

The Many Faces of Google

A Presentation by Joel May

I. Our Goals

- A. To Learn how Google works.
- B. To Discover some Google tricks.
- C. To Explore some advanced features.

II. The Google Home Screen (www.google.com) contains

- A. The Search Term Box
- B. The Action Buttons
 - 1. Google Search
 - 2. I'm Feeling Lucky
- C. Along with links to several Types of Searches
 - 1. Web
 - 2. Images
 - 3. Groups
 - 4. News
 - 5. Froogle
 - 6. "More"
- D. As well as some additional features
 - 1. Advanced Search
 - 2. Preferences
 - 3. Language Tools

III. The Search Term Box Rules

- A. Rule #1: Be specific – if you aren't you'll end up with a lot of garbage
 - 1. condominiums: 2,540,000 hits
 - 2. condominiums florida naples time-share: 1,920 hits
- B. Rule #2: Use Quotes ("") to search for exact phrases
 - 1. Joe Montana: 1,080,000 hits
 - 2. "Joe Montana": 159,000 hits
 - 3. and, to demonstrate the first rule again: "Joe Montana" 1988 San Francisco: 6,730 hits
- C. Rule #3: Use the plus sign (+) or AND to require search terms be included
 - 1. Google ignores common words and characters such as "and" and "how" as well as certain single digits and single letters, because they tend to slow down your search without improving the results.
 - 2. If a common word is essential to getting the results you want, you can include it by putting a "+" sign in front of it.
 - 3. For example, to search for Star Wars, Episode I, use:
 - a. Star Wars Episode +I
 - b. Star Wars Episode AND I

D. Rule #4: Use the minus sign (-) or NOT to exclude search terms

1. Sometimes what you're searching for has more than one meaning (eagles can refer to either a large diurnal bird of prey or an Super Bowl ~~champion football team~~ runner-up)
2. You can exclude a word from your search by putting a minus sign (-) immediately in front of the term you want to avoid
 - a. eagles -football -philadelphia
 - b. eagles NOT football NOT philadelphia

E. Rule #5: Use the vertical bar (|) or OR to find pages containing either of two or more terms

1. A search for lodging dining atlanta: gets pages referring to BOTH lodging AND dining in Atlanta returns 233,000 hits
2. A search for lodging OR dining Atlanta: gets pages referring to EITHER lodging OR dining in Atlanta returns 2,900,000 hits
3. Can you tell why there are more hits with the OR?

F. Rule #6: Use the tilde (~) to find pages containing both the term and its synonyms

1. For example, to search for food facts as well as nutrition and cooking information, use:
 - a. ~food ~facts

G. Rule #7: Use n1 ... n2 to search for numbers in a specified range

1. For example, to search for a digital camera costing between \$250 and \$500:
 - a. digital camera \$250...\$500
2. To search for Olympic Game records for the years 1950 to 1960:
 - a. olympic game records 1950...1960

H. Rule #8: Use an asterisk (*) as a wild card representing one word

1. For example, to search for "Google changed my life," "Google enhanced my life," "Google ruined my life," etc. use:
 - a. google * my life

I. Rule #9: A term with an accent doesn't match a term without an accent, but the reverse is not true

1. Gödel matches *Gödel* but not *Goedel* nor *Godel*
2. But Goedel matches both Gödel and Godel

J. Rule #10: Combine search terms as often as possible (see Rule #1)

IV. The Action Buttons

A. Google Search

1. For a search like: disney fantasyland pirates, Google would use
2. Word hits: Whether the search terms appear on the page
 - a. It would identify all pages that contain *disney* AND *fantasyland* AND *pirates*
3. Adjacency hits: How close together the search terms appear on the page
 - a. A page that says: "My favorite *Disney* attraction, outside of *Fantasyland*, is *Pirates* of the Caribbean" will receive a higher adjacency score than a page that

