Save 30-60% On Our Selection Of Shop Today. Savings End Soon. [JCPenney.com](#)

- User generated content **5:1** vs. mediated content
- **Richer** *search* **experience**
- A **comprehensive** knowledge repository?
- Demand for many **types** of information
- Expanded **feedback loops**



- Perhaps the biggest change in search is the shift toward information-seeking on social networking sites.
- The knowledge possessed by your friends and people you know acts as a supplement to the web's huge amount of other, less verifiable information.
- This knowledge can provide extremely qualified answers to specific queries through a process defined as:  
*information-seeking via a chain of trust.*







- Local, mobile, social, multimedia **convergence**
- Digital communities **counter-balance** one-sided brand messages
- Applications **sidestep web browsers** to deliver specialized content

“As content becomes more diverse, more complex, bigger and more fragmented...getting it through HTTP and HTML may not be the right model anymore.”

**Andrew Tomkins**  
VP, Search Research, Yahoo!

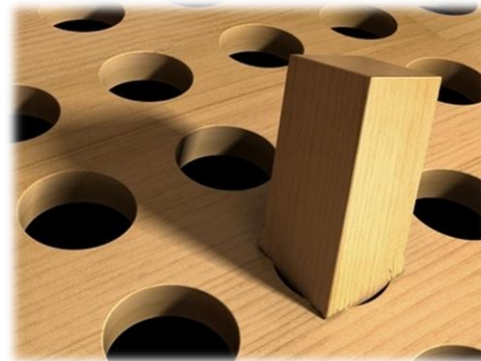


“So how many unique pages does the web really contain? We don't know; we don't have time to look at them all!”

**Official Google Blog**

“It's amazing what people have done with HTML, which was never intended to do rich Internet applications...these technologies were designed for something else, and people are really hacking them to do more.”

**Brad Becker, Microsoft**



“Many years ago your grandparents sat in front of a brown wooden box listening to Franklin D. Roosevelt. Who knows, maybe last week you bought a new HD-ready TV. Guess what? Same airwaves, different protocol.”

**Mike Grehan**



**You wake up and check** your email on your bedside iPad — that's one app. During breakfast you browse Facebook, Twitter, and *The New York Times* — three more apps. On the way to the office, you listen to a podcast on your smartphone. Another app. At work, you scroll through RSS feeds in a reader and have Skype and IM conversations. More apps. At the end of the day, you come home, make dinner while listening to Pandora, play some games on Xbox Live, and watch a movie on Netflix's streaming service.

You've spent the day on the Internet — but not on the Web. And you are not alone.

“The web is no longer just a huge collection of web pages linked together. It is a network of networks of people who are constantly connected to each other, communicating and sharing. And they’re no longer stuck in front of monitors and keyboards in homes and offices – they’re on the move.”

Mike Grehan 2008

# Connected Marketing.



## What is Connected Marketing really about?

---

- Connected Marketing is about the change in consumer behavior than it is about the technology available.
- Much of the received wisdom in marketing circles is undergoing reconsideration. The nature of consumer and business markets is going through major change.
- Decision makers no longer act independently of each other but are all the more connected to other consumers, to other channel members and often to brands.
- In turn, brands and companies are now vying for central positions inside consumer networks.
- Consumers are attracted to rapidly emerging third party information providers , through collaborative product review and price comparison services.

## Complexities of modern marketing.



This is a very new and complex concept applied to marketing.  
It does make us think differently. Not because of technology.

Because we're trying to attract an entirely different kind of consumer.

**Thank You!**

**Q & A**