

Web Search Basics

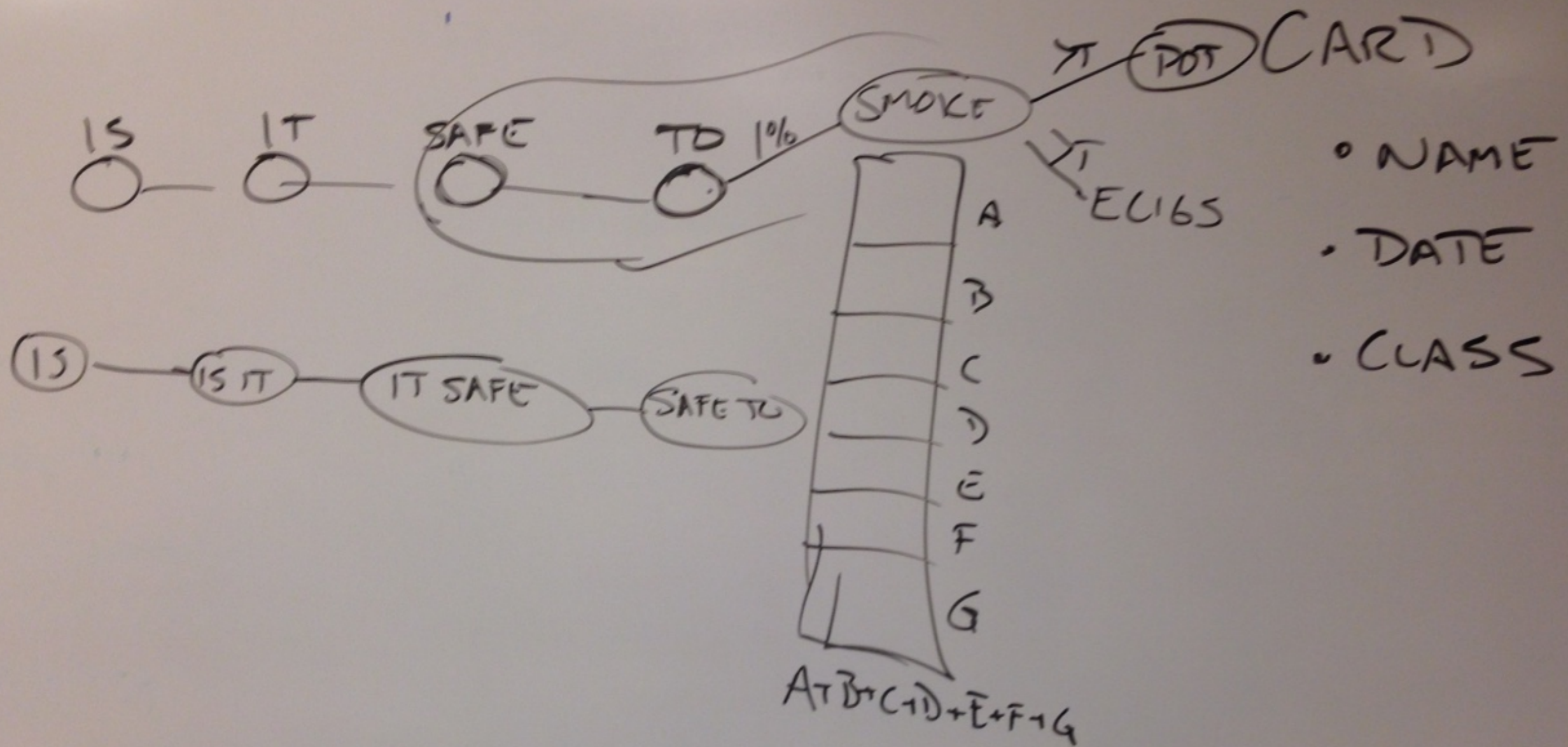
Introduction to Information Retrieval

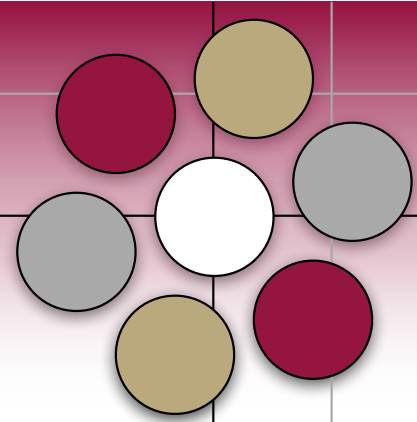
Content adapted from Hinrich Schütze
<http://www.informationretrieval.org>



Top queries

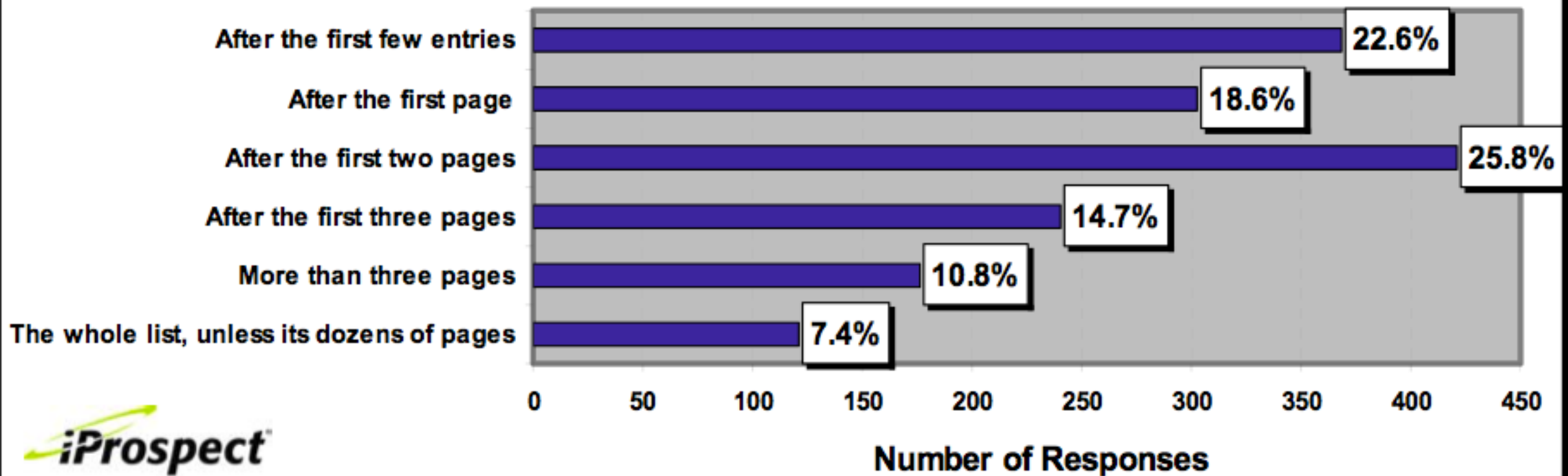
- Live demo - WARNING this is not very safe....
 - “Is it safe to”
 - “Is it legal to”
 - “why does”
 - “why doesn’t
 - “why is there”
 - “why isn’t”
 - “americans are”

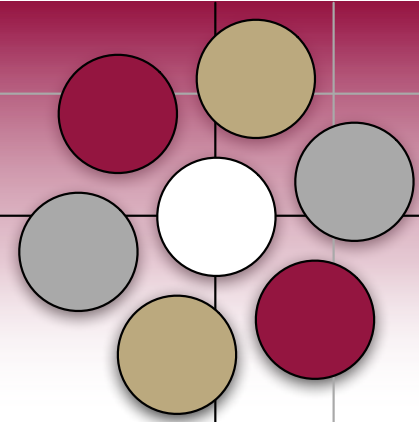




How far do people look for results?

If you don't find what you are looking for, at what point do you move on either to another search engine or to another search on the same engine?





True Example *

The User
flickr:crankyT



Task

Stop the noisy fan in the courtyard

Info Need

Info about EPA regulations

Verbal Form

What are EPA rules on noise pollution?

Query

EPA Sound Pollution

A screenshot of a search bar interface. It features a text input field with the word "search" in blue. Below the input field are two buttons: "Go" and "Search".

"To Google or to GoTo" Business Week Online 9/28/2001

SOURCE
QUESTION

BIG
6AD

MEANING OF
INFORMATION

QUERY

WISDOM

KNOWLEDGE

INFORMATION

FACTS
DAT

How do we help people move from data to wisdom?



How do users evaluate search engines?

- Quality of pages
 - Classic IR relevance
 - Also important:
 - Trust
 - Duplicate elimination
 - Readability
 - Fast Access
 - No pop-ups



How do users evaluate search engines?

- Precision is more important than recall
 - Precision:
 - How precise is a portal in locating relevant results?
- Recall
 - How thorough is the coverage of available relevant results?
- Precision with 1 result, 10 results, 2-3 pages of results.
- When is recall important?