says

- b. "Walt **Disney** was a both a genius and a taskmaster. The team at WDI spent many sleepless nights designing **Fantasyland**. But nothing could compare to the amount of Imagineering work required to create **Pirates** of the Caribbean."
4. **Frequency hits: How often the search terms appear on the page**
 - a. A page that has the word **disney** four times, **fantasyland** three times, and **pirates** seven times would receive a higher weights score than a page that only has those words once.
5. **About 100 other secret variables**
6. **Then, Google applies Page Rank**
 - a. There is a premise in higher education that the importance of a research paper can be judged by the number of times it is cited or referenced in other papers
 - b. Google simply applies this premise to the Web: the importance of a web page can be judged by the number of hyperlinks pointing to it from other pages
 - c. A page with a lot of (incoming) links to it is judged to be more important than a page with only a few links to it
 - d. A page with few (outgoing) links to other pages is judged to be more important than a page with links to a lot of other pages
7. **The pages with the highest ranks appear first on the output list**

B. I'm Feeling Lucky

1. **Displays only the web page with the highest page rank**

V. The Results

A. In a color bar stretching across the top of the page, Google displays

1. **The type of search**
2. **The number of hits (total and displayed)**
3. **The elapsed time of the search**

B. Down the right hand side of the page, Google displays

1. **Paid advertisements (Sponsored Links) related to the search subject**

C. In the main body of the page Google displays, for each item found,

1. **The page title as a hyperlink**
2. **Quotes from the text of the page containing the search terms**
3. **The URL of the page**
4. **Links to cached versions of the page and to similar pages**

VI. Types of Searches

NOTE: All of these searches use the seven rules discussed earlier

A. Web (www.google.com)

1. **This is the standard type of search discussed above**
2. **Currently, Google searches over 8 billion web pages**
3. **It searches the URL and the title as well as the text**
4. **If the searched-for text appears on more than one page in the same web site, each page is displayed separately**
5. **But they may not be contiguous due to differing page rank**

B. Images (www.images.google.com)

1. Searches for the web for pictures related to the specified search terms
2. Displays the image (as a thumbnail), the file name, the file size (in pixels and kilobytes) and the URL in which the picture is located
3. Double-click on the thumbnail to see the picture in its original web site context
4. Click "See Full-Size Image" to (guess what?) see a full-size version of the picture
5. Right click on the picture and choose "Save Picture As" or "Save Image" to copy it to your hard drive

C. Groups (www.groups.google.com)

1. Searches news groups for the specified search terms
2. A newsgroup is a forum for sharing information. Each one is an area on the Internet that allows users to post messages and reply to other users.
3. Thus, searches on news groups are more likely to return opinions and tips than is the case with standard web pages

D. News (www.news.google.com)

1. Searches current news sources for the specified search terms
2. Accesses many news sources (currently 4,500). Groups and summarizes the results for easy browsing
3. Results can be sorted by relevance or by date
4. A news search with a blank search box returns current headlines
5. Clicking on the link labeled News Alerts give you the opportunity to obtain news on a chosen topic from Google on a regular basis (once a day, as it happens, once a week)

E. Froogle (www.froogle.com)

1. Searches for prices and product availability associated with the specified search terms
2. Accesses product reviews and facilitates price comparisons

F. Local (www.local.google.com)

1. Enter the type of business and the city and state
2. Find local businesses and services on the Web

G. "More" (www.google.com/options/index.html)

1. **Answers (www.answers.google.com)**
 - a. Ask a question, tell what you are willing to pay for the answer
 - b. More than 500 carefully screened Researchers are ready to answer your question for as little as \$2.50 -- usually within 24 hours.
2. **Catalogs (www.catalogs.google.com)**
 - a. Search and browse mail order catalogs online
 - b. Choose a catalog from the directory listing and view the pages in full color
 - c. You can also search for specific items in the catalog
3. **Directory (www.google.com/dirhp)**
 - a. Organizes the Web by topic into categories
 - b. "Look Up," rather than search, for information on the Web
 - c. Works like Yahoo and other directories

