

Search Engines

&

**how the world has
changed because of this
technology**

Kristen Scarpa

What are search engines?



“service on the web that allows searching of a large database of Web pages by word, phrase, or other criteria”

-Randolph Hock

What parts make up the search engines?



1.

Spiders or crawlers

These programs are used by search engines. They scan the Internet for new sites and existing sites that have changed.

2.

Index

is how the page is labeled once identified by the crawler. Usually by word, image, url, and metatags. This helps with speedy retrieval.

3.

Engine

A program that retrieves the pages that best match the search query. These are usually listed by relevance, popularity of page, number of times the search term appears on a page, and location of the terms on the page.

4.

HTML-based interface (Hypertext Markup Language)

gathers query info from the search and finds the correct page or pages

Archie

Alan Emtage

Our First Search Engine!



The screenshot shows the Archie Query Form interface. At the top, it says "Welcome to archie.icm.edu.pl". Below that is the title "Archie Query Form" with a magnifying glass icon. There is a search input field with the text "Search for: and" and a "Search" button. Below the search field are several options for search criteria:

- Database: Worldwide Anonymous FTP, Polish Web Index
- Search type: Sub string, Exact, Regular Expression
- Case: Insensitive, Sensitive
- Do you want to look up strings only (no sites returned): NO, YES
- Output Format for Web Index Search: Keywords Only, Excerpts Only, Links Only

changed the way the world shared files because the user could type in a search term and was able to navigate to the servers necessary

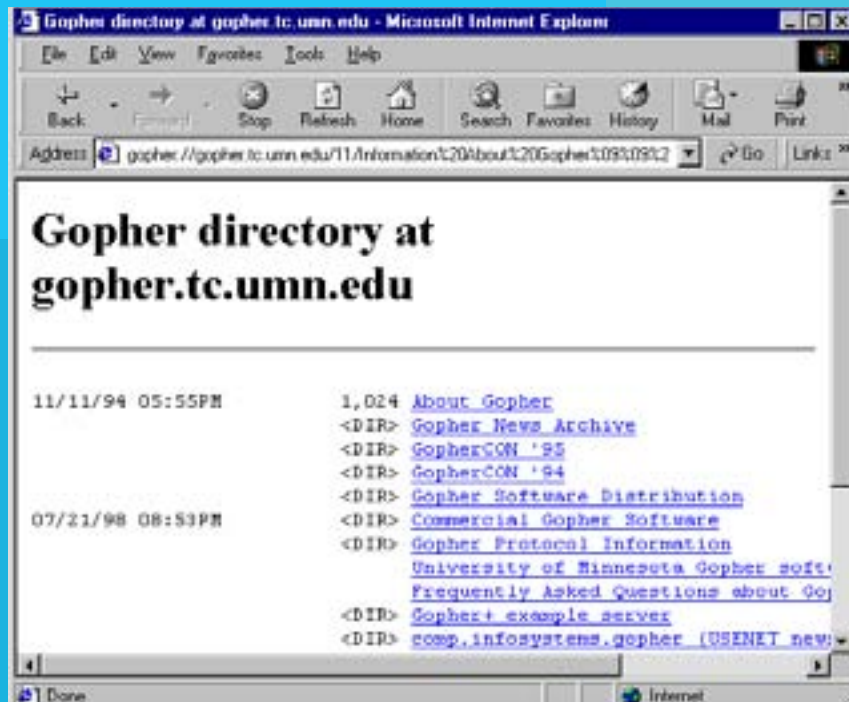
1. the user could type in a search term and was able to navigate to the necessary servers
2. files could be indexed
3. files didn't need to be found by word of mouth, an anonymous server could be found by using Archie

