

How does Google search for everything?

Searching For and Organizing Data

Prof. Susan Rodger
Computer Science Dept
Duke University
Oct. 29, 2010



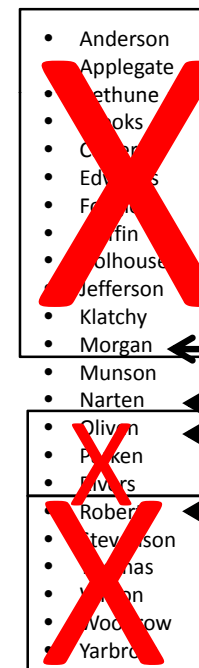
[Advanced Search](#) [Language Tools](#)

Lots of names. How do we find someone?

- Anderson, Mary : 203 Main St. Durham NC
- Williams, Fred : 14 Union Circle, Cary, NC
- Wu, Xin : 57 Wilson Court, Raleigh, NC
- Smith, Doug : 18 Pine Cone Lane, Durham, NC
- Pratt, Sarah: 6 White Lane, Hillsborough, NC
- Chase, Angela: 34 Dogwood Road, Durham, NC
- Brooks, Bolton : 10 Time St., Durham, NC
- French, Melvin : 42 Starship Circle, Durham, NC
- Gao, Bo : 134 Brookside Lane, Durham, NC

Put the names in alphabetical order

- Anderson, Mary : 203 Main St. Durham NC
- Brooks, Bolton : 10 Time St., Durham, NC
- Chase, Angela: 34 Dogwood Road, Durham, NC
- French, Melvin : 42 Starship Circle, Durham, NC
- Gao, Bo : 134 Brookside Lane, Durham, NC
- Pratt, Sarah: 6 White Lane, Hillsborough, NC
- Smith, Doug : 18 Pine Cone Lane, Durham, NC
- Williams, Fred : 14 Union Circle, Cary, NC
- Wu, Xin : 57 Wilson Court, Raleigh, NC



Find Narten

Found!

How many words did we look at?

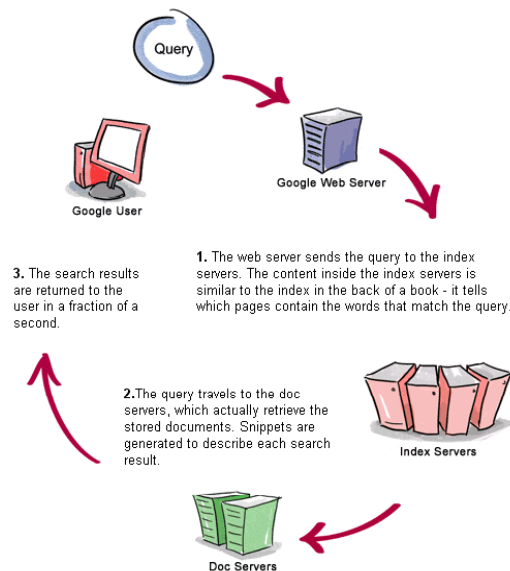
Searching for words

- If we had a million words in alphabetical order, how many would we need to look at worst case to find a word?
- About 20

How does one search for an item?

- Data must be organized in some way
- Sorting alphabetically (or numerically) is one way
- There are other ways to organize data!

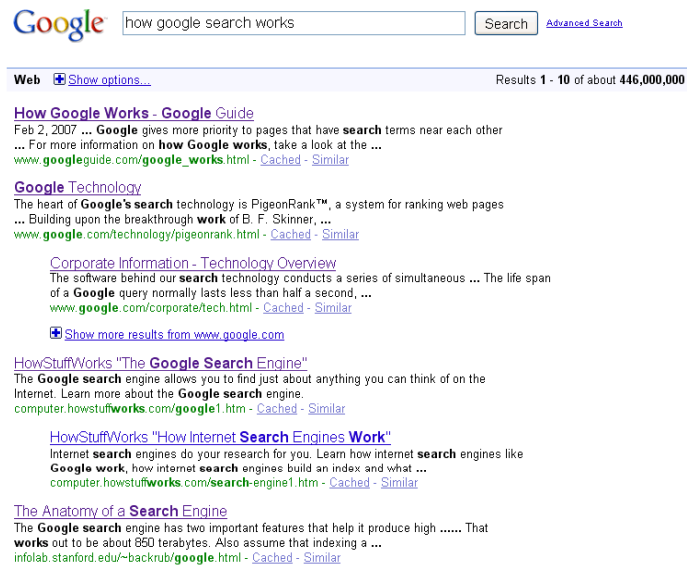
Google Search Query



Computer Science at work behind the scenes!

- Googlebot web crawler
 - Finds and retrieves pages
 - Gives pages to google indexer

“how” “google” “search” “works”



Google how google search works Search Advanced Search

Web Show options... Results 1 - 10 of about 446,000,000

How Google Works - Google Guide
Feb 2, 2007 ... Google gives more priority to pages that have search terms near each other ... For more information on how Google works, take a look at the ...
www.googleguide.com/google_works.html - [Cached](#) - [Similar](#)

Google Technology
The heart of Google's search technology is PigeonRank™, a system for ranking web pages ... Building upon the breakthrough work of B. F. Skinner, ...
www.google.com/technology/pigeonrank.html - [Cached](#) - [Similar](#)

Show more results from www.google.com

HowStuffWorks "The Google Search Engine"
The Google search engine allows you to find just about anything you can think of on the Internet. Learn more about the Google search engine.
computer.howstuffworks.com/google1.htm - [Cached](#) - [Similar](#)