- 4. Google Labs (www.labs.google.com)**
 - a. Provides an access to the latest ideas in the Google pipeline (see below)
 - b. Changes on a regular (sometimes daily) basis
- 5. Google Scholar (<http://scholar.google.com>)**
 - a. Google Scholar enables you to search specifically for scholarly literature, including peer-reviewed papers, theses, books, preprints, abstracts and technical reports from all broad areas of research.
 - b. Use Google Scholar to find articles from a wide variety of academic publishers, professional societies, preprint repositories and universities, as well as scholarly articles available across the web.
- 6. Special Searches (www.google.com/options/specialsearches.html)**
 - a. Uses a Google query word (related:) to search for sites specifically related to
 - (1) Microsoft
 - (2) Apple MacIntosh
 - (3) Linux
 - (4) BSD Unix
 - (5) Government
- 7. University Searches (www.google.com/options/universities.html)**
 - a. Uses a Google query word (site:) to search for sites related to specific colleges and universities
- 8. Wireless (www.google.com/options/wireless.html)**
 - a. Google's search technology can be accessed from any number of devices, such as mobile phones, Palm VII handhelds. Whatever language or platform you're using, Google lets you search the web with ease, speed and accuracy.
 - b. You'll find instructions at <http://www.google.com/options/wireless.html>
- 9. Blogger (www.blogger.com/start)**
 - a. "Blogger" is a word which combines two other words: Web and Log.
 - b. A "Blog" is a web site where you can post personal opinions about anything you wish as often as you wish and others who visit the site to respond to your opinions
 - c. Google provides the technology to enable you to create your own "Blog"
- 10. Google Toolbar (<http://toolbar.google.com>) and Browser Buttons (www.google.com/options/buttons.html)**
 - a. If you use Internet Explorer, you can use a Google Toolbar
 - (1) This toolbar is visible on your screen regardless of what web site you are viewing
 - (2) It gives you access to many of the types of searches described above
 - (3) Provides a pop-up blocker and automatic form completion
 - (4) It's free
 - (5) It's available in 13 languages
 - b. If you use some other browser, you can use the Google Browser Buttons
 - (1) They are visible on your screen regardless of what web site you are viewing
 - (2) Give you access to many of the types of searches described above
 - c. Both of these tools involve downloading software from Google onto your hard drive

11. Google Desktop Search (<http://desktop.google.com>)

- a. Google Desktop Search indexes files on your computer (similar to the ways in which it indexes web pages) and performs searches on them
- b. Find your email, files, web history and chats instantly
- c. View web pages you've seen, even when you're not online
- d. Google Desktop Search finds entries in:
 - (1) Outlook and Outlook Express
 - (2) Internet Explorer
 - (3) Word, Excel and Powerpoint
 - (4) AOL Instant Messenger
 - (5) Text files
- e. This tool involves downloading software from Google onto your hard drive

12. Language Tools (www.google.com/language_tools)

- a. Google in Your Language lets you view the Google home page in any of many available languages including Pig Latin, Elmer Fudd and Klingon
- b. It also lets you search for pages created in any of the available languages or registered in any of the available countries
- c. Translate Text lets you translate words, phrases or entire web pages to and from English, French, German, Spanish, Portuguese or Italian

13. Keyhole (www.keyhole.com)

- a. Keyhole is a web site, recently purchased by Google which provides access to aerial photos of every inch of the globe
- b. Google purchased Keyhole as part of its goal of "organizing the world's information and making it universally accessible and useful"
- c. But it hasn't announced its specific plans as yet
- d. You can purchase the Keyhole software for \$29.95 at www.keyhole.com

14. Picasa Photo Organizer (www.picasa.com)

- a. Picasa is a software program, recently purchased by Google, that helps you instantly find, edit and share all the pictures on your PC.
- b. Every time you open Picasa, it automatically locates all your pictures (even ones you forgot you had) and sorts them into visual albums organized by date with folder names you will recognize.
- c. It enables you to share your pictures – you can email, print photos home, make gift CDs, and even post pictures on your own blog.
- d. It is a free download from Google at www.picasa.com/index.php