use ended in 1995

Gopher

Veronica

Jughead

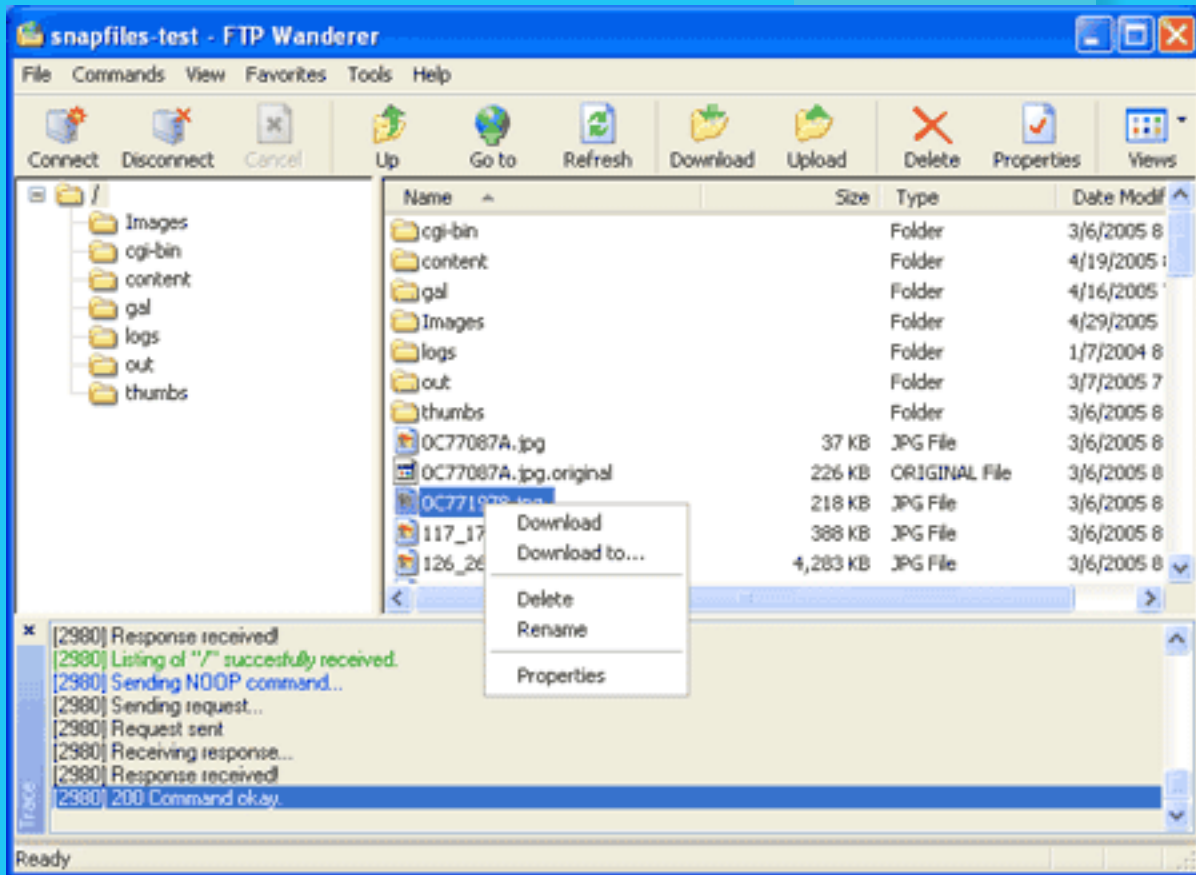


1. used to access documents on a server
2. Veronica (Very Easy Rodent-Oriented Netwide Index) and Jughead (Jonzy's Universal Gopher Hierarchy Excavation And Display) provided a keyword search of most Gopher servers
3. Jughead searches one server at a time and keeps each database in memory.
4. Veronica searches many Gopher servers at the same time

still in use today

World Wide Wanderer

Mathew Gray



1. first robot on the web

2. Wandex - first web database

3. this was very controversial because the robot would access the same sites around 100 times per day. This slowed the sites down for the end user

use ended in 1995

Aliweb

created by Martijn Koster



1. this search engine relied on webmasters to index their own sites, not robots

2. the search engine can only find the text that is within the index file and not in the the actual document