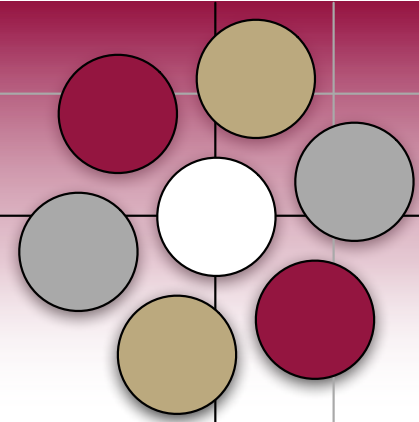
How do users evaluate search engines?

- Recall is sometimes important:
 - Googling for a new doctor
 - Googling a prospective employee
 - Googling your date

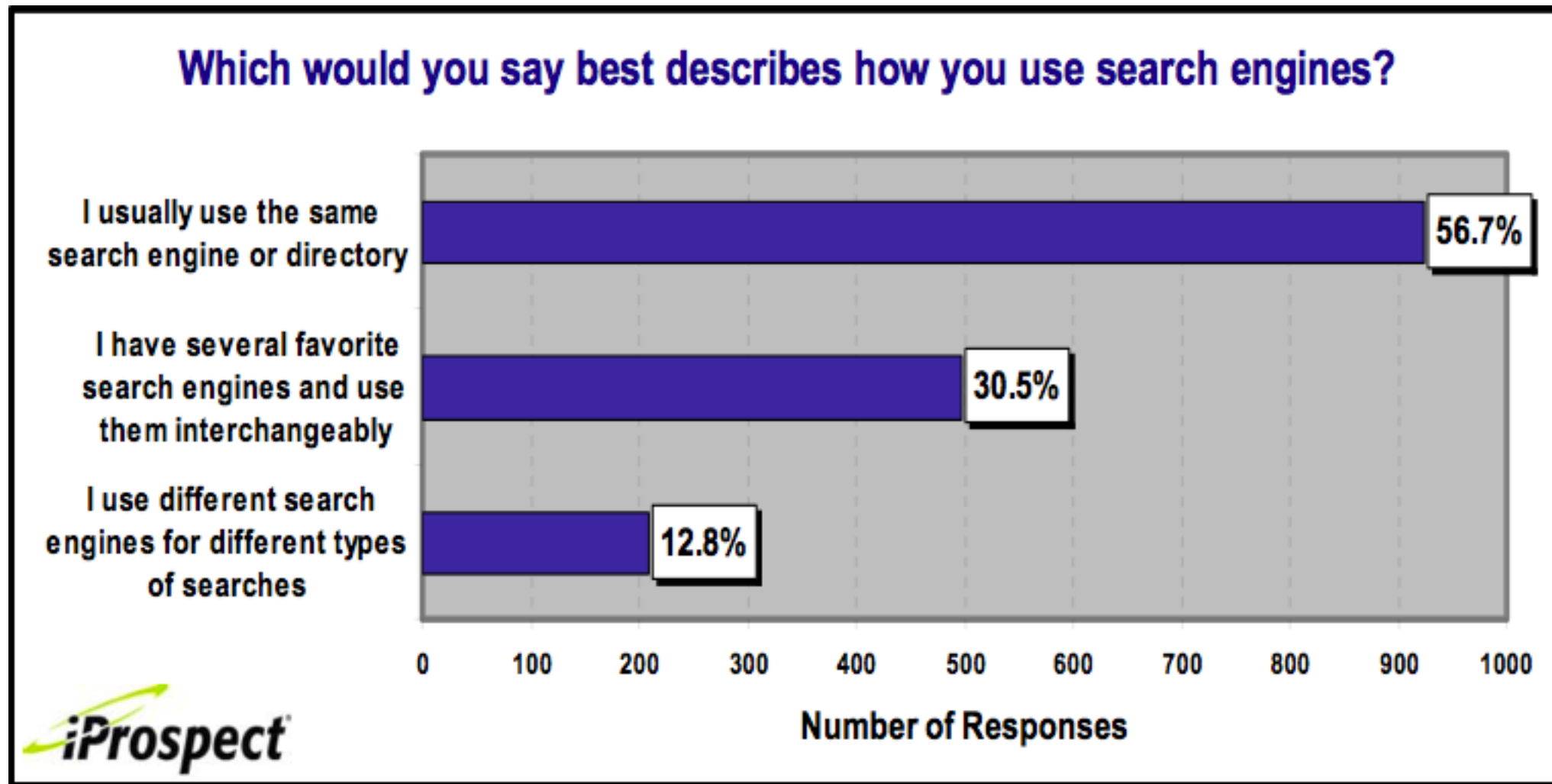


How do users evaluate search engines?

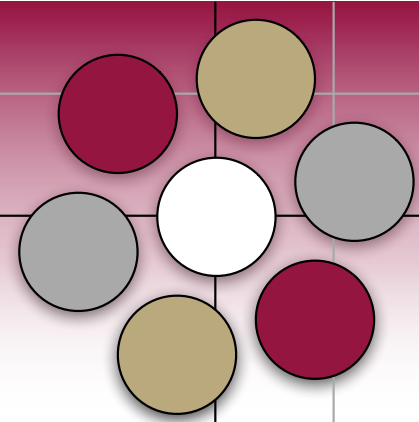
- Good U/I
 - Simple
 - No Clutter
- Pre and post processing tools
 - Spell check (“Did you mean?”)
 - Suggested alternative searches
 - Links to resources (maps, images, stock quotes)
- Able to deal with typical behavior
 - e.g., a URL typed into a search box



Loyalty to a given search engine



Web Users: Helping the User



True Example *

The User
flickr:crankyT



Task

Stop the noisy fan in the courtyard

Info Need

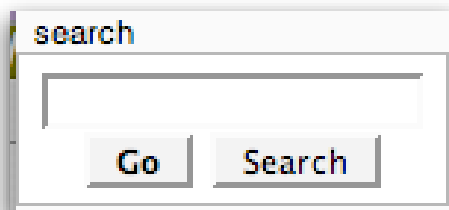
Info about EPA regulations

Verbal Form

What are EPA rules on noise pollution?

Query

EPA Sound Pollution



“To Google or to GoTo” Business Week Online 9/28/2001

Web Users: Helping the User



Answering “the need behind the query”

- The query is often an imprecise indicator of what the user really wants
- What can we do to get a better handle on the underlying information need?
- Query language
 - Adjust rank of English results for a Japanese query
- Use user context
 - In particular geographic context

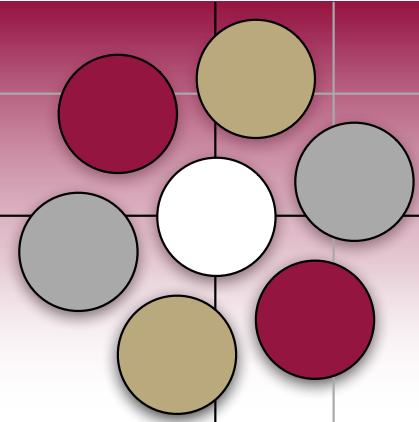
Web Users: Helping the User



Answering “the need behind the query”

- Guess what type of information the user wants
 - a web page?
 - a map?
 - a stock price?
 - what else?
- Correct queries
 - Suggest correct spellings
 - Suggest related searches (google-fu)

Web Users: Helping the User



Example

- Google

-

-

-

-

- C

-

-

kitakurihama - Google Search

http://www.google.com/search?hl=en&cli keitei

Last open ▾ Class Development ▾ Research Links ▾ UCI ▾ Thesis pages ▾ DR ▾ Conferences ▾ Place Lab ▾ >>

Web Images Maps News Shopping Gmail more ▾

donald.j.patterson.iii@gmail.com | Web History | My Account | Sign out

Google kitakurihama Search Advanced Search Preferences

Web Results 1 - 10 of about 276 for kitakurihama. (0.12 seconds)

Did you mean: [kujukurihama](#)

[MySpace.com - Nasty - Kitakurihama, Kanagawa - Hip Hop - www ...](#)
MySpace music profile for Nasty with tour dates, songs, videos, pictures, blogs, band information, downloads and more.
[profile.myspace.com/index.cfm?fuseaction=user.viewprofile&friendid=1000891776](#) - 48k - [Cached](#) - [Similar pages](#) - [Note this](#)

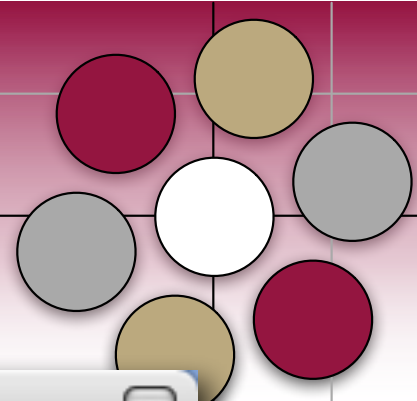
[Kitakurihama](#) - [[Translate this page](#)]
Kitakurihama. 1 to 7 out of 7 Glee land (グリーンランド) Karaoke box HAIR CLINIC TAIRA (ヘアークリニック タイラ) Hairdressing and cosmetics TAKOCHU (たこ忠) ...
[p.bari3.net/.../lookup2.asp?agntcd=1802&ct1=&ct1_nm=&ct2=007&ct2_nm=Kitakurihama&lv=1](#) - 5k - [Cached](#) - [Similar pages](#) - [Note this](#)