15. Web API's (www.google.com/apis)

- a. An API (Application Program Interface) allows you to develop Your Own Applications Using Google
- b. With the Google Web APIs service, software developers can query more than 4 billion web documents directly from their own computer programs.
- c. For more information on this and other advanced Google topics see "Google Hacks" by Tara Calashain and Rael Dornfest, O'Reilly & Associates Inc ISBN: 0596004478

VII. Advanced Search Techniques

A. There are several kinds of advanced search techniques available in Google:

1. Special Search Features
2. Query Words
3. The Advanced Search Screen

B. Special Search Features (<http://www.google.com/help/features.html>)

1. Cached Links (query word – cache:)

- a. Shows the version of the page that Google last found
- b. Useful when attempting to retrieve copies of pages no longer active on the web
- c. Example: “cache:yahoo.com” displays the most recent version of the Yahoo page that Google has found

2. Calculator

- a. Google works as a calculator and unit converter
- b. Enter the calculation you'd like done into the search box and hit the Enter key or click on the Google Search button.
- c. The calculator can solve math problems involving basic arithmetic, more complicated math, units of measure and conversions, and physical constants.
- d. Examples:
 - (1) 5+5*2 returns 15
 - (2) (5+5)*2 returns 20
 - (3) 2^20 returns 1,048,576
 - (4) half a cup in teaspoons returns 24 US teaspoons
 - (5) au/c returns 8.31675359 minutes

3. Definitions (query word – define:)

- a. Google can function as a dictionary
- b. Enter the word you'd like defined preceded by the query word define:
- c. Then press the Enter key or click the Google Search button
- d. Example: “define:jackass” returns two definitions of the word
 - (1) A man who is a stupid incompetent fool
 - (2) A male donkey

4. File types (query word – filetype:)

- a. Google will limit its search to certain types of documents rather than Web pages (or it can exclude certain types of documents from its search)
- b. Google recognizes pdf, Microsoft Office formats, PostScript, Corel WordPerfect, Lotus 1-2-3, and other file types
- c. Google also offers the user the ability to “View as HTML” so you can examine the contents of a file even if the corresponding application is not installed on your computer
 - (1) Example 1: “budget filetype:xls” would return all the Microsoft Excel spreadsheets on the Web which contained the word “trenton”
 - (2) Example 2: “spyware filetype:-pdf” would return all the web sites containing the word “spyware” that were NOT in pdf format

5. Phone Book (query word – phonebook: rphonebook: bphonebook:)

- a. phonebook: searches all of Google's listings
- b. rphonebook: searches residential listings only
- c. bphonebook: searches business listing only

- d. You can access the information by entering various combinations of:
 - (1) LastName
 - (2) FirstName
 - (3) City State ZipCode
 - (4) Phone Number (including area code)
- e. Some examples
 - (1) rphonebook:may 08560 (returns information on everyone with the last name of May living in zip code 08560)
 - (2) bphonebook:pizza pennington nj (returns information on all the pizza parlors in Pennington, NJ)
- f. Note that for each listing returned you can access Yahoo! Maps or Mapquest to get a map of the location

6. Search by Number

- a. Parcel tracking IDs, patents and other specialized numbers can be entered into Google's search box for quick access to information about them.
- b. Examples
 - (1) UPS tracking numbers (example search: "1Z9999W9999999999")
 - (2) FedEx tracking numbers (example search: "9999999999999")
 - (3) USPS tracking numbers (example search: "9999 9999 9999 9999 9999 99")
 - (4) Vehicle ID (VIN) numbers (example search: "AAAAA999A9AA99999")
 - (5) UPC codes (example search: "073333531084")
 - (6) Telephone area codes (example search: "650")
 - (7) Patent numbers (example search: "patent 5123123") Note that the word "patent" is required as part of the search term
 - (8) FAA airplane registration numbers (example search: "n199ua")
 - (9) FCC equipment IDs (example search: "fcc B4Z-34009-PIR") Note that the word "fcc" is required as part of the search term