disadvantage - many of the webmasters didn't do this properly

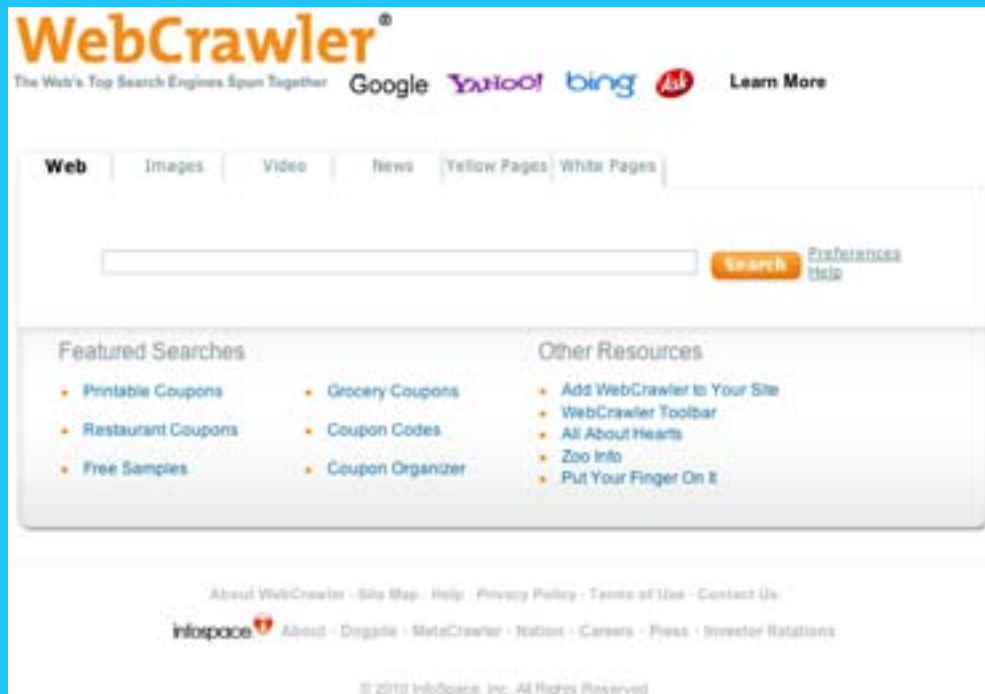
Altavista

Digital Equipment
Corporation



1. unlimited bandwidth at the time
2. recognizes natural language queries
3. use of advanced search and search tips
4. add or delete urls within 24 hours

Webcrawler



1. first crawler that indexed entire pages

2. was very slow during the daytime because there was so much web traffic

1994
Yahoo

David Filo and Jerry Yang



YAHOO!

1. man-made description with each url
2. info sites were added for free, but commercial sites at cost

yahoo.com

still in business today



1. provides ranked relevance

2. prefix matching and word proximity

3. was a large catalog that grew to 60 million documents by 1996

HotBot



1. very powerful search engine
2. can index 10 million pages per day
3. implemenation of cookie technology which stores personal search info

2003 - brought by Yahoo

MSN Search



1. 1998 - displayed search results from Inktomi and later blended results with Looksmart.
2. 1999 - AltaVista search results were used instead of Inktomi.
3. 2004 - created its own search results and not other search engine's results
4. used MSNbot, a crawler to create their own search results

2006 - the msn search was renamed Windows Live

1997
Ask.com



1. natural language search engine
2. human editors tried to index and create reliable searches
3. ranked links which resulted in a lot of spam