[保育所ちびっこランド 北久里浜園](#) - [[Translate this page](#)]
混合自由保育で三育法を実践。
[www.kidslink.jp/chibikko-kitakurihama/](#) - 2k - [Cached](#) - [Similar pages](#) - [Note this](#)

[YuDiary](#)
i waited for my friend on **kitakurihama** station at 1:00pm. there are a lot of fields in misakiguchi. i drove my car from **kitakurihama** station to his house. ...
[idauyutajifu.blogspot.com/](#) - 61k - [Cached](#) - [Similar pages](#) - [Note this](#)

[Takenori Co.,Ltd Kanagawa/Yokosuka Construction work Renovation ...](#)
Miura/Fuzisawa Bank(**Kitakurihama** Branch).Shounan Bank(**Kitakurihama** Branch) ... JA Yokosuka/Hayama farm cooperative(**Kitakurihama** Branch)

Web Users: Helping the User



infermatics - Google Search

http://www.google.com/search?hl=en&client=firefox-a&rls=org.mozilla G keitei

Last open ▾ Class Development ▾ Research Links ▾ UCI ▾ Thesis pages ▾ DR ▾ Conferences ▾ Place Lab ▾ OpenPresence ▾

Web Images Maps News Shopping Gmail more ▾ donald.j.patterson.iii@gmail.com | Web History | My Account | Sign out

Google[™] infermatics Search Advanced Search Preferences

Web Results 1 - 10 of about 87 for infermatics. (0.13 seconds)

Did you mean: [informatics](#)

[DÜŞÜNCELER - dellal international infermatics - Blogcu](#)
dellal international **infermatics**. emrah dellal. Bağlantılar. Ana Sayfa · Profilim · Arşiv · RSS.
Son Yazılar. DÜŞÜNCELER. 18/6/2006 - DÜŞÜNCELER ...
[mavilimm.blogcu.com/717078/](#) - 10k - [Cached](#) - [Similar pages](#) - [Note this](#)

[dellal international infermatics - dellal.com.tr.tc - www dellal ...](#)
emrah dellal,arkadaş arama,sohbet,google,, dellal com tr tc, dellal.com.tr.tc.
[www.dellal.com.tr.tc/](#) - 2k - [Cached](#) - [Similar pages](#) - [Note this](#)

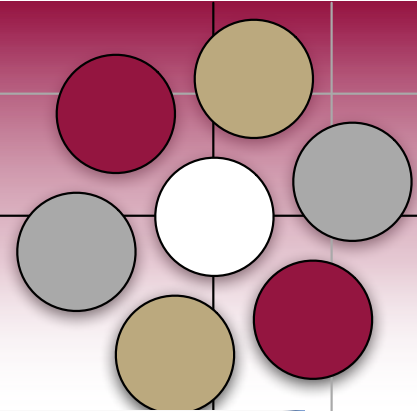
[LexiconWiki: Information](#)
... the academic community, promoting and assessing the use of Information and producing
dedicated **Infermatics** for both academic and general consumption. ...
[kevan.org/lexwiki.pl?action=browse&diff=1&id=Information](#) - 8k -
[Cached](#) - [Similar pages](#) - [Note this](#)

[Hepatology Research : Medical economics—Hepatitis C virus ...](#)
Medical **Infermatics** and Decision Sciences, Yamaguchi University, School of Medicine,
1-1-1 Minami-Kogushi, Ube, Yamagnchi: 755-8505, Japan ...
[linkinghub.elsevier.com/retrieve/pii/S1386634602000414](#) - [Similar pages](#) - [Note this](#)

[IngentaConnect Medical economics-Hepatitis C virus infection and ...](#)

Done

Web Users: Helping the User



Ex

Live Search: rock

http://search.live.com/results.aspx?q=rock&go=Search&mkt=en-1

Last open ▾ Class Development ▾ Research Links ▾ UCI ▾ Thesis pages ▾ DR ▾ Conferences ▾ Place Lab ▾ OpenPresence ▾

Live Search Home Hotmail Spaces OneCare Sign in

rock Search Advanced Options

Web results 1-10 of 441,000,000

See also: [Images](#), [Video](#), [News](#), [Maps](#), [MSN](#), [More](#) ▾

The Rock - Jellyfish.com/Movies Sponsored sites
Shop and compare prices across stores. Rebates up to \$1.10.

Rock.com® - The Official Site of Rock Music®
Free Email & Free Music Downloads by **Rock.com**. Get Free Internet Radio, Music Downloads & more. Get your free @rock.com email today
[www.rock.com](#) · [Cached page](#)

Rock music - Wikipedia, the free encyclopedia
Rock music is a form of popular music with a prominent vocal melody accompanied by guitar , drums , and bass . Many styles of **rock** music also use keyboard instruments such as organ ...
[en.wikipedia.org/wiki/Rock_music](#) · [Cached page](#)

Rock and roll - Wikipedia, the free encyclopedia
Rock and roll is a genre of music that evolved in the United States in the late 1940s and early 1950s, and quickly spread to the rest of the world.
[en.wikipedia.org/wiki/Rock_and_roll](#) · [Cached page](#)
[Show more results from en.wikipedia.org](#)

Rock - Australia's Climbing Magazine
Subscription information. Overview of current issue. Classifieds. Back issues and guide books for sale.

Related searches
[Rock Music](#)
[Rocks Minerals](#)
[Pictures Of Rocks](#)
[Rock Star](#)
[Kid Rock](#)
[Rock Cycle](#)
[Hard Rock Cafe](#)

Sponsored sites
Free Rock Music
Listen To Your Favorite Music Now!
It's Free with the Music Toolbar
[Music.alot.com](#)

The Rock
Interviews with Stars, Newsmakers & Icons at The Biography Channel®.
[www.biography.com](#)

Rock Videos
Enjoy Savings & Selection on **Rock** Videos!
[Shopzilla.com](#)

Done

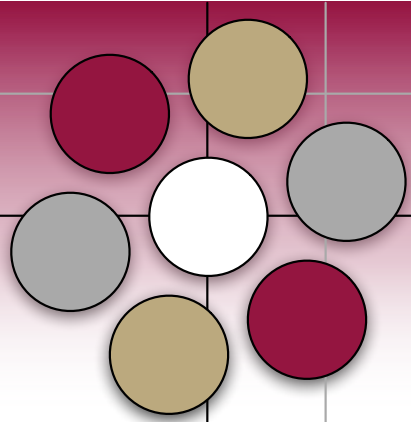
Web Users: Helping the User



Query shortcuts

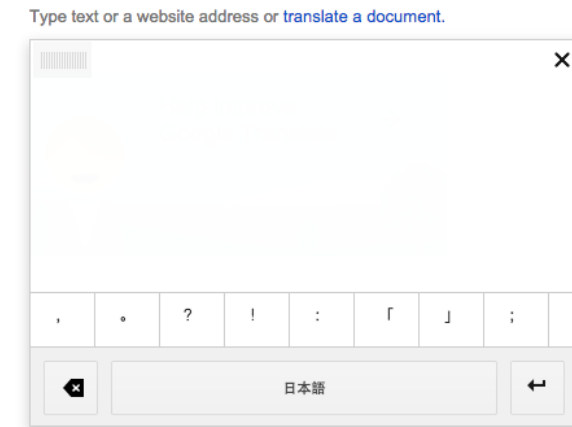
- Map: “irvine, ca 92614”
- Calculation: “5+4”
- Flight Info: “american airlines 715”
- Stock price: “msft”
- Unit conversion: “1 dollar in euro”
- Music: “White Stripes”

Web Users: Helping the User

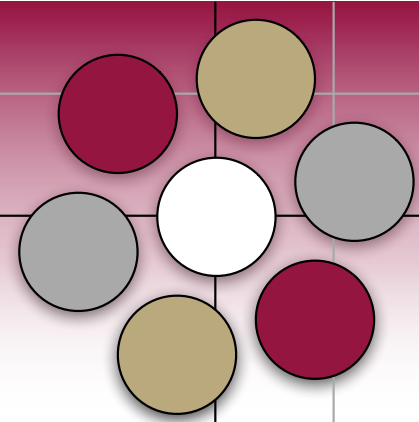


Query shortcuts

- apple vs banana
- tip calculator
- zerg rush
- 88921384914 in english
- atari breakout




Web Users: Helping the User



Query

- Map
- Calc
- Flight
- Stock
- Unit
- Music

 **WolframAlpha**™ computational knowledge engine


when is the next solar eclipse?

