



# Propaganda, Spam and the Evolution of Search Engines

Panagiotis Takis Metaxas

**@takis\_metaxas**

Professor, Computer Science, Wellesley College  
Affiliate, CRCS, Harvard University

AACAC, October 26, 2012

# Who do you consult for answers?

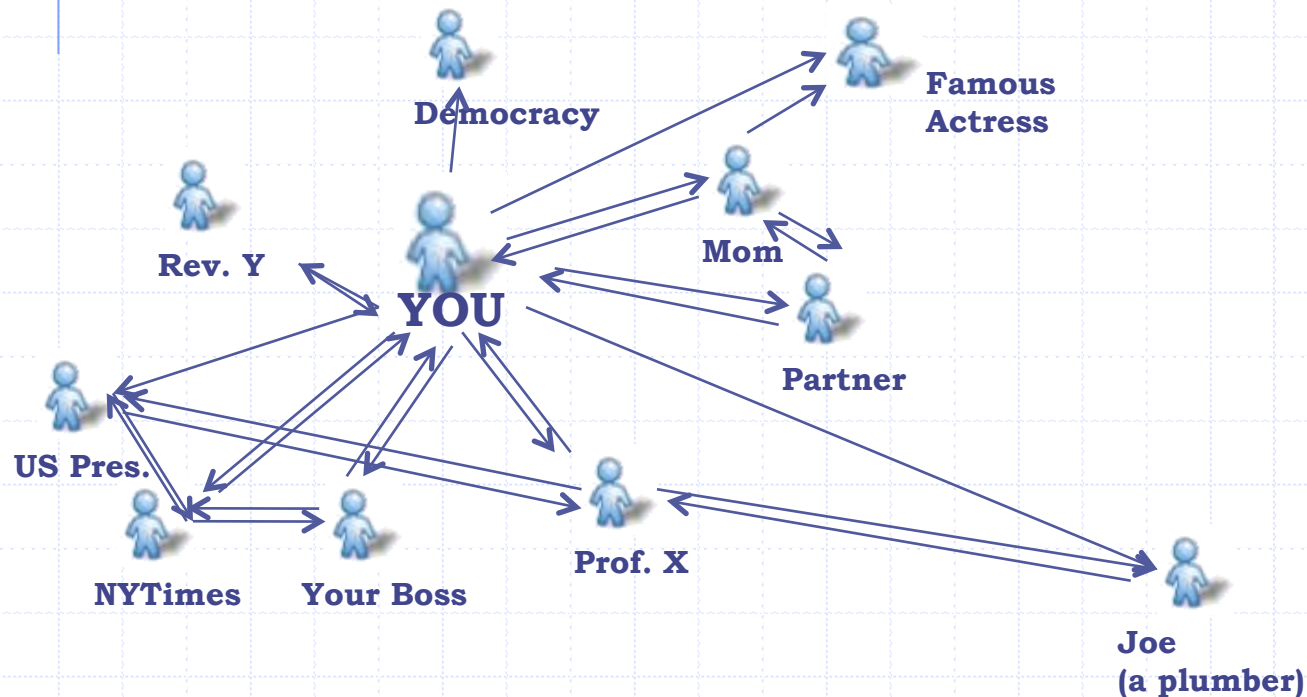
- ◆ Myself and my experiences
- ◆ A friend or family member
- ◆ A religious authority
- ◆ A professor
- ◆ Someone else
- ◆ A book or librarian
- ◆ A newspaper
- ◆ An ad that I like
- ◆ The Internet

## How did you decide who to ask:

- ◆ Are they **knowledgeable** on the issue?
- ◆ Can I **trust** them?
- ◆ Is it **convenient** to ask them?

# Your Trust Network

- ◆ Network of Nodes and Arcs (directed edges)
  - Nodes = social entities (people, entities, sources, ideas)
  - Arcs = **trust relationships** from an entity to another
    - Length of arc = strength of trust
- ◆ We can explore it (mentally)
- ◆ We change/verify/augment it all the time



# CHALLENGES to your Trust Network



# Challenges through Propaganda

## ◆ Propaganda:

- Attempt to **modify** human behavior, and thus **influence** people's actions in ways **beneficial** to propagandists

## ◆ Theory of Propaganda

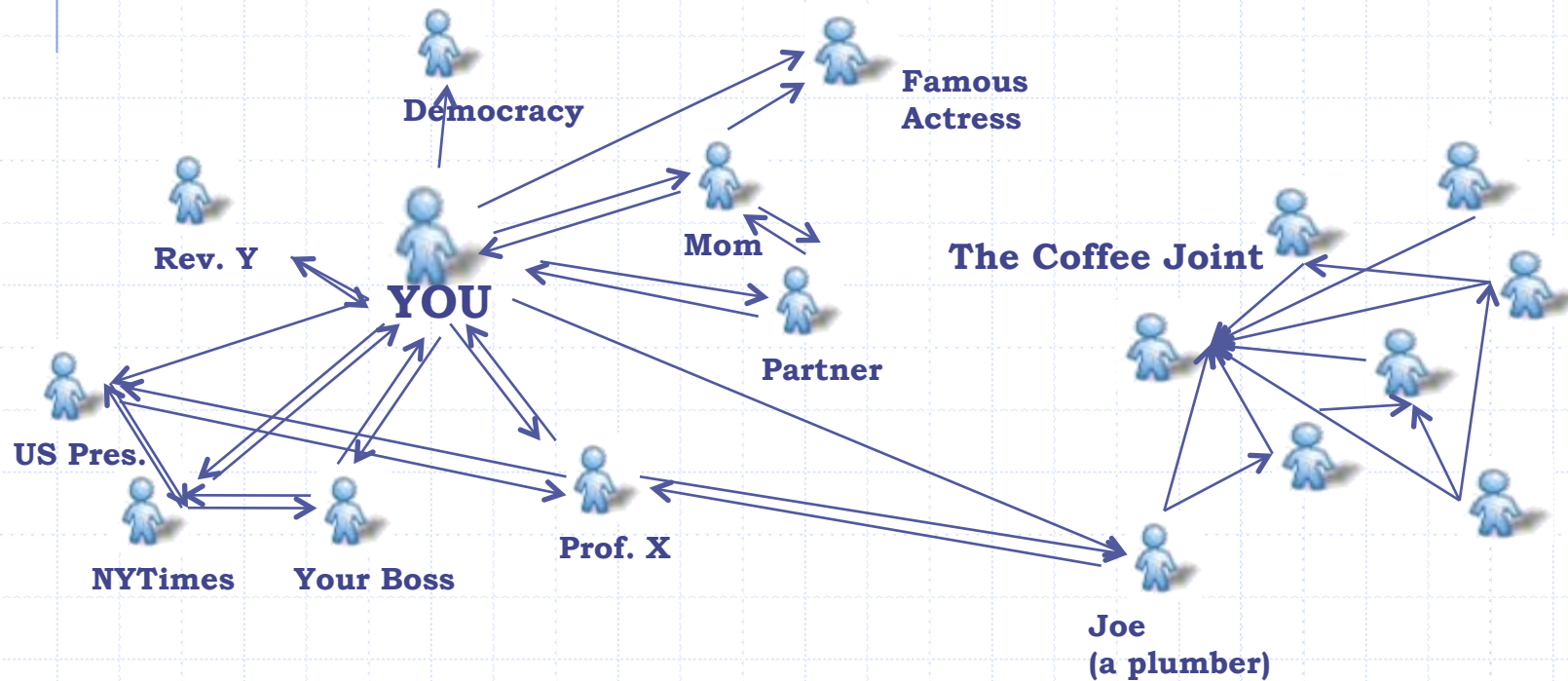
- Developed by the Institute for Propaganda Analysis 1938-42

## ◆ Propagandistic Techniques (and ways of **detecting** propaganda)

- **Word Games (Glittering Generalities & Name Calling)** - associate good/bad concept with social entity
- **Plain Folk** - people like us think this way
- **Bandwagon** - everybody's doing it, jump on the wagon
- **Testimonial** - famous non-experts' claims
- **Card Stacking** - use of bad logic
- **Transfer** - use special privileges (e.g., office) to breach trust

# The Bandwagon Technique

◆ *with it the propagandist attempts to convince us that all members of a group to which we belong are accepting his program and that we must therefore follow our crowd and “jump on the band wagon”*