7. Similar Pages (query word – related:)

- a. Finds pages similar to the page specified in the search term
- b. Enter a URL to find pages that have the same sort of information
 - (1) Example 1: "related:aarp.org" displays all pages with information similar to that on the AARP home page
 - (2) Example 2: "related:Microsoft.com" displays all pages of information similar to that on the Microsoft home page
- c. As you can see from the second example, this is the technique used by google for the Special Searches for Microsoft, Linux, etc. we discussed earlier

8. Site Search (query word – site:)

- a. Limits the search to a specified domain
 - (1) Example 1: "courses site:www.ewingsnet.com" finds pages with the word "courses" on the Ewing SeniorNet site
 - (2) Example 2: "admissions site:www.stanford.edu" finds pages with the word "admissions" on the Stanford University site
- b. As you can see from the second example, this is the technique used by Google for the University Search we discussed earlier

9. Spell Checker

- a. Google's spell checking software automatically looks at your query and checks to see if you are using the most common version of a word's spelling
- b. If it calculates that you're likely to generate more relevant search results with an alternative spelling, it will ask "Did you mean: (more common spelling)?"
- c. Clicking on the suggested spelling will launch a Google search for that term

10. Stock Quotes (query word – stocks:)

- a. Entering the query word stocks: followed by selected stock symbols returns current prices along with much additional information on the stocks chosen
- b. Example: “stocks: goog msft” returns information on Google and Microsoft

11. Street Maps

- a. To use Google to find street maps, enter a U.S. street address, including zip code or city/ state in the Google search box.
- b. Often, the street address and city name will be enough.
- c. Example: 230 Jacobs Creek Road Titusville NJ will return a map to Joel May's house

12. Travel Information

- a. To see delays and weather conditions at a particular airport, type the airport's three letter code followed by the word "airport."
 - (1) Example: “ewr airport” will return information about Newark International Airport
- b. To check the status of a U.S. flight, type the name of the airline followed by the flight number.
 - (1) Example: “US Airways 11” will return the status for the US Airways flight 11 from Barajas, Spain to Philadelphia

13. Who Links to You (query word - link:)

- a. Identifies all pages linking to the Web page you've specified
- b. Example: “link:ewingsnet.com” displays all the pages on the Web that link to the Ewing SeniorNet home page

C. Google Query Words (<http://www.google.com/help/operators.html>)

NOTE: All of these searches use the seven rules discussed earlier

1. daterange:

- a. Limits search to pages concerning a date or a range of dates using Julian dates
- b. Julian dates (abbreviated JD) are a continuous count of days and fractions since noon Universal Time on January 1, 4713 BCE. Almost 2.5 million days have elapsed since this date.
- c. A conversion tool can be found at (<http://aa.usno.navy.mil/data/docs/JulianDate.html>)
- d. Example: “joe montana” daterange:2450815-2451180” would find sites referring to Joe Montana and dates between January 1, 1998 and January 1, 1999

2. intext:, intitle:, and inurl:

- a. Intext searches only within the text of the pages
- b. Intitle searches only in the titles of the pages
- c. Inurl searching only in the URL of the pages
- d. Examples:
 - (1) intext:trenton (finds “trenton” only in the page text)
 - (2) intitle:pirates (finds “pirates” only in the page title)
 - (3) inurl:ewing (finds “ewing” only in the page URL)

3. info:

- a. A catch-all that provides information on the specified domain including
 - (1) Google's cache of the domain page (see cache: above)
 - (2) Web pages that are similar to the domain page (see related: above)
 - (3) Web pages that link to the domain page (see link: above)
 - (4) Web pages that contain the search term