Input interpretation:
next solar eclipse after today

Eclipse type:
annular

Eclipse date:
11:06 pm PST | Thursday, January 14, 2010 (7.344 days from now)

Eclipse path: [Show 3D](#) | [Orthographic projection](#)

A map of the world showing the path of an annular solar eclipse. The path is highlighted in red and starts in the Atlantic Ocean, crosses Africa, and ends in Asia. The map is in an orthographic projection.

Web Users: Helping the User

White Stripes - Yahoo! Search Results

http://search.yahoo.com/search;_ylt=A0oGkmbxeZdH0nsBWtZXNyo

american airlines 715 - Googl... Live Search: white stripes White Stripes - Yahoo! Search ...

Web | Images | Video | Local | Shopping | more

White Stripes Search Options

YAHOO!

1 - 10 of about 53,700,000 for White Stripes (About this page) - 0.11 sec.

The White Stripes - Official Site
www.whitestripes.com
[Albums](#) | [Lyrics](#) | [Photos](#) | [Videos](#)

Watch Music Videos

Play Popular Songs

- Icky Thump
- You Don't Know What Love Is (You Jus...
- 300 M.P.H. Torrential Outpour Blues

[More Songs...](#)

The White Stripes
Official site for The **White Stripes**, the Detroit-based garage/blues duo featuring Jack and Meg White.
www.whitestripes.com - 2k - [Cached](#)

White Stripes v2
Jack and Meg **White** fansite includes news and tour information, forum, links, release details, and more.
www.whitestripes.net - 72k - [Cached](#)

SPONSOR RESULTS

White Stripes Ringtones
Complimentary **White Stripes** Ringtones. Get Them Instantly.
www.TheBombRingtones.com

White Stripes Ringtones
White Stripes Ringtones. Get Them Complimentary Now.
ChartRingers.com/whitestripes

The White Stripes
Find, compare & buy. Compare & Buy from 1000's of Stores.
www.Dealtime.com

The White Stripes at Amazon.com
Low prices on new & used music. Qualified orders over \$25 ship free.
Amazon.com/music

[See your message here...](#)

Done

Web Users: Helping the User

The screenshot shows a web browser window titled "Live Search: Don Patterson irvine ca". The address bar displays the URL <http://search.live.com/results.aspx?q=Don+Patterson+irvine+ca&gc>. The browser has several tabs open: "american airlines 715 - Googl...", "Live Search: Don Patterson irvin...", and "White Stripes - Yahoo! Search ...". The Live Search interface includes a search bar with the text "Don Patterson irvine ca" and a green "Search" button. Navigation links for "Home", "Hotmail", "Spaces", "OneCare", and "Sign in" are visible. The search results are categorized into "Web results" (1-9 of 253,000) and "Sponsored sites". The "Web results" section includes a link to "Top local listings for Don Patterson near Irvine, CA" and a list of three entries for Don Patterson with their phone numbers and addresses. The "Sponsored sites" section includes links to "Irvine CA Book Hotel", "Locate Don Patterson", "Don Patterson at Amazon", and "Irvine CA - Cheap Rates". The "Jerry M. Patterson" entry is also visible at the bottom.

Live Search: Don Patterson irvine ca

http://search.live.com/results.aspx?q=Don+Patterson+irvine+ca&gc

american airlines 715 - Googl... Live Search: Don Patterson irvin... White Stripes - Yahoo! Search ...

Live Search Home Hotmail Spaces OneCare Sign in

Don Patterson irvine ca Search Advanced Options

Web results 1-9 of 253,000

See also: [Images](#), [Video](#), [News](#), [Maps](#), [MSN](#), [More](#) ▼

» [Top local listings for Don Patterson near Irvine, CA](#) (1 more) Is this useful? [Yes](#) | [No](#)

- Don Patterson (626) 918-1974 15626 Dubesor St, La Puente
- Don Patterson (213) 389-7583 226 S Berendo St Apt 205, Los Angeles
- Don Patterson (760) 729-6941 2768 Dundee Ct, Carlsbad

[DBLife: University of California-Irvine](#)

Jadwiga Indulska, University of Queensland, Australia Don Patterson, University of California Irvine, USA Tom Rodden, Nottingham University, UK Program Committee:
dblife.cs.wisc.edu/search.cgi?entity=entity-13655&begin=60 · [Cached page](#)

[DBLife: University of California-Irvine](#)

Jadwiga Indulska University of Queensland, Australia Don Patterson University of California Irvine, USA Tom Rodden Nottingham University, UK...
dblife.cs.wisc.edu/search.cgi?entity=entity-13655&begin=90 · [Cached page](#)

[Show more results from dblife.cs.wisc.edu](#)

[Jerry M. Patterson](#)

Professor: University of California at Irvine (2000-) Moulton Patterson Associates ... Partner (1967-75) State Bar of California 1967. Do you know something we don't?
www.nndb.com/people/406/000163914 · [Cached page](#)

Sponsored sites

[Irvine CA Book Hotel](#)
Visiting California & Need a Hotel?
Reserve Rooms for Irvine CA!
www.calibex.com/hotels

[Locate Don Patterson](#)
Current address and phone available.
Instant results.
www.usa-people-search.com

[Don Patterson at Amazon](#)
Low prices on new & used music.
Qualified orders over \$25 ship free
amazon.com/music

[Irvine CA - Cheap Rates](#)
Visiting California & Need a Hotel?
Compare Top Sites for Irvine CA!
www.nextag.com/hotels

[See your message here...](#)

Done

Web Users: Helping the User

The screenshot shows a web browser window titled "diabetes - Yahoo! Search Results". The address bar displays the URL "http://search.yahoo.com/search;_ylt=A0oGkl24d5dHeL0A2gZXNyo/". The search bar contains the word "diabetes" and a "Search" button. The page shows search results for "diabetes", with a count of "1 - 10 of about 222,000,000 for diabetes (About this page) - 0.09 sec.".

SPONSOR RESULTS

Free Diabetes Supplies
www.LibertyMedical.com - Medicare Diabetics - Your supplies Could be 100% covered. Learn more.

Diabetes and Pain
www.Cymbalta.com - Diabetic peripheral neuropathic pain (DPNP) -learn about managing it.

Learn About Diabetes
www.Type2Diabetes-Info.com - Learn more about type 2 **diabetes** including treatment options.

Diabetes Learning Center
Educational site explains how to manage type 1 and type 2 **diabetes**.
www.bddiabetes.com

Manage Diabetes Naturally
Sugar Free Psyllium Fiber Products for Type 2 **Diabetes**. Order Online.
www.konsyl.com

Diabetes
Get a Free Meter from Liberty if You have **Diabetes** & on Medicare.
www.Diabetic-Testing-Supplies.com

Diabetes Is Info
Get Info on **Diabetes** Is from 14 search engines in 1.
www.info.com

Losing Income Due to Diabetes?
Unable to work due to **Diabetes**?

Diabetes Information - Health Results

Articles	Community	Other Resources
Symptoms	Yahoo! Groups	National Institute of Diabetes &...
Treatments	Ask a Question	American Diabetes Association
Medications	Read Q & A	MedlinePlus: Diabetes

Yahoo! Shortcut - [About](#)

CDC: Basics About Diabetes
Fact sheet about **diabetes** covering its common symptoms, risk factors, and management options.
www.cdc.gov/diabetes/faq/basics.htm - 36k - [Cached](#)

American Diabetes Association
Provides extensive news and information about **diabetes** including symptoms, nutrition, weight

Done

Web Users: Helping the User



Context

- Context determination
 - Spatial
 - User Location
 - IP geocoding
 - cell phone localization
 - User input (“make this my home”)
 - Web page location
 - Target Location
 - “SNA” maps to a latitude and longitude