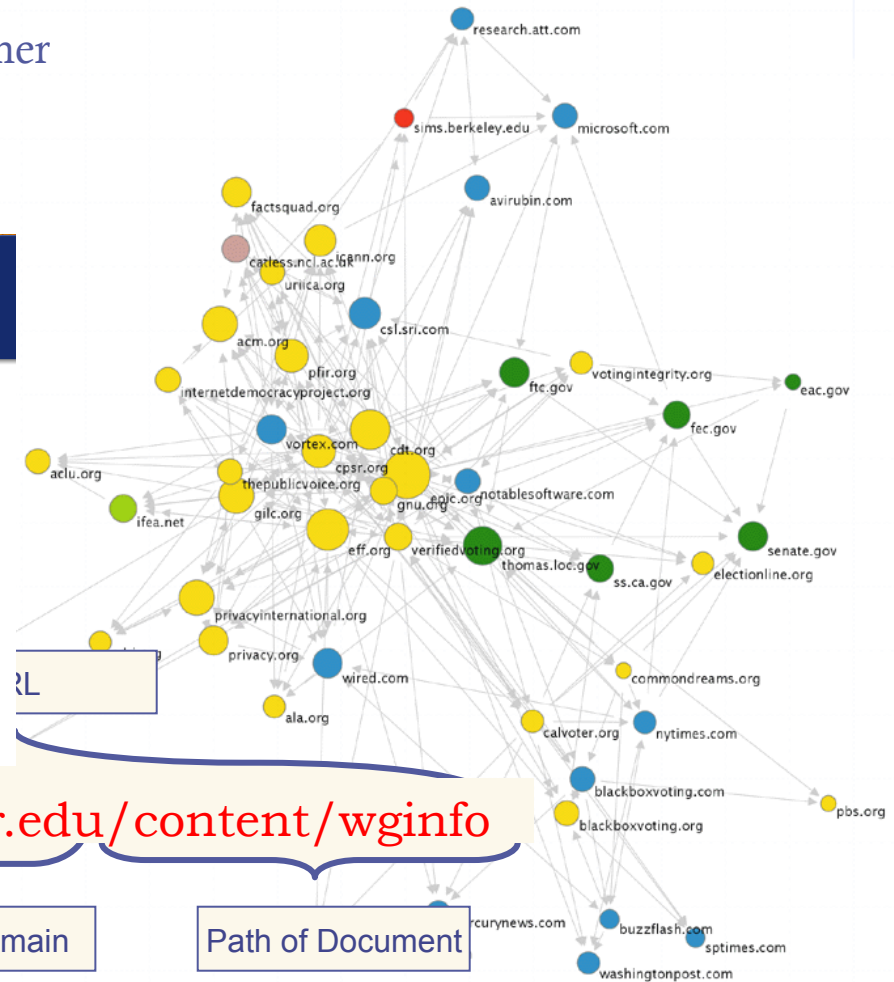
# The Testimonial Technique

- ◆ *having some respected person say that a given idea or program or product or person is good or bad*



# The Web is also a network

- ◆ Network of Nodes and Arcs (directed edges)
  - Nodes = web pages
  - Arcs = **hyperlinks** from a page to another
- ◆ A network can be **explored**
- ◆ A network can be **indexed**



**BRYN MAWR** Department of Computer Science

Academics Admissions & Financial Aid Student Life Campus Offices About Us

Academics · Fields of Study  
Department of Computer Science

Home  
About Us  
People  
Curriculum  
Research  
Labs  
Resources

**AALAC/Mellon 23 Working Group on Information**  
Fall 2012 Meeting  
**AALAC/Mellon 23 Working Group on Information**  
Co-Sponsored by the Center for Science of Information

October 26-27, 2012  
Bryn Mawr College  
Bryn Mawr, PA 19010  
Maps & Directions  
Parking: Between Buildings 38 & 40 on Campus Map

**Keynote speaker: Prof. Sanjeev Kulkarni, Princeton University & Center for Science of Information, Title: Science of Information Meets the Liberal Arts**

**Workshop Leaders**  
Deepak Kumar, Computer Science – Bryn Mawr College (Project Leader)  
Duane Bailey, Computer Science – Williams College  
Tzu-Yi Chen, Computer Science – Pomona College  
Panagiotis T. Metaxas, Computer Science – Wellesley College  
Richard Wicentowski, Computer Science – Swarthmore College

**Agenda:**  
All Sessions will be held in Room 300, Dalton Hall (Building 28 on Campus Map)

<http://cs.brynmawr.edu/content/wginfo>

Access method

Server and domain

Path of Document



# Use Web, Search Engines, Twitter?

- ◆ [ ] of the Americans had internet access in May, 2008
  - More than half of connected Americans use broadband
- ◆ [ ] of all US adults went online to get news or information about the 2010 midterm elections
- ◆ [ ] of all internet users use a search engine regularly
- ◆ [ ] of US adults received candidate or campaign information on social networking sites or Twitter





# Reliability of Information on the Web



*Anyone* can be an author on the web!

In the past, the greatest problem facing researchers was **finding** information; now, with the advent of the Internet, the greatest problem is **evaluating** the vast wealth of information available.

Students [...] placed greater emphasis on the **process** of finding an answer than on **analyzing** the actual information.

# “OF COURSE IT’S TRUE; I SAW IT ON THE INTERNET!”

## *Critical Thinking in the Internet Era*

*Students use the Net as a primary source of information, usually with little or no regard as to the accuracy of that information.*

**T**he Internet is revolutionizing research methods at colleges and universities around the world. Though it can be extremely useful to researchers, the Net presents a significant challenge in that it is quite different from traditional sources. The lack of uniform standards and the ease of access have made the Internet a powerful but uncertain medium. Substantial effort is required to adequately evaluate its information, and this may not always be apparent to users [5]. This is particularly challenging for students, as many have come to rely on the Net as a primary source of information without formal instruction about the difficulties involved. The Internet has gained a primary place in research methods, and it is vital that students become able to critically evaluate the information it provides.

Several solutions have been suggested to determine accuracy in Internet research. In [1], Jerry Campbell supports the Association of Research Libraries' plan to develop an Internet portal to "trustworthy" information. This portal would "promote the development of and provide access to the highest quality content on the Web." Many colleges have also adopted this approach by pro-

viding lists of approved online sources to students. While it appears to provide a practical alternative to information.coms that focus more on advertising than accuracy, this approach suffers from several drawbacks. First, it is impossible to continually monitor all the content found using these portals. Web sites change overnight and expand at exponential rates, and attempting to continuously ver-

ILLUSTRATION BY LISA HANEY

# The Web has misinformation!



Web Images Groups <sup>New!</sup> News Froogle more »

can human growth hormone increase muscle mass?

Search

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

Web

Results 1 - 10 of about 313,000 for [can human growth hormone increase muscle mass?](#). (0.26 seconds)

## [Human Growth Hormone](#)

[www.HGH-Info-Buyers-Guide.org](http://www.HGH-Info-Buyers-Guide.org)

Doctor Recommendations & Warnings. Find Info on Effective & Safe HGH!

Sponsored Link

Tip: Have a question? Ask the researchers at [Google Answers](#).

## [Human Growth Hormone Natural HGH Supplement Decrease Fat and Gain ...](#)

... **Increase** stamina. Improve immune functioning. ... Now, you **can** benefit from the premiere HGH ... information you need on hGH, **human growth hormone** hGH releaser, youth ...

[www.hghplus.net/](http://www.hghplus.net/) - 23k - [Cached](#) - [Similar pages](#)

## [Human Growth Hormone - "Rejuvenation" Medicine?](#)

... the list of the age-depleted **hormones** that **can** be replaced ... is natural; it is plentiful in the **human** body during ... weight gain of 30-50 pounds occur when HGH wanes ...

[www.usdoctor.com/gh.htm](http://www.usdoctor.com/gh.htm) - 17k - [Cached](#) - [Similar pages](#)

## [Secretagogue-One can increase - hgh, hGH - Human Growth Hormone](#)

Maximum **Human** Performance (MHP ... almost embarrassed to carry Secretagogue-O **hormone** and IGF ... Enhances all other **hormone** activities; **Increases** hGH and IGF1; ...

[www.powersupplements.com/hgh.htm](http://www.powersupplements.com/hgh.htm) - 12k - [Cached](#) - [Similar pages](#)

## [Best HGH for Bodybuilding - HGH Supplement](#)

... Yes, You **Can** Now Get Actual **Human Growth Hormone** ... **Human Growth Hormone** (HGH) is one of the most powerfully effective anabolic **hormones** available. ...

[www.trulyhuge.com/hgh.htm](http://www.trulyhuge.com/hgh.htm) - 14k - Jan 29, 2005 - [Cached](#) - [Similar pages](#)

## [Human Growth Hormone HGH - Pro HGH Supplement | Human Growth ...](#)

... HGH) over the counter, however, you **can** improve your ... The same study found that administering **human growth hormone** (hgh) resulted in a 8.8% **increase** in lean ...

[www.bodybuildingforyou.com/pro-hormones/human-growth-hormone-hgh.htm](http://www.bodybuildingforyou.com/pro-hormones/human-growth-hormone-hgh.htm) - 33k - [Cached](#) - [Similar pages](#)

## [Buy HGH human growth hormone supplement, oral HGH hormone](#)

Somatrop Rx is a legal HGH **hormone** supplement with the strongest potency allowed by law. ... Now you **can** get actual **human growth hormones** legally without ...

[www.anabolic.com/ghk.htm](http://www.anabolic.com/ghk.htm) - 12k - [Cached](#) - [Similar pages](#)

Sponsored Links

### [Human Growth Hormone](#)

US Prescribed HGH and Testosterone  
Amazing results. Great service.

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

### [Human Growth Hormones](#)

Don't buy HGH until you read this

Get the latest research and reviews

Search results steroid drug HGH  
(human growth hormone)

### [HGH - The Truth](#)

What They Don't Want You To Know  
What It Does To Your Body Affiliate

[www.GenF20.com](http://www.GenF20.com)

### [Muscle Growth](#)

Want Faster **Muscle Growth**?

Learn The Tips & Secrets

[www.steroid-encyclopaedia.com](http://www.steroid-encyclopaedia.com)

### [Build muscle fast](#)

System used by 150,000+ worldwide

Put on 30+ pounds of pure **muscle**!