D. The Advanced Search Screen (http://www.google.com/advanced_search)

- 1. Number of Results (from 10 to 100 per page)**
- 2. Find Results (Works like the six rules we talked about earlier)**
 - a. All of the words (AND)
 - b. Exact Phrase (Quotes)
 - c. At Least one of the words (OR)
 - d. Without the Words (NOT)
- 3. Language (Can limit search to pages composed in one of the 36 available languages) (Like the Language Tool)**
- 4. File Format (Can limit or exclude search to certain types of documents rather than web pages) (like the filetype: query word)**
 - a. Adobe Acrobat (.PDF)
 - b. Adobe Postscript (.PS)
 - c. Microsoft Word (.DOC)
 - d. Microsoft Excel (.XLS)
 - e. Microsoft PowerPoint (.PPT)
 - f. Rich Text Format (.RTF)
- 5. Date (Can limit search based on how recently the pages were updated)**
 - a. Past three months
 - b. Past six months
 - c. Past year
- 6. Numeric Range (Can limit search to pages concerning a range of numbers or dates) (like the daterange: query word but without the need to compute Julian dates)**
- 7. Occurrences (Can limit search according to where the search term(s) occur on the page)**
 - a. In the Title (like the intitle: query word)
 - b. In the URL (like the inurl: query word)
 - c. In the Text (like the intext: query word)
 - d. In Links to the page (like the links: query word)

E. Domain (Can limit or exclude search to a specific domain)(like the domain: query word)

- 1. If you know the website you want to search but aren't sure where the information is located within that site, you can use Google to search only that domain.**
- 2. Enter the domain name in the appropriate box on the Advanced Search page**

F. Safe Search (Allows filtering for “inappropriate” pages)

G. Froogle (Find products for sale and their prices – discussed earlier)

H. Page Specific Search

- 1. Find pages similar to the page - Enter a URL to find pages that have the same sort of information (like the related: query word)**
- 2. Find pages that link to the page – Enter a URL to find pages with links to the chosen page (like the links: query word)**

I. Topic Specific Searches

- 1. Local - Find local businesses and services on the web (discussed earlier)**
- 2. Catalogs - Search and browse mail-order catalogs online (discussed earlier)**
- 3. (the following use the “Find pages similar to the page” technology discussed above)**
 - a. Apple Macintosh - Search for all things Mac
 - b. BSD Unix - Search web pages about the BSD Unix operating system
 - c. Linux - Search all penguin-friendly pages
 - d. Microsoft - Search Microsoft-related pages
- 4. (the following use the “Domain” technology discussed above)**
 - a. U.S. Government - Search all .gov and .mil sites
 - b. Universities: Stanford, Brown, BYU, & more - Narrow your search to a specific school's website

J. Similar Advanced Search Screens with differing specialized options are available for:

- 1. Google News**
- 2. Google Groups**
- 3. Google Image Search**
- 4. Google Catalog Search**
- 5. Froogle**

VIII. Google Preferences (<http://www.google.com/preferences>)

A. Interface Language

- 1. Choose any of 104 (and growing) languages for the Google Home Screen**

B. Search Language

- 1. Limit your searches to pages written in certain languages**

C. Safe Search Filtering

- 1. Choose the level of filtering to use**
 - a. Strict
 - b. Moderate
 - c. None

D. Choose the number of results (from 10 to 50) to appear on each results screen

E. Results Window

- 1. Choose to open results in a new window (useful for browsers like Firefox that use tabbed windows)**

IX. Google Labs (<http://labs.google.com>)

- A. Google Labs is Google’s “technology playground”**
- B. Every Google employee is given one day a week to work on a project of her/his choice**
- C. Some of these ideas make their way into Google Labs where both Google technologists and general users can try them and comment on them**

- D. Those that succeed in being useful are added to the Google product mix**
- E. The list changes on a regular (sometimes daily) basis – check it out regularly for new features**