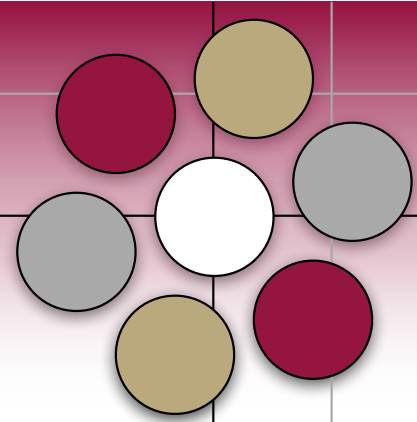
Web Users: Helping the User



Context

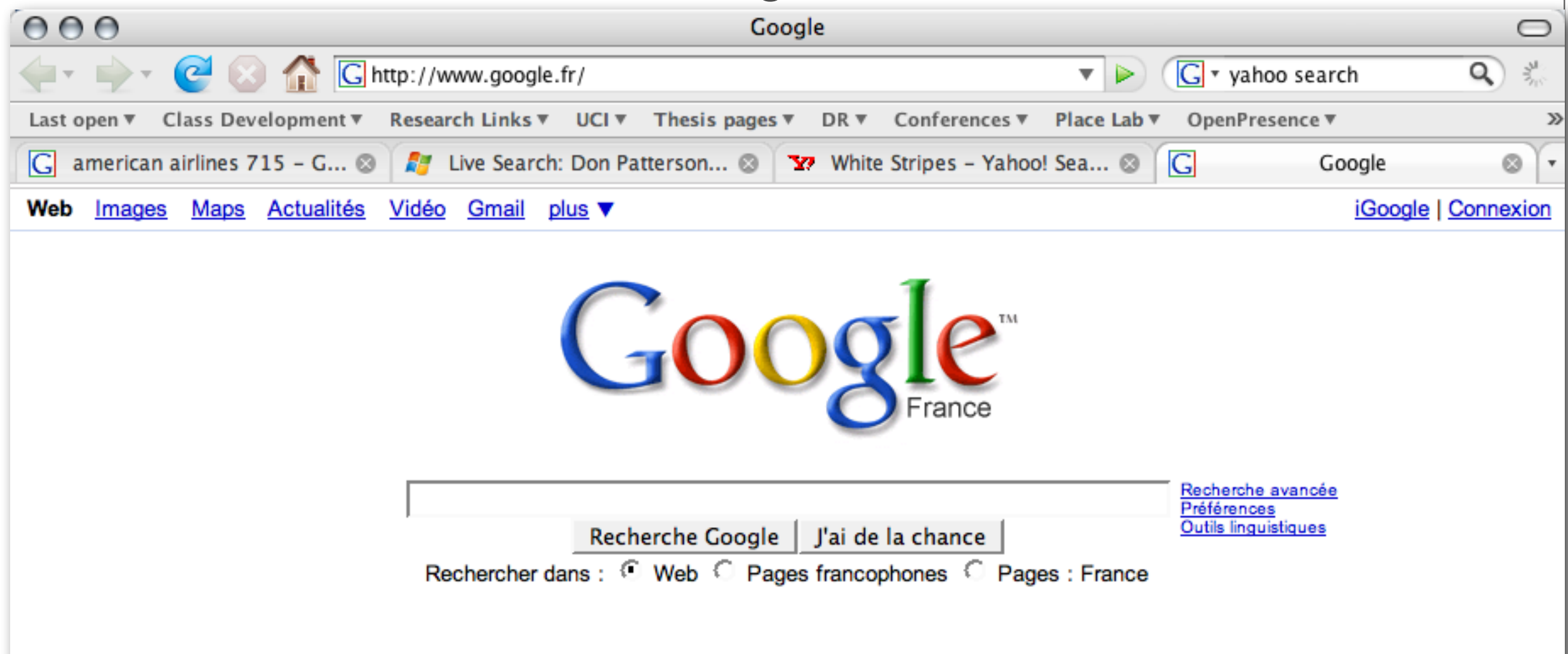
- Query Stream
 - Previous queries have influence
- Personal Search
 - Your profile and history have influence
- Explicit
 - You choose to use a vertical search engine

Web Users: Helping the User



Context

- Implicit Context
 - google.fr is the default in france
 - User Location and Target Location





Context Use

- Restrict Results / Delete inappropriate results
- Rerank results based on context
- Change the content presentation
 - Language etc.
- Specifying a context is built into some of the search APIs

Web Users: Helping the User

C

Amazon.com: single cookbook

http://www.amazon.com/exec/obidos/search-handle-url/index=blend... yahoo search

Last open ▾ Class Development ▾ Research Links ▾ UCI ▾ Thesis pages ▾ DR ▾ Conferences ▾ Place Lab ▾ OpenPresence ▾ Blogs and Wikis ▾

Amazon.com: single cook... Live Search: Don Patterson ... White Stripes - Yahoo! Sear... Google


amazon.com Hello. Sign in to get [personalized recommendations](#). New customer? [Start here](#).


Your Amazon.com Today's Deals ▾ Gifts & Wish Lists ▾ Gift Cards ▾ Your Account | Help

Shop All Departments ▾ Search Amazon.com single cookbook GO Cart Your Lists ▾

Category
Any Category
Books (395)
Software (2)
Home & Garden (1)

Listmania!


Cooking - Just the two of us/you: A list by Laura

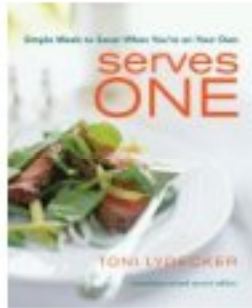

Best Books for Health and Diet in 2007: A list by Moni Jeddian "PWR"

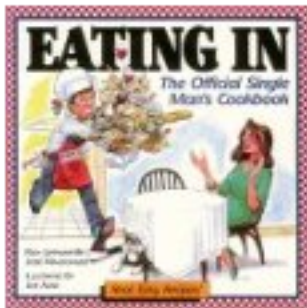
► Create a Listmania! list


Search Listmania! GO!

"single cookbook"

Showing Top Results [« Previous](#) | **Page: 1** 2 3 | [Next »](#)

- 

Serves One: Simple Meals to Savor When You're on Your Own by Toni Lydecker (**Paperback** - Dec 25, 2005)
Buy new: ~~\$16.95~~ **\$11.53** [26 Used & new](#) from **\$7.94**
Get it by **Thursday, Jan 24** if you order in the next **5 hours** and choose one-day shipping. Eligible for **FREE** Super Saver Shipping.
★★★★☆ (8)
Books: See all 395 items
- 

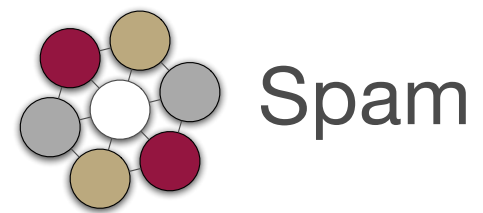
Eating in: The Official Single Man's Cookbook by Rich Lippman, Jose Maldonado, and Joe Azar (**Paperback** - Oct 1988)
Buy new: **\$8.95** [11 Used & new](#) from **\$5.92**
Not in stock; order now and we'll deliver when available
Eligible for **FREE** Super Saver Shipping.
★★★★☆ (2)
Books: See all 395 items
- 

Going Solo in the Kitchen by Jane Doerfer (**Paperback** - Aug 25, 1998)
Buy new: ~~\$16.95~~ **\$11.53** [33 Used & new](#) from **\$7.45**
Get it by **Thursday, Jan 24** if you order in the next **5 hours** and choose one-day shipping. Eligible for **FREE** Super Saver Shipping.

Done

Overview

- Introduction
- Classic Information Retrieval
- Web IR
- Sponsored Search
- Web Search Basics
 - Size of the Web
- Web Users





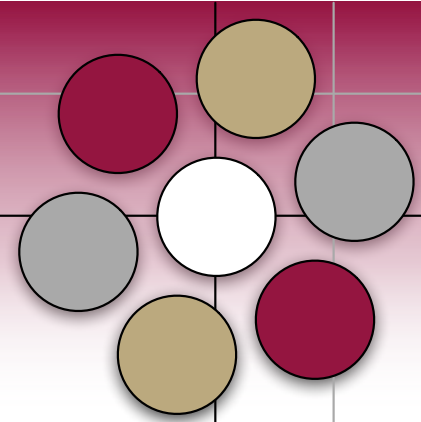
The trouble with paid placement (aka ads):

- It costs money... so instead
- Search Engine Optimization (“SEO”)
 - define: “Tuning” your web page to rank highly in the search results for select queries
- Alternative to paying for placement
- It is marketing. Getting your content to your audience.



Search Engine Optimization

- Motives
 - Commercial
 - Political
 - Religious
 - Lobbying
- Who does this?
 - Internally: webmasters
 - Commercially: companies, consultants
 - Hosting services



Search Engine Optimization

- Learn more about how to do it online:
 - Web-Master World
 - <http://www.webmasterworld.com>
 - Search Engine Specific Tricks
 - Discussions about academic papers and results

Content creation vs link building

Let's get the IncrediBills and the like to pitch in on this one

Moved: 2 pages on site %55 similar - duplicate content or not?

Post was moved here: <http://www.webmasterworld.com/google/3537427.htm>

Ideal length of content for SEO purposes?

How long should an article be?

Source of Competitor's Traffic from a Major .com?

No apparent outlinks from the site to competitor's site

Moved: week old website ranked #14 in google for main keyword



Search Engine Optimization

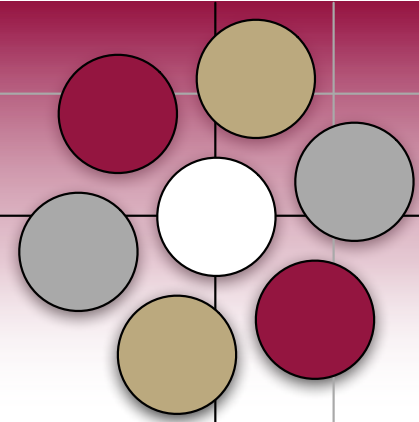
- There are ethical and inethical ways to approach SEO
- Legitimate approach is to:
 - create valuable content
 - make it widely accessible
 - clearly organize it
 - keep it up to date
 - use web standards
 - use web validation tools
 - get high visibility sites to link to your content



Search Engine Optimization

- Inethical approaches (aka spam):
 - lots of tricks
 - make lots of fake pages which point to your site
 - make lots of fake comments on sites which point to your site
 - In a nutshell, “lie”
- Sometimes legitimate and illegitimate techniques are hard to differentiate. It can be a fine line between them.