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

### [Muscle Growth Secrets](#)

# Any controversial issue will be targeted

The screenshot shows a Google search interface. At the top, the Google logo is on the left, and navigation links for Web, Images, Groups, News, Froogle, and more are on the right. A search bar contains the text "is adhd a real disease?". Below the search bar, a message states: "The following words are very common and were not included in your search: is a. [details]". The search results section is titled "Web" and shows "Results 1 - 10 of about 200,000 for is adhd a real disease?. (0.18 seconds)". A tip suggests asking researchers at Google Answers. The first result is titled "Texans for Safe Education Resolution" and discusses Dr. Fred Baughman's discovery of "real diseases". The second result is titled "Truths About ADHD and Stimulant Drugs" and states that ADHD is not a real disease. A red callout box highlights the text "Search results for mental disease ADHD (attention-deficit/hyperactivity disorder)". The third result is titled "To all website visitors: I urge one and all concerned with" and discusses the CHADD answer to the question "IS ADHD A REAL DISEASE—YES OR NO?". The fourth result is titled "ADD and ADHD Fraud. Find out the truth about ADD and ADHD. ..." and discusses the fraud of ADHD. The fifth result is titled "ADHD a Real Disorder" and discusses the controversy over whether ADHD is a real disorder. The sixth result is titled "Baughman Dispels The Myth of ADHD" and discusses Dr. Baughman's insight into the ADHD diagnosis.

Web Images Groups <sup>New!</sup> News Froogle more »

Google™ is adhd a real disease? Search Advanced Search Preferences

The following words are very common and were not included in your search: is a. [details]

**Web** Results 1 - 10 of about 200,000 for [is adhd a real disease?](#). (0.18 seconds)

Tip: Have a question? Ask the researchers at [Google Answers](#).

[Texans for Safe Education Resolution](#)  
Texans for Safe Education Is **ADHD A Real Disease?** Dr. Fred Baughman is a neurologist who has discovered **real** diseases. By directly ...  
[www.wildestcolts.com/safeEducation/real.html](#) - 8k - [Cached](#) - [Similar pages](#)

[Truths About ADHD and Stimulant Drugs](#)  
... Informative Websites. [www.adhdfraud.org](#) - Dr. Fred Baughman's excellent website, containing the best of his essays revealing that **ADHD** is not a **real disease**. ...  
[www.wildestcolts.com/mentalhealth/stimulants.html](#) - 14k - [Cached](#) - [Similar pages](#)  
[ [More results from www.wildestcolts.com](#) ]

[To all website visitors: I urge one and all concerned with](#)  
MAKE CHADD ANSWER: IS **ADHD A REAL DISEASE—YES OR NO?** ...  
Jr., MD, May 26, 2001. Is **ADHD a real disease—yes or no?** ...  
[www.adhdfraud.org/commentary/5-27-01-1.htm](#) - 5k - [Cached](#) - [Similar pages](#)

[ADD and ADHD Fraud. Find out the truth about ADD and ADHD. ...](#)  
... the fraud of Attention Deficit Hyperactivity Disorder (**ADHD**)—Compiled by ... Making "**disease**" (**real** diseases—epilepsy, brain tumor, multiple sclerosis, etc.) or ...  
[www.adhdfraud.org/](#) - 21k - Jan 29, 2005 - [Cached](#) - [Similar pages](#)  
[ [More results from www.adhdfraud.org](#) ]

[ADHD a Real Disorder](#)  
... Some people say, **ADHD** is not a **real** disorder because ... He says : "**ADHD** is a disorder that cannot be identified in the same way as polio, heart **disease** or other ...  
[web4health.info/en/answers/adhd-real-disorder.htm](#) - 20k - [Cached](#) - [Similar pages](#)

[Baughman Dispels The Myth of ADHD](#)  
... The "**disease**," Baughman tellsInsight, "is a total 100 percent ... to bring an end to the **ADHD** diagnosis ... as an adult and child neurologist, diagnosing **real** diseases. ...  
[www.becomehealthynow.com/article/dirty/209](#) - 62k - [Cached](#) - [Similar pages](#)

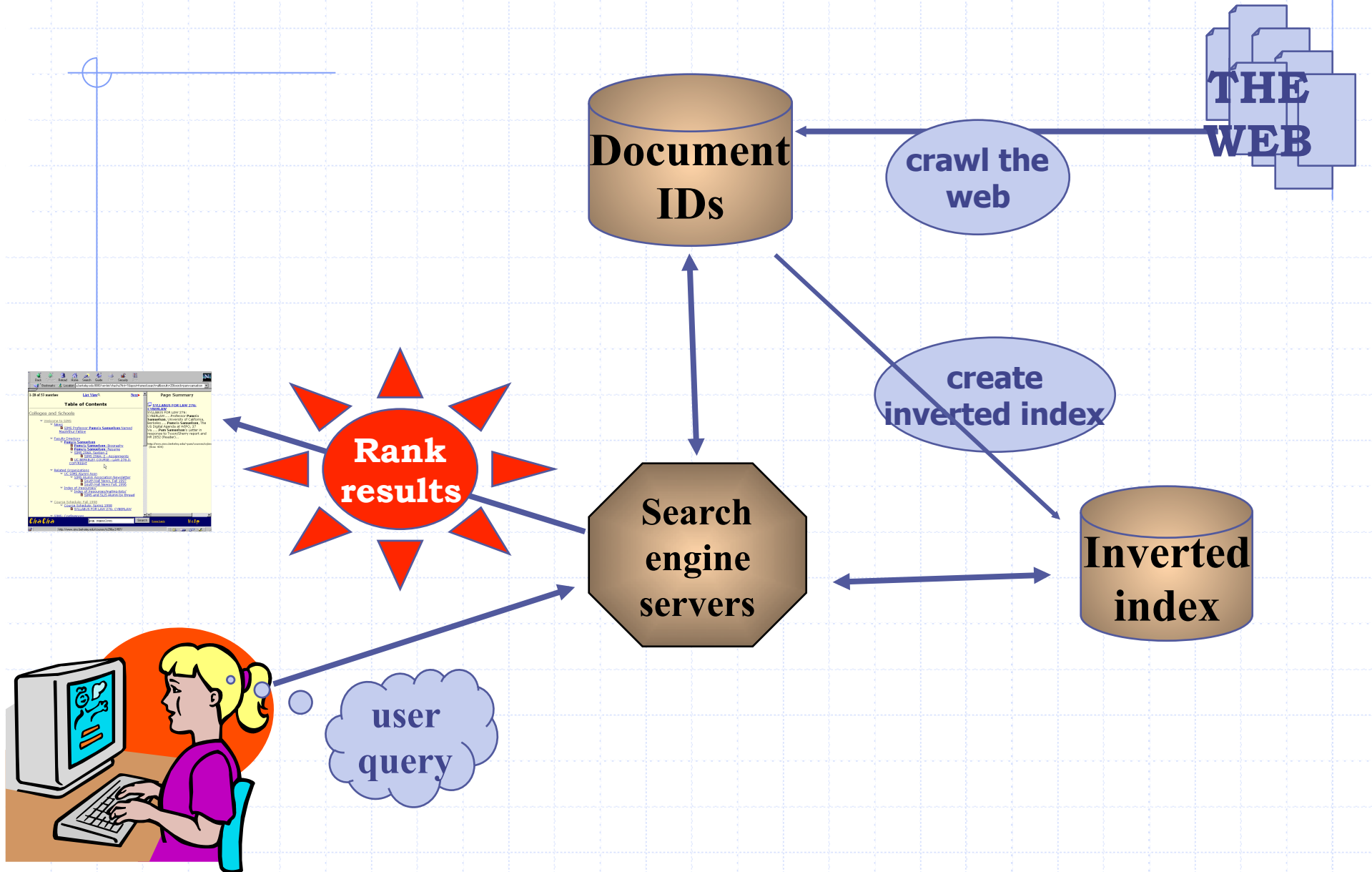
**Search results for mental disease ADHD (attention-deficit/hyperactivity disorder)**

# Why?



- ◆ If you do not **rank high** in search results,  
you may well not exist!
- ◆ **Web spam:**
  - Attempt to **modify** the web (its structure and contents),  
and thus **influence** search engine results  
in ways **beneficial** to web spammers

# Ranking is The Important Decision





# Evolution of Search Engines

## ◆ 1st Generation (ca 1994):

- AltaVista, Excite, Infoseek...
- Ranking based on **Content**:
  - ◆ Pure Information Retrieval

## ◆ 2nd Generation (ca 1996):

- Lycos
- Ranking based on **Content + Structure**
  - ◆ Site Popularity

## ◆ 3rd Generation (ca 1998):

- Google, Teoma, Yahoo
- Ranking based on **Content + Structure + Value**
  - ◆ Page Reputation

## ◆ 4<sup>th</sup> Generation? (In the Works)

- Ranking based on “the need behind the query”

# 1st Generation: Content Similarity

## ◆ *Content Similarity Ranking:*

*The more **rare** words two documents share, the more similar they are*

	Antony and Cleopatra	Julius Caesar	The Tempest	Hamlet	Othello	Macbeth
Antony	157	73	0	0	0	0
Brutus	4	157	0	1	0	0
Caesar	232	227	0	2	1	1
Calpurnia	0	10	0	0	0	0
Cleopatra	57	0	0	0	0	0
mercy	2	0	3	5	5	1
worser	2	0	1	1	1	0

- ◆ Documents are treated as “**bags of words**”  
(no effort to “understand” the contents)
- ◆ Term frequencies are computed  
(Above: a few rare words in W.S. books)
- ◆ Very common words (“a”, “of”, “the”, “in”, ...) are ignored

# 1st Generation: Content Similarity

## ◆ Content Similarity Ranking:

*The more **rare** words two documents share, the more similar they are*

	Antony and Cleopatra	Julius Caesar	The Tempest	Hamlet	Othello	Macbeth
Antony	13.1	11.4	0.0	0.0	0.0	0.0
Brutus	3.0	8.3	0.0	1.0	0.0	0.0
Caesar	2.3	2.3	0.0	0.5	0.3	0.3
Calpurnia	0.0	11.2	0.0	0.0	0.0	0.0
Cleopatra	17.7	0.0	0.0	0.0	0.0	0.0
mercy	0.5	0.0	0.7	0.9	0.9	0.3
worser	1.2	0.0	0.6	0.6	0.6	0.0

- Above: **Normalized** term frequencies.
- Results are **ranked** by  
sorting the sum of normalized frequencies of the query terms

Query: “In the mercy of Caesar”

Rank Results:

**How To Spam?**

# 1st Generation: How to spam

◆ **“Keyword stuffing”:**  
*Add keywords, text, to increase content similarity*

The screenshot displays the BMW website's navigation and content structure. At the top, there is a navigation bar with links for Home, Automobile, Produkte & Services, Dienstleistungen, Finanzierung, and Modelle. Below this is a secondary navigation bar with model identifiers: 1, 3, 5, 6, 7, X3, X5, Z4, M. The main content area features a large advertisement for the BMW 5er Limousine, titled "Neue Automobile" and "Automobile Leidenschaft." To the right of the main ad is a sidebar with several sections: "Beste Automobile direkt" (listing options like configuring vehicles, finding dealers, test drives, and financing), "Weitere Spezialis" (listing various BMW models like 1er, 3er, 5er, X3, etc.), "Gewisse Leistungen" (listing services like BMW Financial Services, BMW Service, etc.), and "BMW Deutschland" with a logo and "Partners am Platz".

