



Getting the Most out of Google

<http://www.google.com>

WHAT GOOGLE INCLUDES

- ❑ Over 5 billion pages
- ❑ Does not index: more than the first 101 kilobytes (17,000 words or 22 pages) of a document; dynamically generated Web pages ("invisible web")
**If no one links to a page, Google is unlikely to list it!*
- ❑ Re-indexes every 28 days, compared to 7-11 days for AlltheWeb (<http://www.alltheweb.com>)
- ❑ "National" versions of Google: Google in U.S. (<http://www.google.com>) versus Google in U.K. (<http://www.google.co.uk>) or Germany (<http://www.google.de>)

SEARCH TECHNIQUES

- ❑ Google does not use natural language searching, so do not use a question as your search statement
"effect of bilingual education on students' test scores" ⇒ bilingual education AND test scores
- ❑ Do not use non-essential words when searching; also avoid words that a writer on the topic would not necessarily use
"articles about," "discussion of," "documentation on"
born vs. birth ⇒ "John F. Kennedy was born"
- ❑ Not case sensitive: *New Mexico = new mexico*
- ❑ Phrase searching: Need to put quotation marks around units you want treated as such
"human cloning" vs. human cloning
- ❑ AND is default operator for connecting words (unless there are quotation marks)
iraq "weapons of mass destruction" = iraq AND "weapons of mass destruction"
- ❑ Synonym operator (-) looks for the search term plus any of its synonyms
google -guide = guide, tutorial, etc.
- ❑ "I'm feeling lucky" button tells Google to go directly to the first result for the search instead of showing a list of results; good for when you think the page you want is the best result for your search
"Paul McCartney" + "I'm feeling lucky" gets
<http://www.paulmccartney.com>
- ❑ Google only looks at the first 10 words in the search box
- ❑ Common words (stop words) excluded from searches: the, who, of, are, etc.
**Unless they are placed in quotes: "to be or not to be," "the who"*
- ❑ Most punctuation marks and special characters are excluded: ? ; @, etc.
- ❑ Accents are used in searching: *Schröder Schroder*
- ❑ Use OR (|) to connect variants of a search term: *tahiti OR hawaii = tahiti | hawaii*
- ❑ Proximity searching: wildcard (*) will match any word in a phrase enclosed in quotes
*george * bush = George W. Bush*
*george ** bush = George H. W. Bush*
- ❑ Not (-) excludes items having certain word(s) from the results list: *dolphins -football*
- ❑ Advanced search page: Can limit search by language, file format (pdf, ps, doc, ppt, rtf), date, occurrences (anywhere, title, text, url, links), domain, and No filtering/Filtering using SafeSearch
- ❑ Page-specific search: From a given URL, can find other pages that are similar to or link to that page
- ❑ "Search within results" option: Can do new search within only results of prior one
- ❑ Can search Web, images, groups, directory, or news
- ❑ Number range syntax: Using two periods (..) between two numbers finds any number within that range
1776..1781=1777, 1778, 1779, 1780

RESULTS LISTS

- ❑ Spelling corrections suggested before start of results list **Suggestions NOT always right*
- ❑ Google automatically does stemming
child bicycle helmet finds child, children, children's, bicycle, bicycles, bicycle's, bicycling, bicyclists, helmet, helmets
**Use quotes or + to find exact word(s)*
- ❑ Ranks by relevancy, but relevancy calculated by 100 or so closely guarded algorithms
 - ❑ Repeating same search word can change results
 - ❑ Results that have search terms nearer each other higher than other results
snake grass finds the plant of that name before pages with "snake in the grass"
 - ❑ Results that have search terms in the same order as your search terms rank higher
 - ❑ PageRank lists results by the number and nature of the other pages linking to a page. Capable of being manipulated!
- ❑ Preferences lets you customize the language of results, or number of results per page
- ❑ Cache option: Google makes a copy of each page it indexes and stores it on a Google server. Google's result list links to the "live" version of the page on its own server—but you can view the cached version to (1) get a better sense of why you found a page (it could have changed between the time it was indexed and the time you searched) and (2) speed up response time.
**Page owners can request pages be removed from cache. Does not cache more than 101 kilobytes of a page.*
- ❑ Can view pages in other formats as HTML. Often a good idea for security reasons.
- ❑ Paid ads are listed ONLY in the column on the right side of the page titled "Sponsored Links"

ADDITIONAL FEATURES

- ❑ Link to dictionary; definition option
- ❑ Translation option
- ❑ Calculator option
- ❑ Stock quotes
- ❑ Travel conditions: Enter three letter code followed by word airport: *HLN airport*
- ❑ Search by number:
- ❑ Google Answers: Can ask a question and get an answer from a person. People answering are NOT necessarily experts. You specify a fee of \$2-200 per question answered and Google adds a \$0.50 listing fee.
- ❑ Google Alert (<http://www.googlealert.com>): Runs specified Google search every day and email results
- ❑ Google Viewer (<http://labs.google.com/gviewer.html>): will show your results full-screen, one after the other, every 5 seconds
- ❑ Searching "inside" selected books: <http://print.google.com> –looks at some 6,000 books, mostly Random House

GOOGLE-RELATED THINGS

- ❑ For Google results updated today, yesterday, within the last 7 days, or within the last 30 days, use FreshGoo.com (<http://www.FreshGoo.com>)
- ❑ Can search Google results between two specific dates at Fagan Finder's Search Tool (<http://www.faganfinder.com/engines/google.shtml>)
- ❑ Google Deskbar: Direct access to Google from the toolbar of your Web browser (for Windows, <http://toolbar.google.com/deskbar>) (for Mozilla or Netscape, <http://googlebar.mozdev.org>)
- ❑ Browser Buttons: Add buttons to your browser's toolbar (<http://www.google.com/options/buttons.html>)
- ❑ "Find in library option" when searching books: Links from Google to OCLC, creators of WorldCat
- ❑ Soople (<http://www.soople.com>) as an easy way to do advanced searches on Google
- ❑ MoreGoogle (<http://www.moregoogle.com>): Gives thumbnails of results pages