Spam



Search Engine Optimization

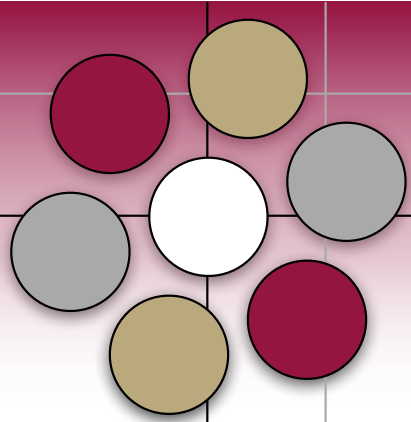
- Ranking depends on the data center
 - http://www.flickr.com/photos/the_impression_that_i_get/1321041609/
- Examine the different results:
 - <http://www.mcdar.net/dance/index.php>





Keyword Stuffing

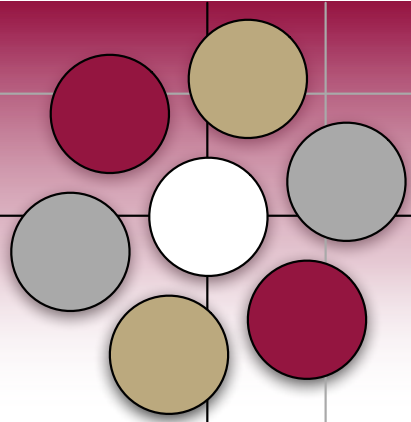
- First Generation Search Engines
 - Heavily relied on tf/idf ratio.
 - E.G. The highest ranking page for the query “brilliant computer scientist” had the most examples of those words.



Keyword Stuffing

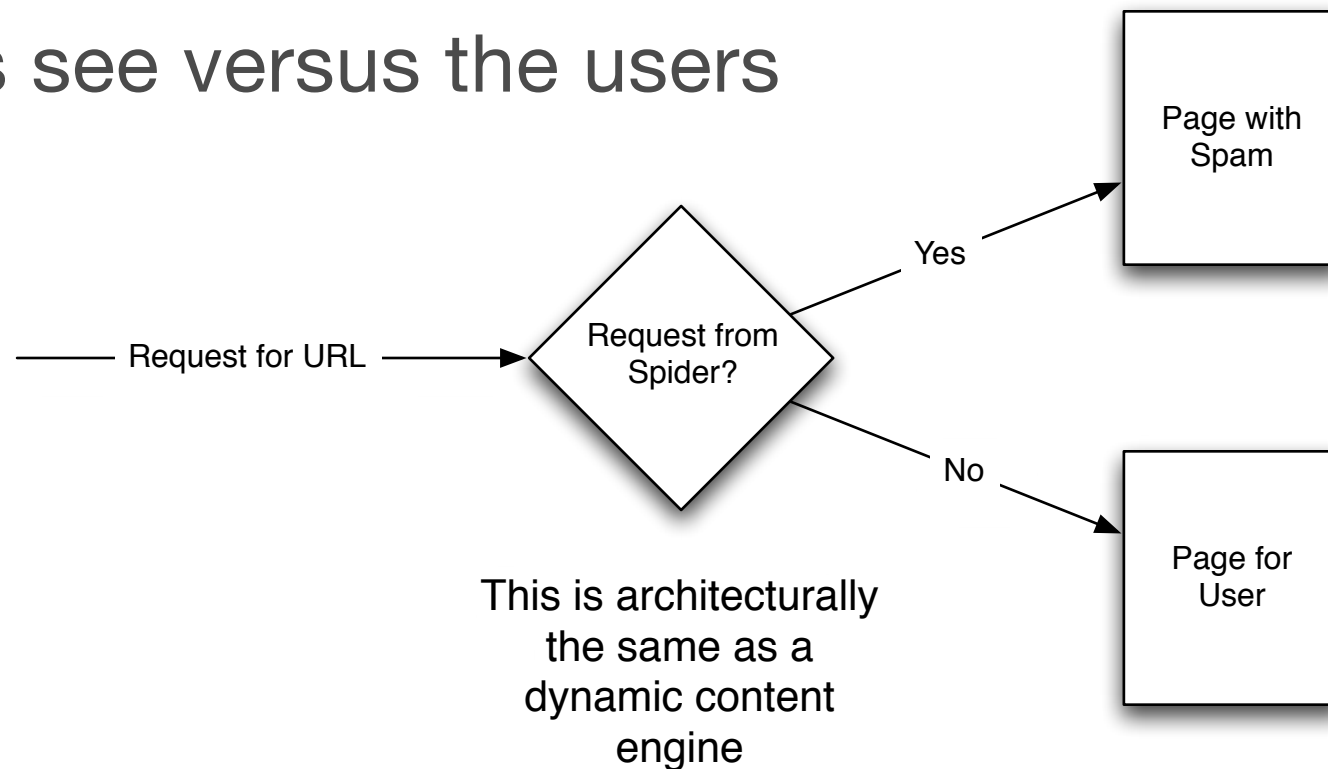
- So SEOs responded by screwing around with keywords
 - Misleading meta-tags
 - Repeating keywords over and over and over and....
 - Playing games with colors. (white on white keywords)
 - visible to spiders but not users in browsers

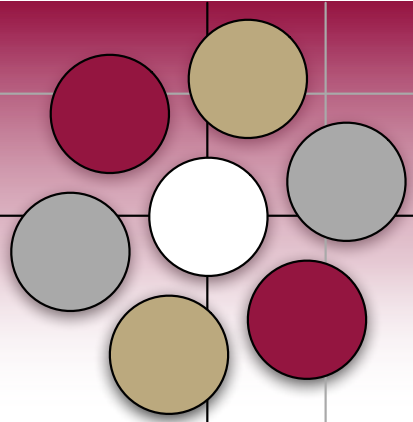
```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.c
<html xmlns="http://www.w3.org/1999/xhtml">
<head>
<COMMENT TITLE="MONITOR"></COMMENT>
<meta http-equiv="Content-Language" content="en-us" />
<meta http-equiv="Content-type" content="text/html; charset=iso-8859-1"/>
<META NAME="ROBOTS" CONTENT="NOODP"><meta name="verify-v1" content="aeVxP6zTHeQzT620ipj5+ikXd/VXcdlKoYUJ/C6vVdY=" />
<META NAME="keywords" content="Expedia, Travel, Cheap Airfare, Car, Hotels, Vacations, Airfare, Car Rental, Cruises,
<META NAME="description" content="Purchase airline tickets, make hotel reservations, find vacation packages, car rent
```



Keyword Stuffing

- Cloaking
 - define: Serving different content to a spider than to a user.
 - More sophisticated versions of differentiating what the spiders see versus the users





Other spam techniques

- Doorway pages
 - Like cloaking but using a redirect
 - Initial page is optimized for a keyword then a redirect takes the user to the “real” page
- Link spamming
 - Programs that search for blogs and automatically leave comments with links
- Robot Clicker-Fraud
 - Programs that “click” on query results to up their value.

Spam

Spam Industry

Advanced Traffic:

Get a **first page listing on Google** - **GUARANTEED!** For maximum search engine traffic - the best of SEO and search advertising. Visitors in just 48 hours from \$7/day. Discover the traffic potential!

[Find out more](#)



WARNING: This site contains sneaky, underhanded Black Hat Seo tactics.

Black Hat Seo is responsible for more online fortunes than you'd care to imagine but it's NOT for everybody.

Make Money Blogging

See How I Earn Over [Six Figures](#) a year Blogging

I Will Get Your Website to the Top of Google!

The art of search engine optimization...gaining **top spots on Google**...is no easy chore. I know...this is my job...

I assist people in getting top positions for their websites on Google, Yahoo, MSN and all the other major search engines.