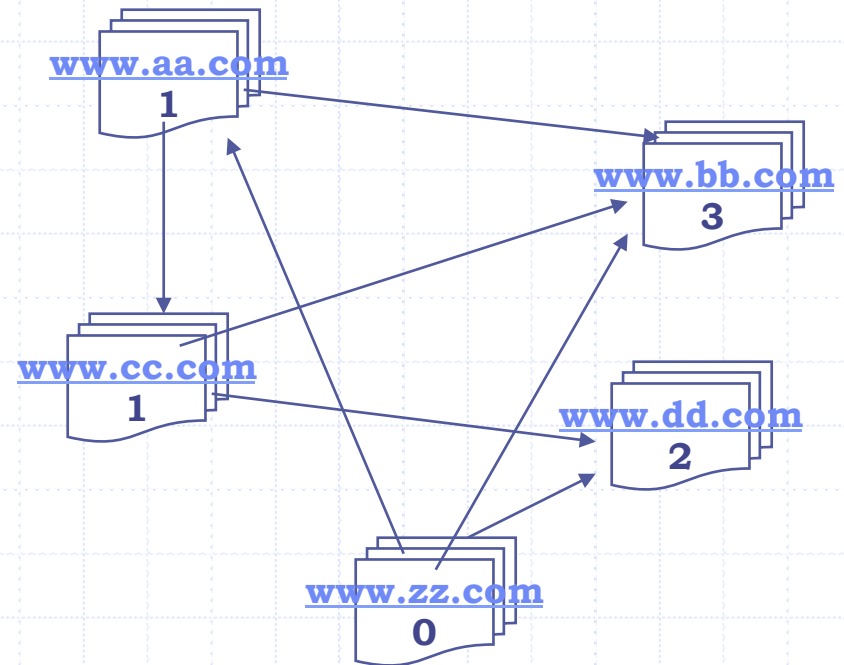
**Der BMW**  
**Zer BMW - Fahrspaß pur!**  
Autohaus BMW  
Sie suchen ein BMW Autohaus?  
Behörde Fahrzeuge Anschaffung  
Fahrzeugverkauf an Behörden  
Behörde Fahrzeuge Beschaffung  
Fahrzeugbeschaffung für Behörden bei BMW  
Beschaffung Fahrzeuge Behörde  
Beschaffung von  
Behördenfahrzeugen bei BMW  
Staatlichfahrzeug BMW  
Für Polizei, Notarzt und Feuerwehr -  
Blaulichtfahrzeuge von BMW  
Rechtsfahrzeuge Notarzt  
Schnell im Einsatzort - mit einem  
Notarzwagen von BMW  
BMW 318  
Informationen zum BMW 318  
BMW 318  
Informationen zum BMW 318  
BMW 318i  
BMW 318i - Design und Fahrkultur  
BMW 320  
Probefahrt mit einem BMW 320T  
BMW 330  
Finanzieren BMW 330  
BMW 3er  
BMW 3er Editionen  
BMW 5er  
BMW 5er Serie  
BMW 6er  
Das BMW 6er Coupé  
BMW Ausbildung  
Stellenangebote bei BMW  
BMW Autohändler  
BMW Autohändler Adressen

**BMW Neuwagen**  
Sie suchen einen BMW Neuwagen? Unsere Suche nach BMW Händlern in Ihrer Nähe bietet schnellen Zugriff auf BMW Autohäuser in Ihrer Nähe, wo Sie sich die BMW Neuwagen in aller Ruhe und Ausführlichkeit ansehen können. BMW Neuwagen - Sie erhalten von uns Adresse, Telefon und Website der BMW Händler in Ihrer Nähe. Suchen Sie über Postleitzahl, Stadt oder Name des BMW Partners. BMW Neuwagen - In jeder Abteilung unserer BMW Niederlassung arbeiten Fachleute für Sie. Unsere Mitarbeiter werden durch intensive Schulungen der BMW AG immer auf dem aktuellen Stand des Wissens gehalten. BMW Neuwagen - Mit diesem Know-how erarbeiten sie garantiert immer die Lösung, die sich am besten an Ihre Bedürfnisse anpasst. In einem unserer vielen Autohäuser in ganz Deutschland können Sie sich rasch und unproblematisch für eine Probefahrt in Ihrem Lieblings-BMW Neuwagen anmelden. Egal, worum es geht: um die Absprache eines Service-Termins, eine Reparatur oder die Finanzierung Ihres Neuen oder Ihres neuen Gebrauchten. BMW Neuwagen - Sympathisch und kompetent. Unser Team ist immer im Einsatz für Sie. Bei uns ist immer was los! Langeweile kommt in unserer Niederlassung nicht auf. Hier finden Sie Informationen zum Thema: BMW Neuwagen gesucht!

**Ein BMW Neuwagen gesucht?**  
BMW Neuwagen - Regelmäßig bieten wir Ihnen neue Angebote, bringen Ihnen Aktuelles über die neuesten BMW Modelle nahe und organisieren Veranstaltungen aller Art. Bei uns werden Sie gut informiert und gut unterhalten. BMW Neuwagen - Egal, welche Frage Sie an unsere Profis haben: In unseren Filialen sind Sie als unser Kunde oder als neugieriger Interessent immer herzlich willkommen. BMW Neuwagen - Wenn Ihnen das BMW Portal im Internet Appetit gemacht hat, besuchen Sie doch einmal eine Niederlassung ganz in Ihrer Nähe - hier können Sie unseren Mitarbeitern Lächer in den Bauch fragen. BMW Neuwagen - Haben Sie Interesse an einer Probefahrt in Ihrem Wunsch-BMW? Kein Problem! Bei unseren Niederlassungen können Sie sich jederzeit für eine Probefahrt in einem unserer Automobile anmelden. In unseren Niederlassungen bekommen Sie einen Vorgeschmack auf die Freude am Fahren. BMW Neuwagen - Der Hof- und Bring-Service. Nutzen Sie die Flexibilität Ihres BMW Partners. So verlieren Sie keine unnötige Zeit. Viele BMW Autohäuser sehen nicht nur die Wartungs- und Reparaturarbeiten an ihrem BMW fachgerecht aus. Sie holen das Fahrzeug auch direkt bei Ihnen zu Hause oder im Büro ab und bringen es Ihnen nach Beendigung der Arbeiten wieder zurück.

# 2nd Generation: Add Popularity

- ◆ A hyperlink from a page in site A to some page in site B is considered a **popularity vote** from site A to site B
- ◆ Rank **similar** documents according to popularity
- ◆ For example, if these sites had the same content similarity, the highest ranking result would be:

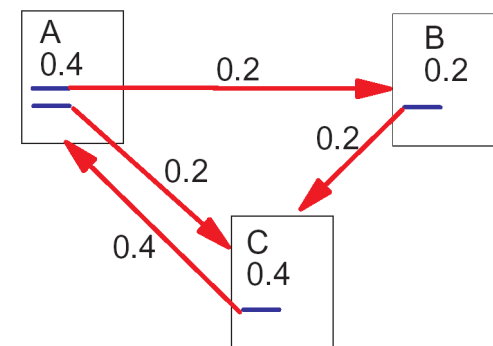


**How To Spam?**



# 3rd Generation: Add Reputation...

- ◆ The **reputation** “PageRank” of a page  $P_i$  = the sum of a fraction of the reputations of all pages  $P_j$  that point to  $P_i$
- ◆ Idea similar to academic **co-citations**
- ◆ Beautiful Math behind it
- ◆ The PageRank of a page
  - Can be calculated by a sequence of matrix multiplications
  - It is equivalent to the probability of randomly surfing to a page



Simplified PageRank Calculation

**How To Spam?**

# 3rd Generation: How to spam

## ◆ *Organize Mutual Admiration Societies: “link farms” of irrelevant reputable sites*

### Resource Partner Additions

Currently we are only adding websites to this Resource page with PR6 or higher AND WHOSE LINKING PAGE IS PR5 OR HIGHER. The first thing we check when we receive a link request is the PageRank. If the page you are planning to put our link on does not have a PR5 or higher we delete the request without a response. Please honor this request. Those websites that meet the above requirements are added within a few days.