HowStuffWorks "How Internet Search Engines Work"
Internet search engines do your research for you. Learn how internet search engines like Google work, how internet search engines build an index and what ...
computer.howstuffworks.com/search-engine1.htm - [Cached](#) - [Similar](#)

The Anatomy of a Search Engine
The Google search engine has two important features that help it produce high That works out to be about 650 terabytes. Also assume that indexing a ...
infolab.stanford.edu/~backrub/google.html - [Cached](#) - [Similar](#)

Page Rank Algorithm



Google miley cyrus Search Advanced Search

Web Show options... Res

News results for miley cyrus

 **Billy Ray Cyrus Raps for Miley's Twitterers** - 11 hours ago
Miley Cyrus rapped about leaving Twitter, so why shouldn't Billy Ray Cyrus rap about stickin' around? Anything Miles can do, dad can do better, right? ...
[El Online - 276 related articles >](#)

Miley Cyrus raps about why she left Twitter - CTV.ca - 914 related articles >
Are Miley Cyrus & Penelope Cruz filming with the ladies of Sex ... - Examiner.com - 66 related articles >

MileyCyrus.com - Step Into... a Miley State of Mind
Features news, pictures, audio, video, biography, events, and information on her fan club.
www.mileycyrus.com/ - [Cached](#) - [Similar](#)

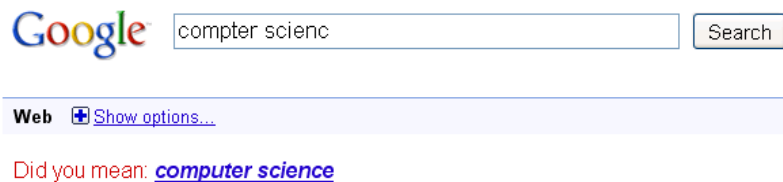
Miley Cyrus - Wikipedia, the free encyclopedia
Miley Ray Cyrus (born Destiny Hope Cyrus; November 23, 1992) is an American pop-rock singer and television and film actress. Cyrus is best known for ...
[Early life](#) · [Career](#) · [Entrepreneurship](#) · [Philanthropy](#)
en.wikipedia.org/wiki/Miley_Cyrus - [Cached](#) - [Similar](#)

Miley Cyrus
John Mayer Endorses Miley Cyrus' Exit From 'The Hate-Rix': The MTV News Quote Of The Day Find where Miley Cyrus is credited alongside another name ...
www.imdb.com/name/nm1415323/ - [Cached](#) - [Similar](#)

Image results for miley cyrus - [Report images](#)



Correction Algorithms



Google computer scienc Search

Web Show options...

Did you mean: [computer science](#)

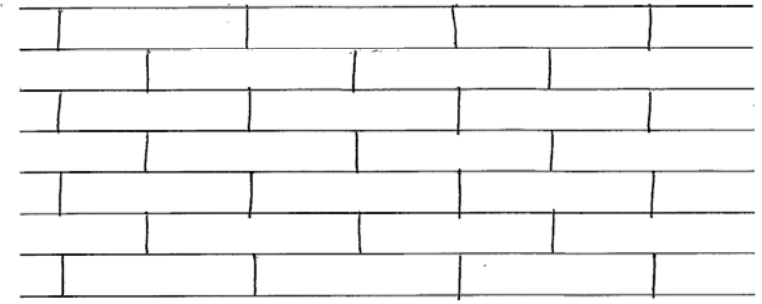
Google is all about problem solving and writing algorithms

- Algorithms must happen fast!
- Can Google put all the web pages it finds in alphabetical order to search?
- Want efficient, fast algorithms!
- No one wants to wait on a search query!

Activities

- Given numbers – sort yourselves
- Redistribute numbers – sort using selection sort
- Parallel Sort
- Hashing with buckets
 - Hash function is last digit – remainder when you divide by 10

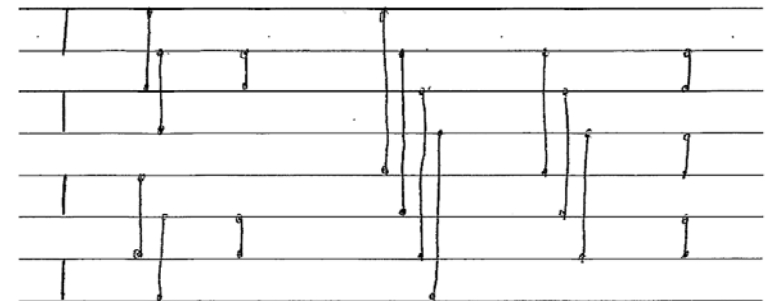
Sorting Network



Sort numbers (largest at bottom)
using comparators in parallel

8	4	4	4	4	2	2	1
4	8	6	6	2	4	1	2
7	6	8	2	6	1	4	3
6	7	2	8	1	6	3	4
3	2	7	1	8	3	6	5
2	3	1	7	3	8	5	6
5	1	3	3	7	5	8	7
1	5	5	5	5	7	7	8

Sorting Network
different setup for comparators



Sort numbers (largest at bottom)

8	4	4	4	1	1	1
4	8	7	6	2	2	2
7	6	6	7	3	3	3
6	7	8	8	5	5	4
3	2	1	1	4	4	5
2	3	3	2	6	6	6
5	1	2	3	7	7	7
1	5	5	5	8	8	8

My research - Making theoretical concepts come alive – visualize and interact with!