There are a few givens on the internet when it comes to trying to market goods and services:

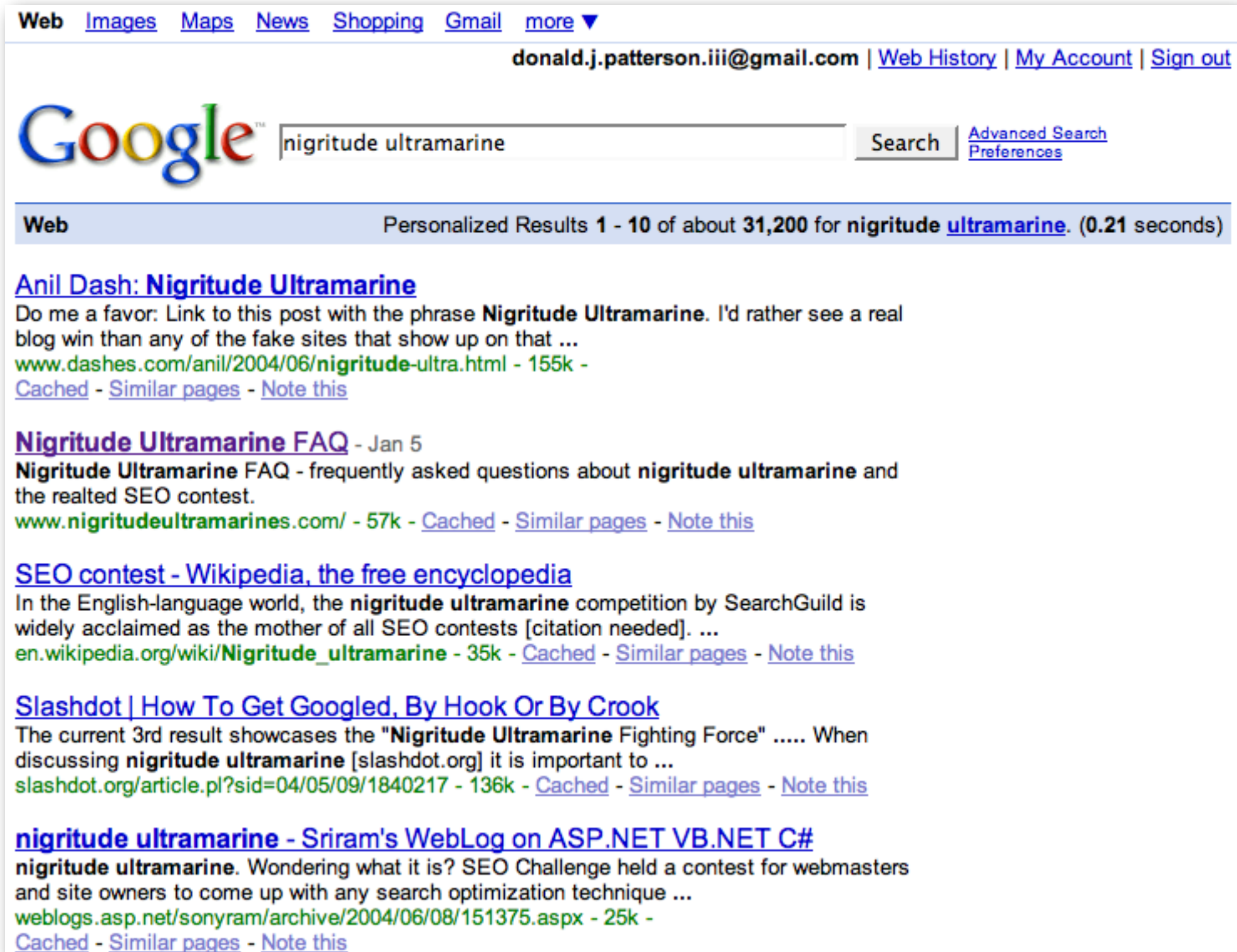
No Traffic=No Sales!

End of story...that's it...bottom line!

If you have a website...

Spam

Spam Contest



The image shows a screenshot of a Google search results page. At the top, there are navigation links: Web, Images, Maps, News, Shopping, Gmail, and a 'more' dropdown arrow. Below these, the user's email address 'donald.j.patterson.iii@gmail.com' is displayed, along with links for 'Web History', 'My Account', and 'Sign out'. The Google logo is on the left, and the search bar contains the text 'nigrITUDE ultramarine'. To the right of the search bar is a 'Search' button and links for 'Advanced Search' and 'Preferences'. Below the search bar, a blue bar indicates 'Personalized Results 1 - 10 of about 31,200 for nigrITUDE ultramarine. (0.21 seconds)'. The search results are listed below, each with a title, a brief description, and a URL. The first result is 'Anil Dash: NigrITUDE Ultramarine' with a description asking for a link to a post. The second result is 'NigrITUDE Ultramarine FAQ' dated Jan 5, with a description about frequently asked questions. The third result is 'SEO contest - Wikipedia, the free encyclopedia' with a description about the NigrITUDE ultramarine competition. The fourth result is 'Slashdot | How To Get Googled, By Hook Or By Crook' with a description about the 'NigrITUDE Ultramarine Fighting Force'. The fifth result is 'nigrITUDE ultramarine - Sriram's WebLog on ASP.NET VB.NET C#' with a description about a contest for webmasters.

Web Images Maps News Shopping Gmail more ▼

donald.j.patterson.iii@gmail.com | Web History | My Account | Sign out

Google™ nigrITUDE ultramarine Search Advanced Search Preferences

Web Personalized Results 1 - 10 of about 31,200 for nigrITUDE ultramarine. (0.21 seconds)

Anil Dash: NigrITUDE Ultramarine
Do me a favor: Link to this post with the phrase **NigrITUDE Ultramarine**. I'd rather see a real blog win than any of the fake sites that show up on that ...
www.dashes.com/anil/2004/06/nigrITUDE-ultra.html - 155k -
[Cached](#) - [Similar pages](#) - [Note this](#)

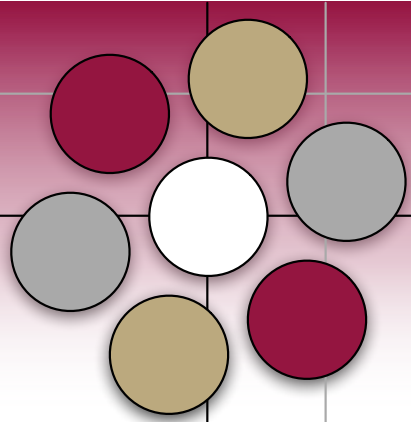
NigrITUDE Ultramarine FAQ - Jan 5
NigrITUDE Ultramarine FAQ - frequently asked questions about **nigrITUDE ultramarine** and the realted SEO contest.
www.nigrITUDEultramaries.com/ - 57k - [Cached](#) - [Similar pages](#) - [Note this](#)

SEO contest - Wikipedia, the free encyclopedia
In the English-language world, the **nigrITUDE ultramarine** competition by SearchGuild is widely acclaimed as the mother of all SEO contests [citation needed]. ...
en.wikipedia.org/wiki/NigrITUDE_ultramarine - 35k - [Cached](#) - [Similar pages](#) - [Note this](#)

Slashdot | How To Get Googled, By Hook Or By Crook
The current 3rd result showcases the "**NigrITUDE Ultramarine** Fighting Force" When discussing **nigrITUDE ultramarine** [slashdot.org] it is important to ...
slashdot.org/article.pl?sid=04/05/09/1840217 - 136k - [Cached](#) - [Similar pages](#) - [Note this](#)

nigrITUDE ultramarine - Sriram's WebLog on ASP.NET VB.NET C#
nigrITUDE ultramarine. Wondering what it is? SEO Challenge held a contest for webmasters and site owners to come up with any search optimization technique ...
weblogs.asp.net/sonyram/archive/2004/06/08/151375.aspx - 25k -
[Cached](#) - [Similar pages](#) - [Note this](#)

Spam



The war on spam

- Quality Indicators
 - Statistical Analysis of Links (aka PageRank)
 - votes from authors
 - Usage indicators (users visiting a page)
 - votes from users
- Anti-Robot techniques
 - “Captchas”
 - Completely Automated Public Turing Test to Tell Computers and Humans Apart





The war on spam

- Limits on meta keywords
- Spam Recognition by machine learning
- “no-follow” attribute
- Family Friendly filters
 - Automatic Detection of Pornography
 - Often the spammers desired landing page
- Text Analysis
 - Look for keywords and variants



The war on spam

- Robust Link Analysis
 - Ignore statistically improbable links
 - Use link analysis to detect spammers
 - “Guilt by association”



The war on spam

- Editorial Intervention
 - Blacklists
 - Query Reviews
 - Customer Complaints
 - Visualization Tools