**To add your website to our Directory simply do 3 things:**

**1)** Copy this code to your website:

<!-- Start Copying here -->

```
<P><A HREF="http://www.1st-Hgh.com" TARGET="_blank"><B>1st Hgh, Homeopathic Human Growth Hormone Spray</B></A><B></B>- The primary hormone in the body, information and sales. Hgh is the ultimate youth reviving hormone. For everyone over 30 years old. This product is safe, inexpensive, and comes with a full money back guarantee.</P>
```

<!-- End Copying here -->

It should look like this:

[1st Hgh, Homeopathic Human Growth Hormone Spray](http://www.1st-Hgh.com) - The primary hormone in the body, information and sales. Hgh is the ultimate youth reviving hormone. For everyone over 30 years old. This product is safe, inexpensive, and comes with a full money back guarantee.

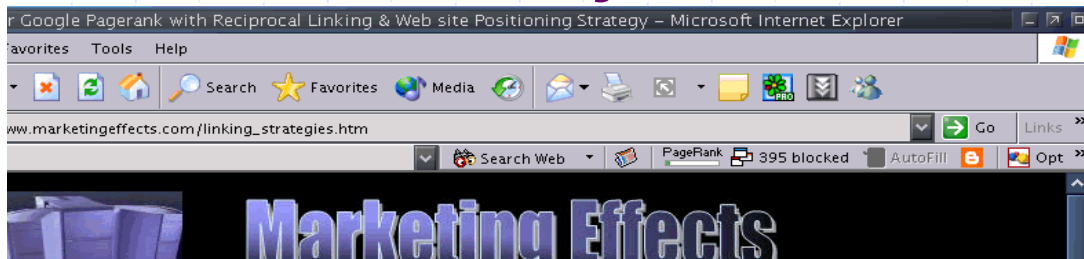
**2)** [E-Mail us](#) the url where we can view our link on your website (it must be PR4 or higher!);

**3)** Include in this e-mail your website url, title, and description.


Thank You, Partner!



# An Industry is Born



- ◆ “Search Engine Optimization”
- ◆ Advertisement Consultants
- ◆ Conferences



## Reciprocal Link Tools and Website Positioning Strategies

**Drive Thousands of Targeted Visitors to your website and boost your Google Pagerank**

### Services

- [Linking Strategies](#)
- [Search Engine Optimization](#)
- [Professional Internet Marketing Services](#)
- [Turnkey Internet Businesses for Sale](#)
- [Expired Domain Names](#)
- [Free Newsletter](#)
- [Hosting](#)
- [88 Domain Registration](#)
- [Resources](#)
- [FAQ](#)
- [Contact Us](#)

### Why Having Links to Your Website Are Important

Having links from other websites is very important, yet, it's one of the most overlooked strategies for increasing traffic and for increasing your position in the search engines.

**As soon as you launch a website, you should spend some time working on a strategic web site linking strategy.**

By working on getting sites to link to you, you'll not only see an increase in traffic, but you'll also see your website achieve better listings in the search engines, especially the most important search engine - Google.

### How to find good sites to link to you

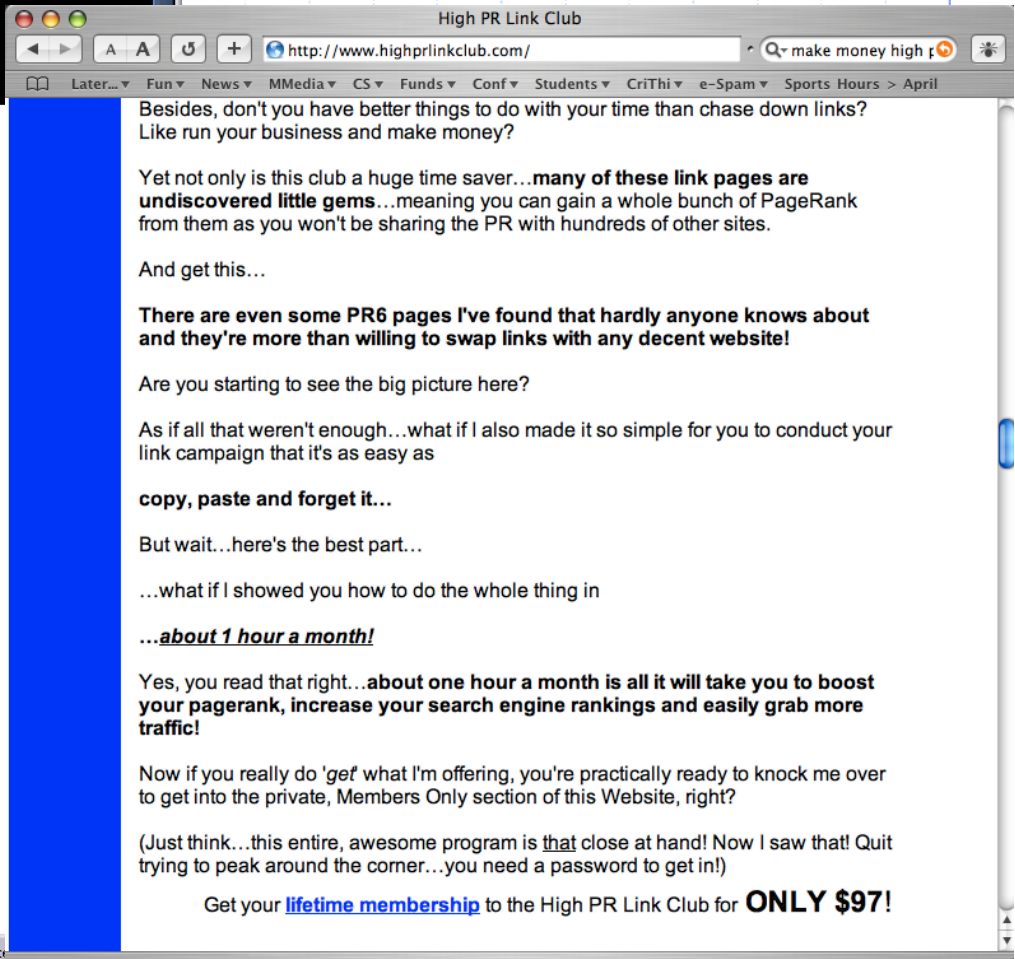
It's all about getting the right sites to link to you. Just remember **quality before quantity**. That means to avoid links farms and FFAs like the plague. Many search engines will penalize you for having a link from a link farm. Ask yourself: "Would I want people to associate me with this website?" If the answer is no, then move on to the next site. With over 40,000,000 websites, you should be able to find some high quality websites willing to link to you.

Before you get started, I recommend these two tools, as they'll make finding the quality sites you want to be linked to much easier.

The first is the [Google Toolbar](#). With the Google Toolbar, you'll be able to see how important Google thinks the page is.

The second is [Alexa Toolbar](#). With Alexa, you can see how popular the site is. Alexa ranks every site based on how often other toolbar users visit the site. It's an excellent tool, and the best part about both of these tools are that they're absolutely FREE.

The next step is to get on your favorite search engine (I recommend Google since it's



## High PR Link Club

http://www.highprlinkclub.com/

Besides, don't you have better things to do with your time than chase down links? Like run your business and make money?

Yet not only is this club a huge time saver...many of these link pages are **undiscovered little gems**...meaning you can gain a whole bunch of PageRank from them as you won't be sharing the PR with hundreds of other sites.

And get this...

**There are even some PR6 pages I've found that hardly anyone knows about and they're more than willing to swap links with any decent website!**

Are you starting to see the big picture here?

As if all that weren't enough...what if I also made it so simple for you to conduct your link campaign that it's as easy as

**copy, paste and forget it...**

But wait...here's the best part...

...what if I showed you how to do the whole thing in

**...about 1 hour a month!**

Yes, you read that right...**about one hour a month is all it will take you to boost your pagerank, increase your search engine rankings and easily grab more traffic!**

Now if you really do 'get' what I'm offering, you're practically ready to knock me over to get into the private, Members Only section of this Website, right?

(Just think...this entire, awesome program is that close at hand! Now I saw that! Quit trying to peak around the corner...you need a password to get in!)

Get your [lifetime membership](#) to the High PR Link Club for **ONLY \$97!**

## CURRICULUM VITAE

### About me

I was born July 23, 1977 in Albany, NY. I recieved a BA from Reed College in biochemistry and molecular biology in 1999, an MS from Cornell University in biophysics in 2002, and an MS from Stanford University in computational and mathematical engineering in 2009. I'm currently working at Genia Technologies Inc., analyzing single-molecule measurements of DNA threaded through protein nanopores embedded in phospholipid bilayers. I also have a part-time position in the Stanford University Radiology Department with the Informatics in Imaging Sciences (ISIS) group, developing the [BIMM](#) biomedical image metadata manager and [ATM](#) Aim Template Manager applications (pre-release version only).

# Link exchanges

The image shows three overlapping browser windows. The leftmost window is titled 'Buy Links Buy Link Popularity - Micro...' and features a large graphic with the text 'The Link Monger' and 'AFFORDABLE RELEVANT L...'. Below the graphic are navigation links for Home, Contact us, and Site map. The middle window is titled 'Link Exchange Directory: MEDICINE & PHAR...' and contains a search bar, a site map with links to Home, Link Exchange Directory, Submit Your Site, SEO Resources, and Contact Us, and sections for 'Featured Articles' and 'Partners'. The rightmost window is titled 'Lifestyle Link Exchange - Microsoft Internet Explorer' and shows the homepage of 'Links-Pal.com'. The page header includes the site name, a tagline 'IMPROVE YOUR LINK POPULARITY!', and a navigation menu. The main content area is titled 'LINK EXCHANGE PARTNERS - RECIPROCAL LINK EXCHANGE' and 'LIFESTYLE'. Below this is a table of 'HIGHLY RECOMMENDED LINK EXCHANGE PARTNERS' with columns for partner name and description.