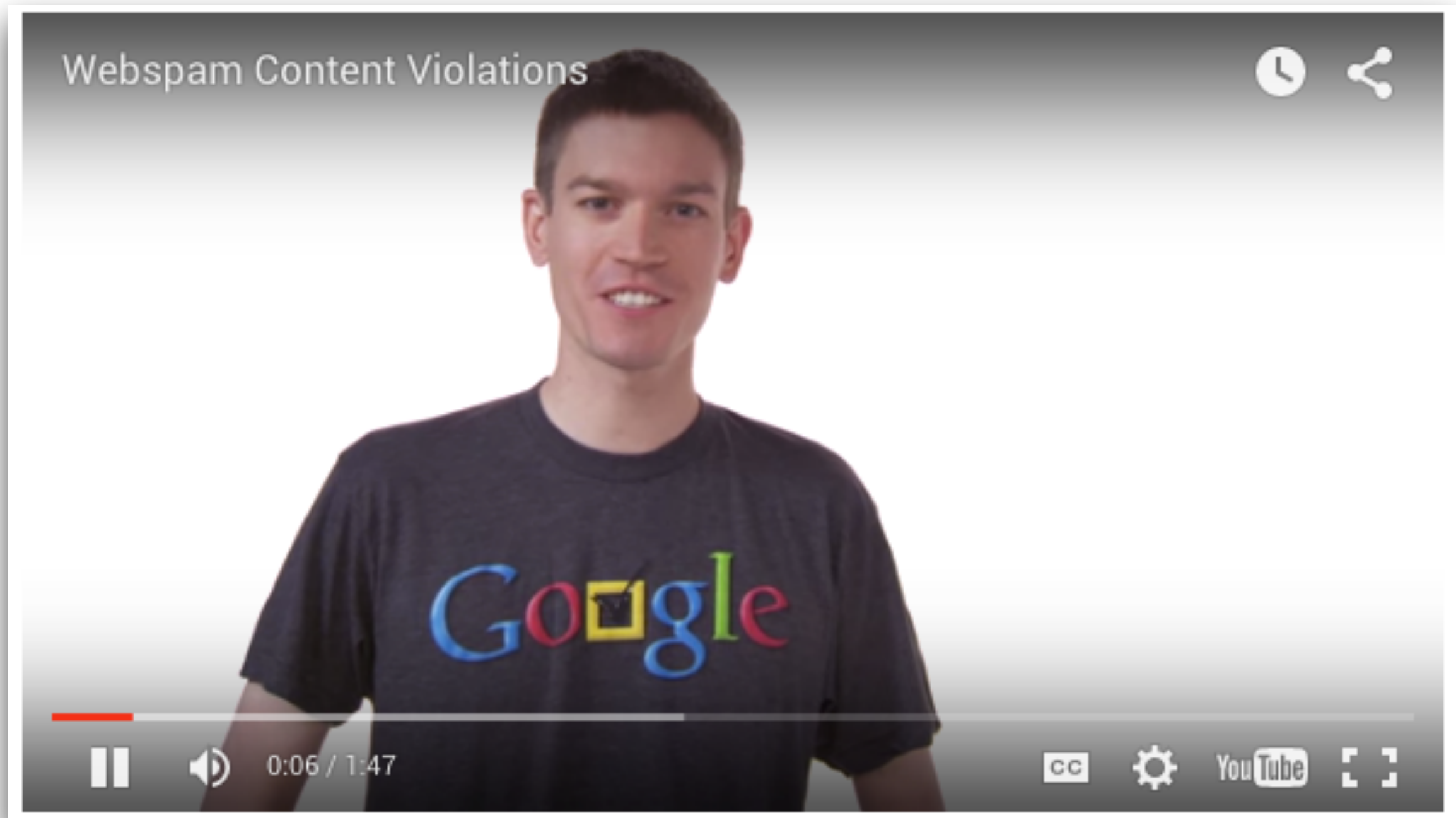
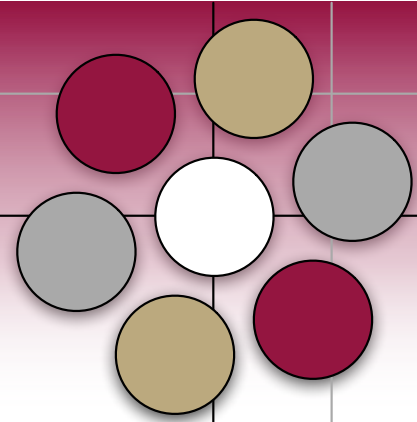


Webmaster Guidelines

- Search Engines have SEO policies
 - What is allowed and not allowed
- Example: Search for “google webmaster guidelines” or “msn guidelines for successful indexing”
- Ignore them at your own risk
- Once you are blacklisted by a search engine you will disappear from the web
 - Remember how search engines enable scalability?
- Adversarial IR Research:
 - <http://airweb.cse.lehigh.edu/>

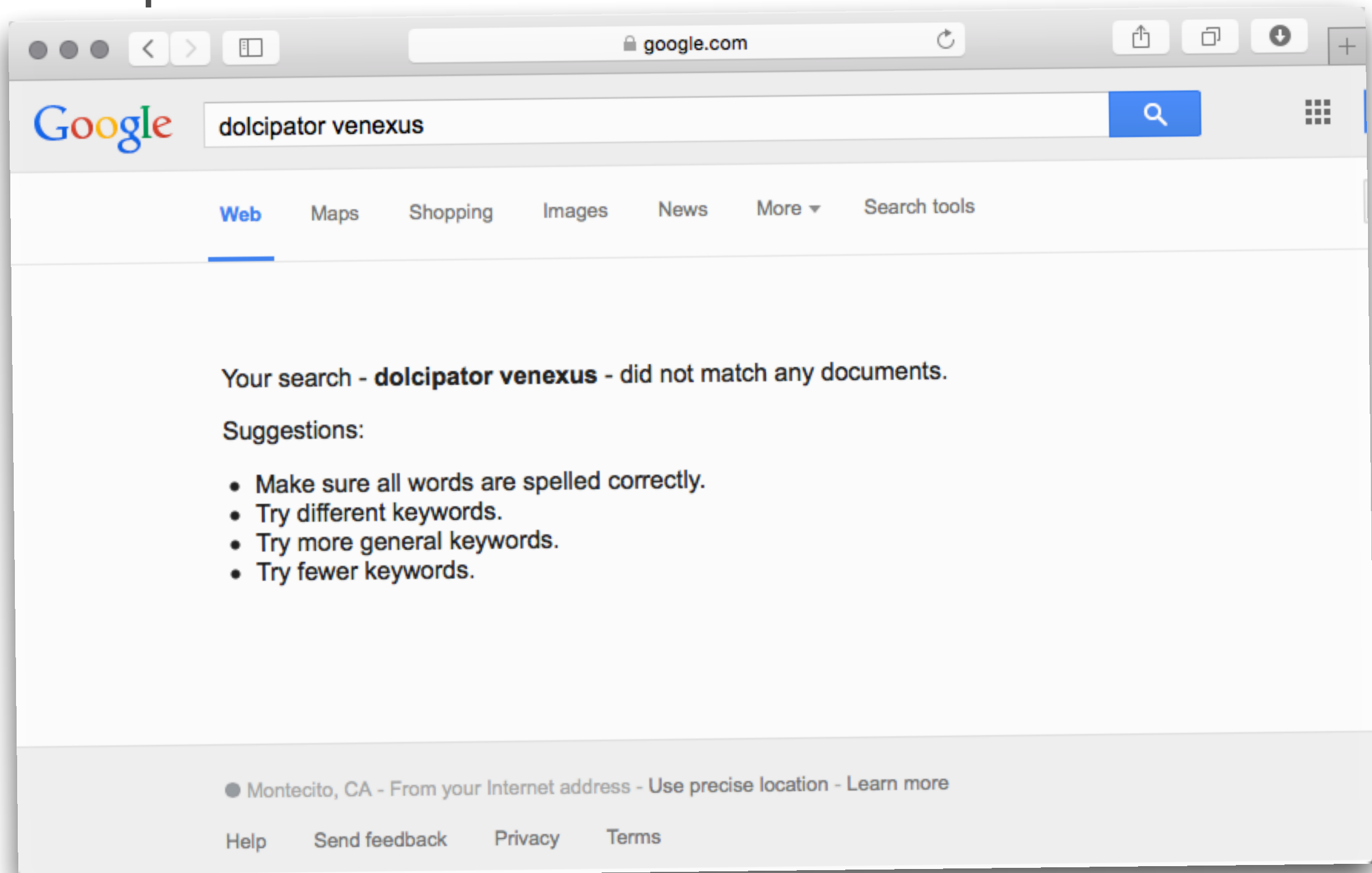
Spam

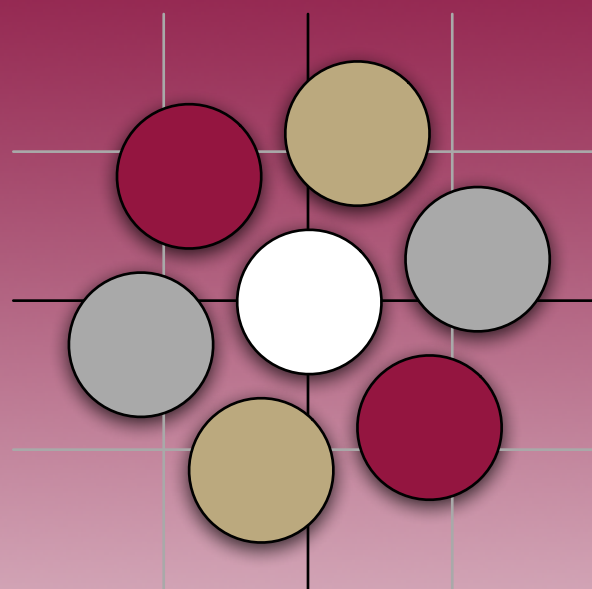
Webmaster Guidelines



Spam - Extra Credit

dolcipator venexus





WESTMONT COMPUTER SCIENCE