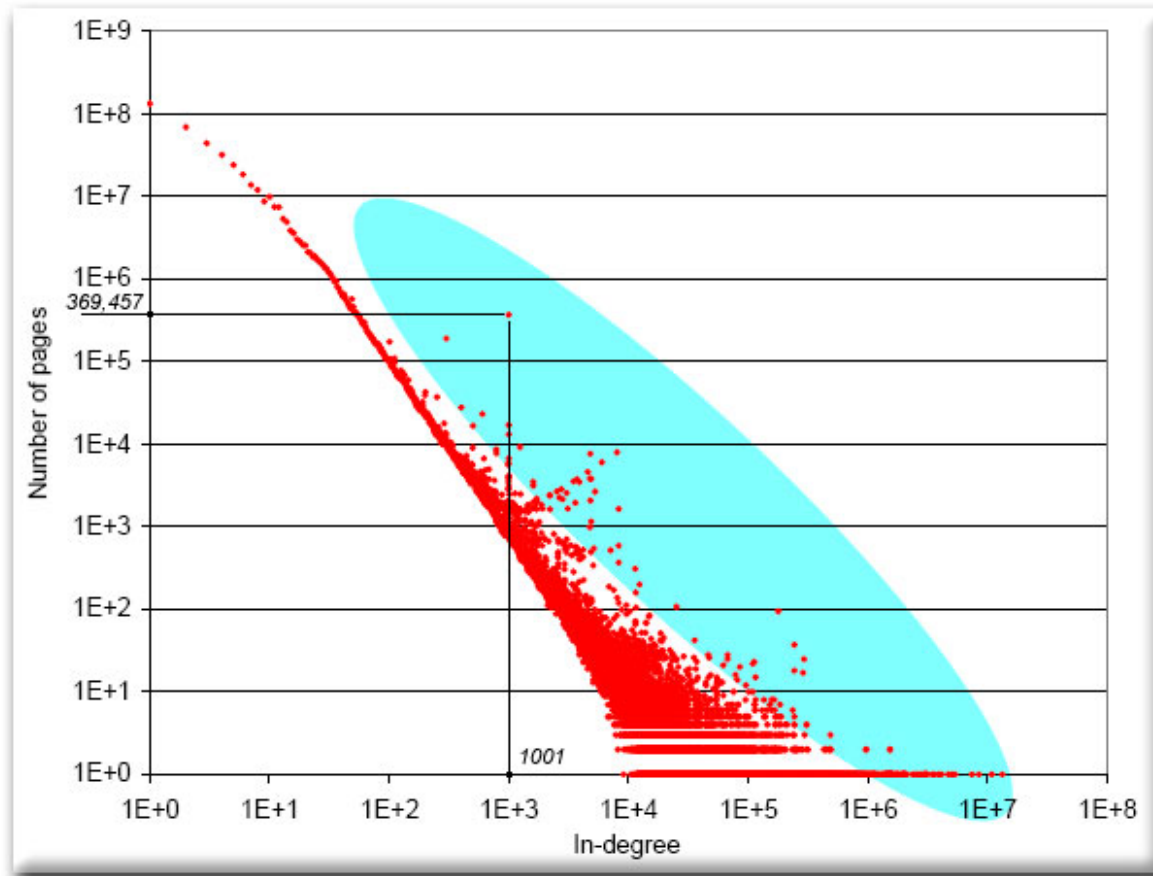
:: HIGHLY RECOMMENDED LINK EXCHANGE PARTNERS ::	
<b>HOODIA GORDONII AN ALL NATURAL APPETITE SUPPRESSANT</b>	Hoodia Gordonii is the all natural appetite suppressant that has been featured on CBS News 60 Minutes. Visit Hoodia-Advice.Org to read about the science and clinical studies that have been done on Hoodia. Learn if Hoodia can help you lose weight.
<b>THE FLOWER ARRANGEMENT ADVISOR</b>	Offers varieties of flower arrangements and advice on how to choose flower arrangements for special occasions like Christmas, Thanksgiving, etc. Also contain creative home decorating ideas using flower arrangements as well as steps on some very easy-to-do flower arrangements.
<b>THE SPORTS NEWSPAPER</b>	We are an Online sports newspaper, providing Handicapped free picks, Fantasy sports, Real Time Scores and odds, stats, trends, forums, game matchups, injuries and much much more!
<b>TULSA REAL ESTATE</b>	Tulsa Area Real Estate, Homes, Search all Tulsa Area Listings.
<b>NARCISSISM AND NARCISSISTIC PERSONALITY DISORDER BOOK</b>	The book sold on this site will change your life forever.
<b>KITCHEN SUPPLIES</b>	Cookware, cutlery, coffee makers, dinnerware, food processors, blenders, vacuums and more.
<b>GIFT IDEAS FOR ROMANCE</b>	Gift Ideas for Romance Romantic Gifts for Love, Friendship and Passion. 24K Gold Roses are featured along with many unique gifts and gift ideas to romance the heart.
<b>A PLACE TO MEET BEAUTIFUL RUSSIAN WOMEN &amp; UKRAINIAN WOMEN</b>	FSM helps you to meet beautiful Ukrainian girls & Russian girls for love, romance, friendship and more. Check our catalog of beautiful ladies. They are waiting to hear from you, just take that first step.
<b>WHOHITS.COM - FREE WEB TRACKER AND TRAFFIC ANALYZER</b>	Sign up free for a complete web tracker, web traffic analyzer and hit counter and track your website traffic.
<b>CATCH A CHEATING WIFE</b>	Cheating wife. Find out Now Is she Cheating. Diet is everything, extramarital affair. Save your Relationship. Married but looking Read about Cheating Wife Stories.
<b>FUTON PLANET</b>	Futons, Futon Covers, and Other Contemporary Furniture for the Home and Office.
<b>ROMANTIC HORIZONS</b>	Romantic gifts for love, romance, intimacy and passion. Bring love and romance to new heights with a wide assortment of gifts that both please and thrill the senses.
<b>NEW YORK SEARCH ENGINE OPTIMIZATION COMPANY</b>	Are you tired of poor rankings and bad web design? Let our team of Web Design Experts design you a great site that is XHTML / CSS compliant, and then let our Search Engine Specialists, rank your site as high as possible in all major search engines.
<b>SEXY ADULT COSTUMES</b>	Find your adult costume this holiday season and scare the sexy out of 'em. Choose from our super collection of sexy adult costumes, and sexy shoes to match!

# Web forum and blog spam

The image shows a screenshot of a Microsoft Internet Explorer browser window. The address bar displays the URL: <http://www.margieandaustinswedding.com/wwwboard/messages/345.html>. The main content area features a large, bold title: "Creative ideas for valentine's day gifts and christening gift idea including holiday office party ideas". Below the title are two empty text input fields and a set of links: "[ Follow Ups ] [ Post Followup ] [ Wedding Forum for Comments ]". The post is attributed to "Holiday office party ideas and Creative Ideas for Valentine's Day gifts" and is dated August 29, 2005 at 01:26:26. The body of the post is a dense, repetitive list of links and keywords, including "Wedding decorations ideas", "christening gift idea", "wedding centerpiece ideas", "Bedroom Design Ideas", "free scrapbook ideas", "kids birthday party ideas", "Gift ideas for mom", "Girl Slumber Party Ideas", "scavenger hunt ideas", "special event ideas", "Halloween Costume Ideas", "scrapbook layout ideas", "Romantic gift ideas", "work at home home business idea", "gift ideas for cat lovers", "Girlfriend Gift Ideas", "Deck ideas", "Creative Ideas for Valentine's Day gifts", "gift idea unique ideas for wedding reception", "bachelorette ideas", "business idea", "decoration halloween idea party", "christening gift idea", "Free Tattoo Ideas", "wedding planning idea", "wedding invitation ideas", "Baby Shower Cake Ideas", "graduation gift ideas", "Recipe ideas", "christmas gift idea", "Tattoo Design Ideas", "Girl Slumber Party Ideas", "design ideas", "gourmet gift basket ideas", "vacation ideas", "home decor ideas", "Home Decorating Ideas", "Cake Ideas", "cost saving ideas", "bridal shower ideas", "outdoor advertising idea", "arts and crafts ideas", "kitchen design ideas", "over the hill party ideas", "home based business idea", "Back to School Bulletin Board Ideas", "Fireplace Design Ideas", "proposal ideas", "free home decorating idea", "business ideas for home Party Ideas", "The romantic gift idea of these porcelain rose nightlights", "unique event ideas", "fund raiser ideas", "fathers day ideas", "home decoration ideas", "free craft ideas", "Ideas for Company Name", "bridal shower gift ideas", "home business ideas", "marriage proposal ideas", "baby nursery ideas", "talent show ideas", "home decorating ideas", "patio ideas", "social event ideas", "Ideas for Decorating Cupcakes", "Birthday", "Birthday Gift Ideas", "great gift ideas", "Ideas for Self Employment", "christmas gift ideas", "marketing idea", "Business Ideas", "family reunion ideas", "team building ideas", "Baby Shower Ideas", "wedding shower ideas", "party ideas", "Decorating Ideas", "bachelorette ideas", "fall home decorating ideas", "baby shower ideas".

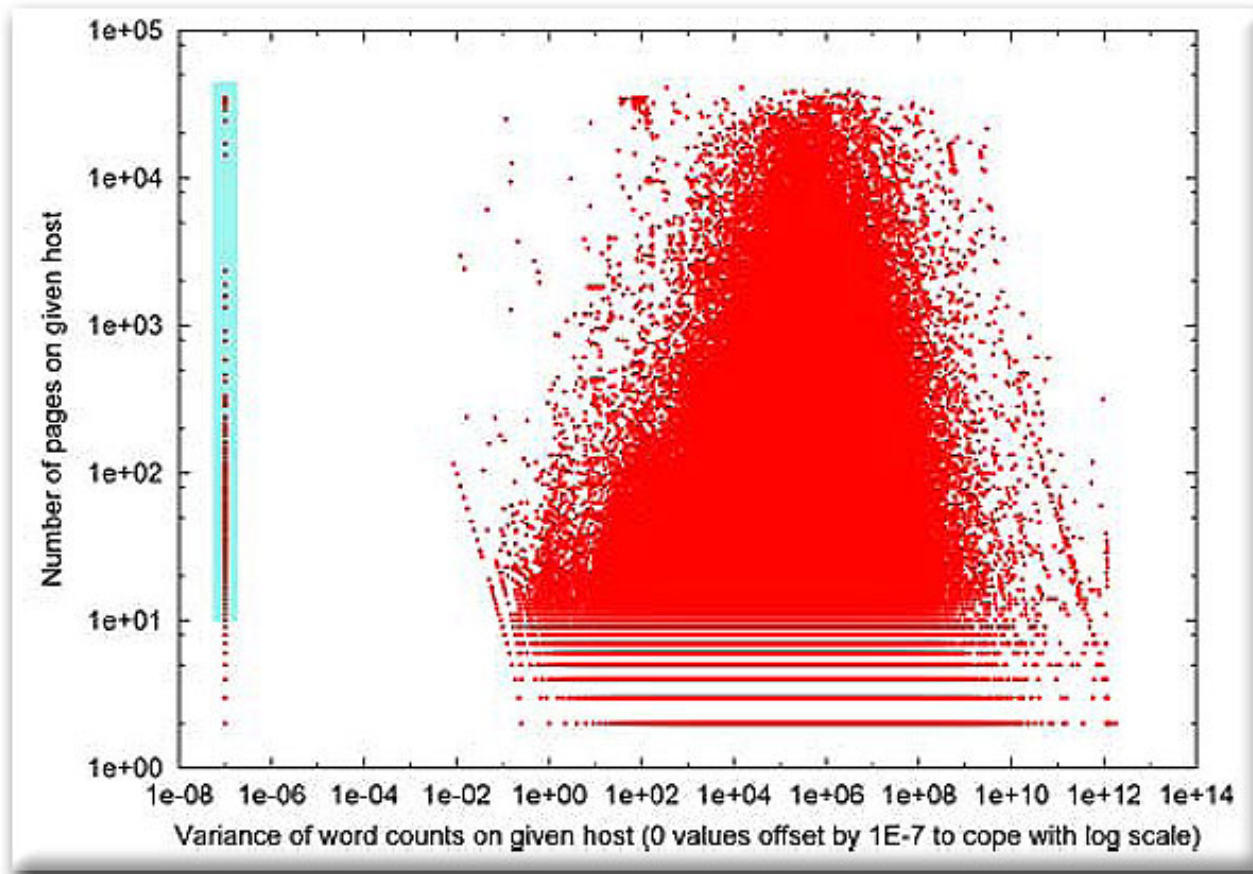
On the left side of the browser window, there is a sidebar for a website titled "Dr. B.'s Blog of classroom activities". It includes a "User login" section with fields for "Username:" and "Password:" and a "Log in" button. Below this are links for "Create new account" and "Request new password". There is also a "Links" section with several links, including "Archived blog posts", "Academic page", "iTunes Playlist", "Reading Marathon Pictures", "C&W 2004 Pictures", and "WPA 2005: Alaska Pictures". A "Who's online" section indicates "There are currently 1 user and 20 guests online." and lists "Online users: dr. b.". A "Navigation" section includes links for "Archives", "books", "content", "search", "categories", and "syndication". At the bottom of the sidebar, there is a "Buttons and Such" section with a link to "Technorati Profile".

# Spam Detection through Statistics



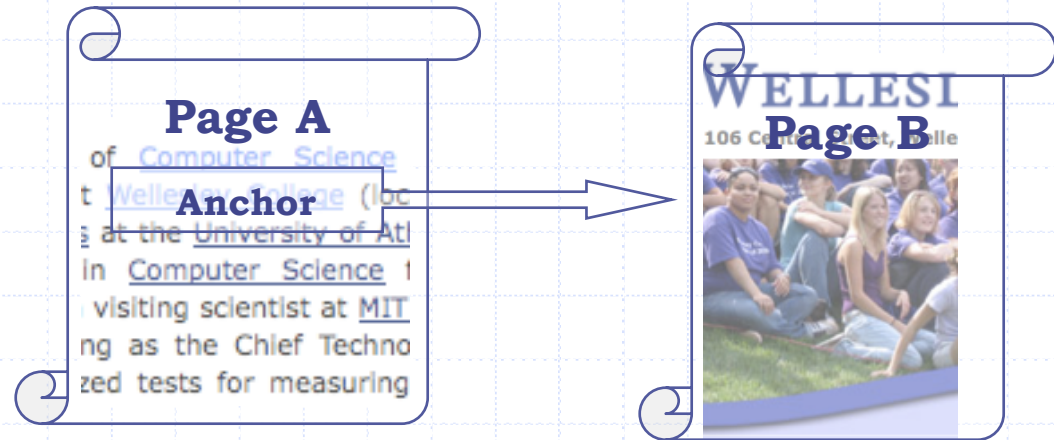
Web page in-degrees follows power law

# Word count variance



# 3rd Generation: Reputation & Anchor Text

- ◆ Anchor text tells you what the reputation is about
- ◆ It adds synonyms for a term



Founded in 1875,  
Wellesley College is known...

US News Best Liberal Arts Colleges:  
Williams College  
Wellesley College  
Pomona College

[www.wellesley.edu](http://www.wellesley.edu)

The contributions of the nation's  
women's colleges, in the education...

**How To Spam?**

# “Google-bombs” spam Anchor Text...



Web Images Groups News Froogle<sup>New!</sup> more »

miserable failure

Search

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

Web

Results 1 - 10 of about 255,000 for [miserable failure](#).

## [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 ...

[www.whitehouse.gov/president/gwbbio.html](http://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"

[www.whitehouse.gov/history/presidents/jc39.html](http://www.whitehouse.gov/history/presidents/jc39.html) - 35k - [Cached](#) - [Similar pages](#)

## [Michael Moore.com](#)

Click Here To Continue To MichaelMoore.Com "The War President". The  
War President Photo courtesy of American Leftist, (click here ...

[www.michaelmoore.com/](http://www.michaelmoore.com/) - 5k - [Cached](#) - [Similar pages](#)

## [Senator Hillary Rodham Clinton: Onli](#)

Dear Friend,. Thank you for visiting my on-line  
your interest in the issues before the United States

[clinton.senate.gov/](http://clinton.senate.gov/) - 9k - Apr 17, 2004 - [Cached](#)

Political Google-bombs

“miserable failure hits  
Obama in January 2009

Web Images Video Local Shopping more ▾

miserable failure

Search

Options ▾

Customize ▾

YAHOO!

1 - 10 of 1,790,000 for miserable failure (About) - 0.03 s | [SearchScan](#)<sup>NET</sup>

Also try: [miserable failure google](#), [miserable failure not bush](#), [More...](#)

## [President Barack Obama](#)

WhiteHouse.gov is the official web site for the White House and President Barack ... Barack  
H. Obama is the 44th President of the United States. ...

[www.whitehouse.gov/administration/president\\_obama](http://www.whitehouse.gov/administration/president_obama) - [Cached](#)

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

Features a biography of the 43rd president, President George W. Bush, from the  
White House website.

[www.whitehouse.gov/president/gwbbio.html](http://www.whitehouse.gov/president/gwbbio.html) - [Cached](#)



# ... for fun and profit!

- ◆ Business weapons
  - “more evil than satan”
- ◆ Political weapon in pre-election season
  - “miserable failure”
  - “waffles”
  - “Clay Shaw” (+ 50 Republicans)
- ◆ Misinformation
  - Promote steroids
  - Discredit AD/HD research
- ◆ Activism / online protest
  - “Egypt”
  - “Jew”
- ◆ Other uses we do not know?
  - “views expressed by the sites in your results are not in any way endorsed by Google...”

# *There is a pattern on How to Spam!*

- ◆ Web spammers are employing propagandistic techniques in order to modify the Trust Network of Search Engines

S.E.'s	Ranking	Spamming	Propaganda
1 <sup>st</sup> Gen	Doc Similarity	Keyword Stuffing	Word Games
2 <sup>nd</sup> Gen	+ Site Popularity	+ Link Farms	+ Bandwagon
3 <sup>rd</sup> Gen	+ Page Reputation	+ Mutual Admiration Societies	+ Testimonials
	+ Anchor Text	+ Link Bombs	+ Card Stacking
4 <sup>th</sup> Gen	+ "user need behind search"		

# 4<sup>th</sup> Generation's great idea?

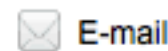
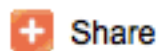
## Google to pay US \$500m in drug ad inquiry

Admits aiding Canadian sales

August 25, 2011 | By Hiawatha Bray, Globe Staff



2



In one of the biggest settlements of its kind in US history, Google Inc. has agreed to pay \$500 million to the Justice Department and admitted that the company's advertising service helped online pharmacies sell prescription drugs illegally.

US attorney for Rhode Island Peter F. Neronha revealed the agreement with the Internet search giant at a press conference in Providence yesterday. Neronha's office led an investigation of Google's role in the online pharmacy ads that began in 2009 and involved the Food and Drug Administration and the Internal Revenue Service, as well as Rhode Island authorities.

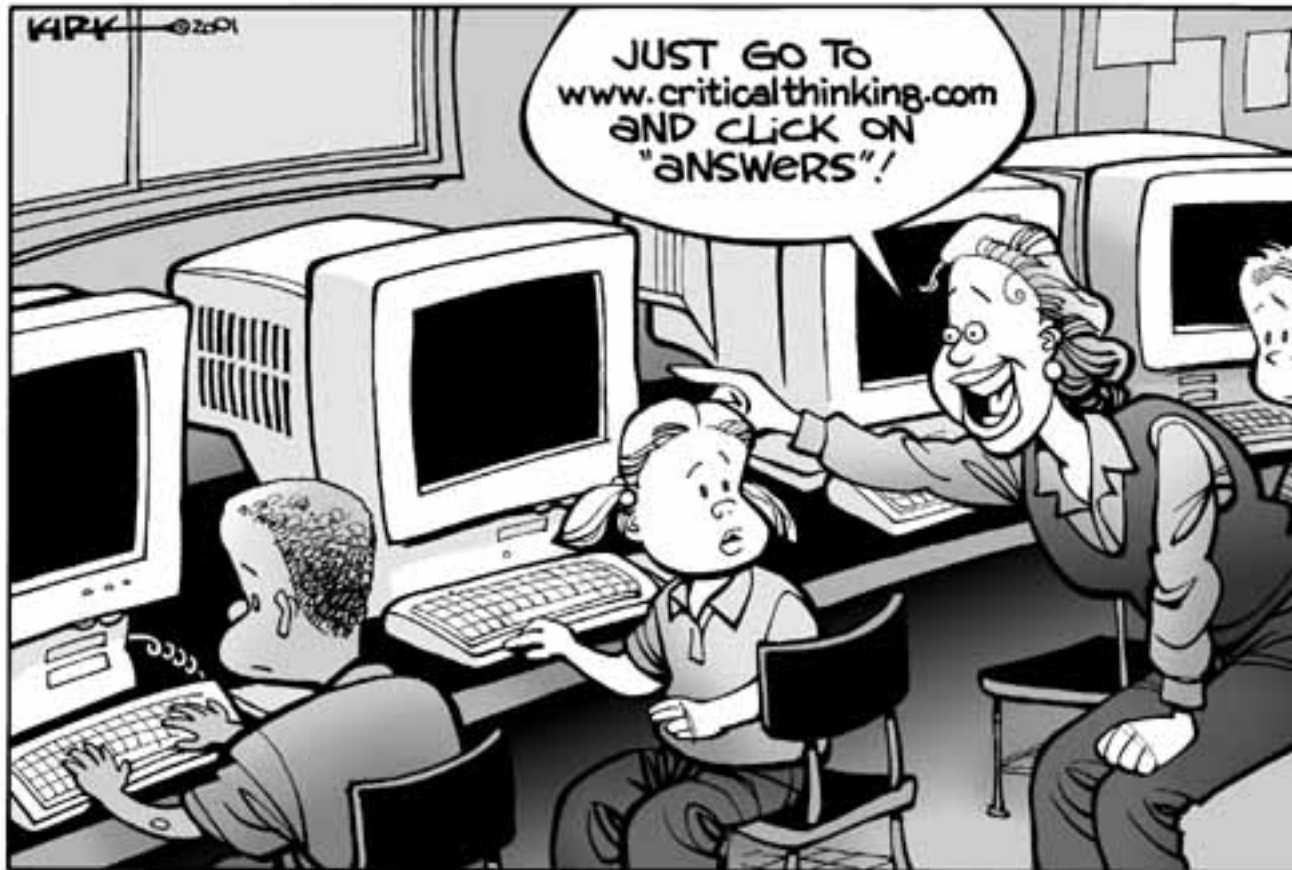
Ads by Google

### [Free Prescription Savings](#)

Save up to 75% on your prescription Just print our 56,000 discounts

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

# So, how do you decide what is reliable?

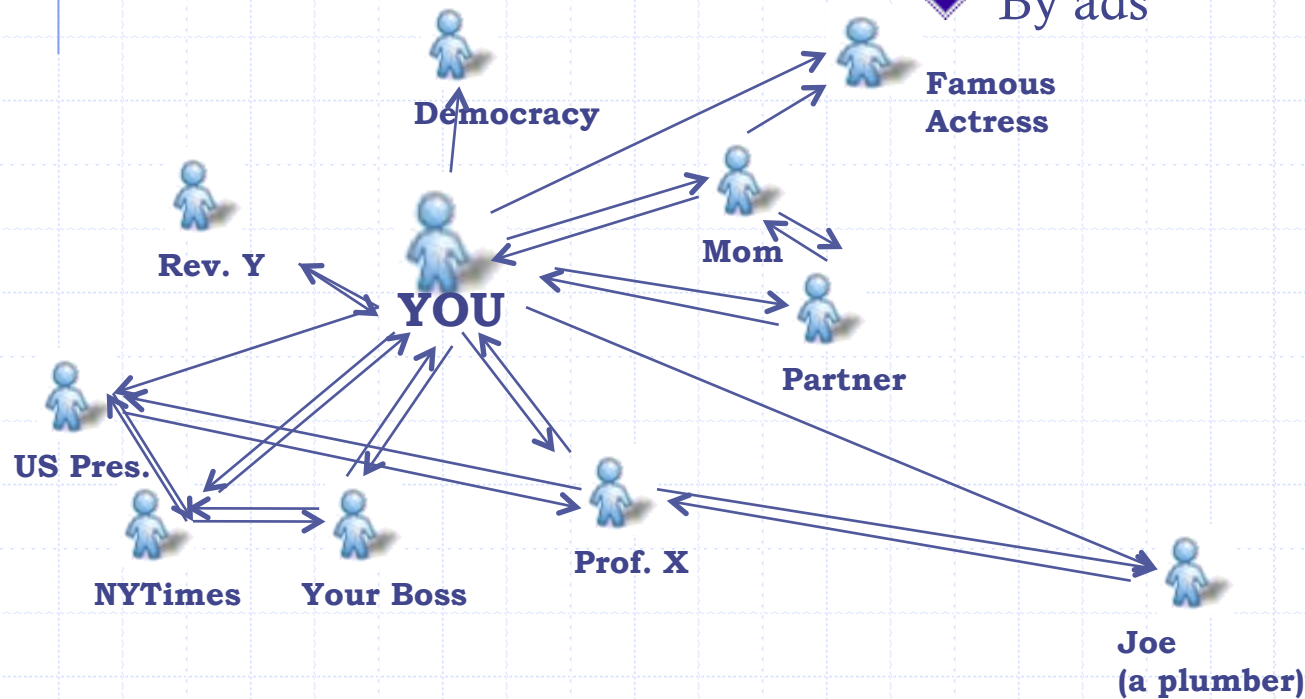


# CHALLENGES to your Trust Network

**By your own abilities**

**Devastation** ↑

- ◆ By your friends and family
- ◆ By teachers and colleagues
- ◆ By trusted advisors
- ◆ By trusted sources
- ◆ By others
- ◆ By ads



# How do you know what you know?

## ◆ Extrinsic reasons

- Trust in the entity supporting the information
- The majority of people use it extensively
- Technology can help here

## ◆ Intrinsic reasons

- Own experience
- Own ability to **think critically**,  
which means:
  - **Understanding the Scientific Method  
and apply it habitually on important matters**

## ◆ But this is tough and requires Education

# We also “know” ...

- ◆ What we learned as children
- ◆ What we think we remember correctly
  - There is no database in our brain, but we recreate memories every time we remember them
- ◆ What we misunderstood
  - Our brain is a pattern matching machine, we find similarities even where there is none
- ◆ What we think under the influence of substances, of voices in sleep, lack of sleep
- ◆ What we decide under fear, anger, personal interest, using illogical processes
- ◆ \_\_\_\_\_ (Add your own examples here)

# Are we so stupid?

- ◆ Our brain is impressively complicated, but it is not perfect
- ◆ It is influenced by construction limitations and errors, our feelings, our senses, our environment
- ◆ It was created through an ongoing evolutionary process. Some parts are old and are activated immediately, others are newer and demand great energy to get activated.
- ◆ We need to feel that we are in control of our environment. We do not easily accepting random phenomena, we want to discover a couple reasons explaining them.
- ◆ Critical thinking uses neocortex, large in size and requiring lots of energy to operate. We try to avoid it by creating heuristics, stereotypes, personal ways of “thinking”.



# The take-home message

- ◆ **Web misinformation** (like propaganda and advertisement) **is here to stay**
- ◆ **Finding** information is easy, **evaluating** it is more difficult than it used to
  - Our trusted sources are being challenged, fooled
- ◆ **Think critically** in terms of your trust network
  - Understand what you believe and why
- ◆ In the 21<sup>st</sup> century the basic survival skills
  - Reading & Writing
  - **Thinking critically**
  - Understand technology

# Thank you!

## 2012

Knowledge-Based Education – We oppose the teaching of Higher Order Thinking Skills (HOTS) (values clarification), critical thinking skills and similar programs that are simply a relabeling of Outcome-Based Education (OBE) (mastery learning) which focus on behavior modification and have the purpose of challenging the student's fixed beliefs and undermining parental authority.

**Report of  
Platform Committee**