Cuil

Anna Patterson, Russell Power,
and Tom Costello

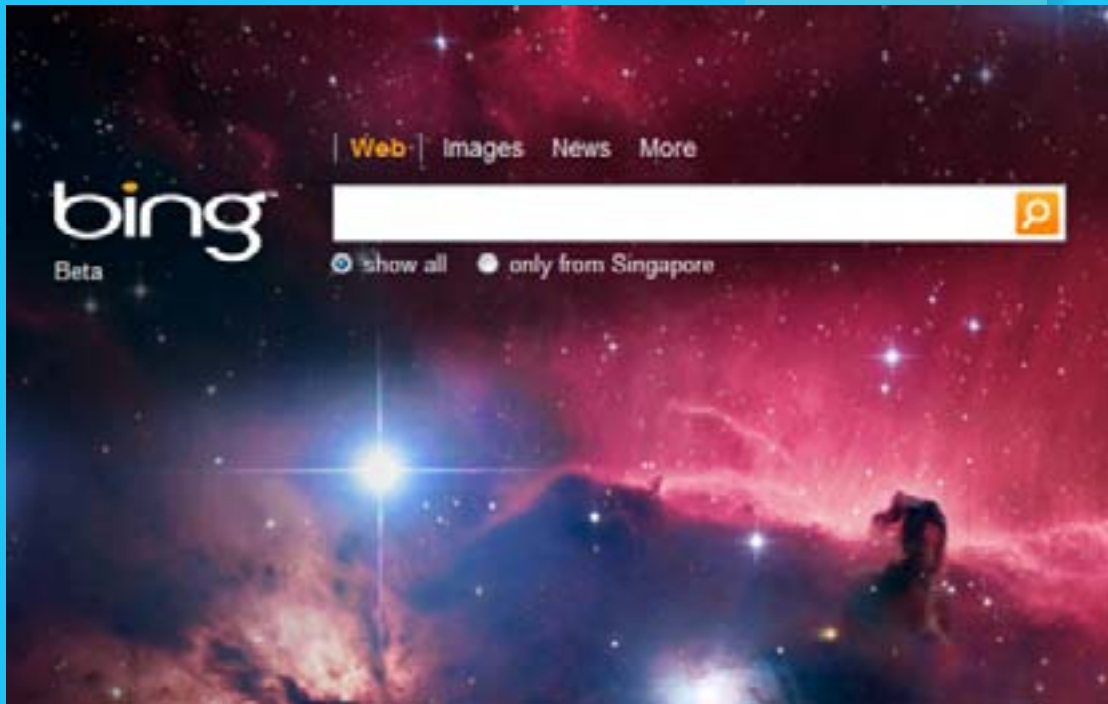


1. organized websites by content and showed images next to the searches

2. by 2009 there were 127 billion web pages

went out of business in 2010

Bing



1. gives search suggestions for related searches in the search results

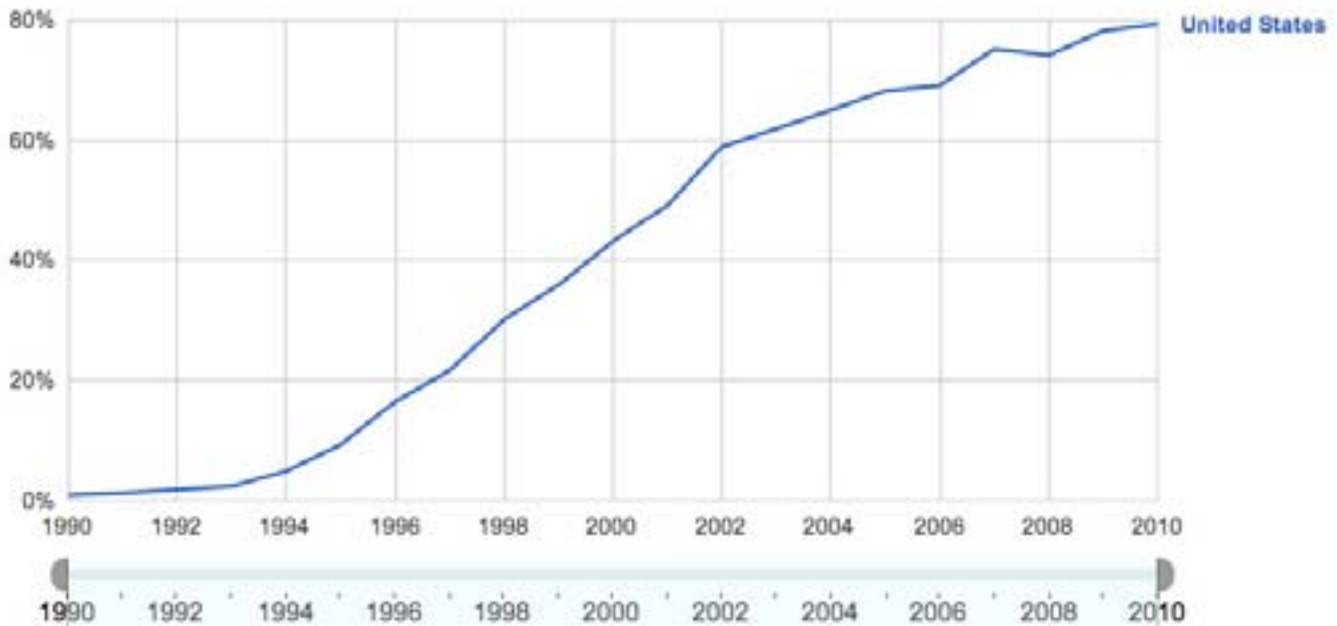
2. MSN and Live Search turned into Bing, which is now the engine for MSN.com

1990-2010

How much has internet usage grown in the last few years?



Internet users as percentage of population ?



Data from World Bank Last updated: Feb 16, 2012

©2012 Google - [Help](#) - [Terms of Service](#) - [Privacy](#) - [Disclaimer](#) - [Discuss](#)

Are there really positive or
negative effects or is it
subjective?

What are the Effects
of the growth of
search engines?



Negative?

1. way too much incoming information contributes to multi-tasking and less productivity
2. Too much information makes it difficult for someone to make a decision, stay on track, and see details
3. may be risking security: Business, banking, and finance are all online
4. people make mistakes when there is too much going on
5. lack of communication with the outside world
6. prevents people from going outside and exercising
7. decreased sense of well-being and eyesight and rise in blood pressure which creates cardiovascular stress
8. inappropriate advertisements

Positive!

- 1.** Online Banking can be very useful if the user takes necessary precautions such as the verification of a site key and use of a secure password. It is also necessary to be sure that the site starts with https, which defines the site as secure.
- 2.** People are in contact with others online and keep in touch with more people. They talk to friends who live in other cities and also meet new friends or romantic partners.
- 3.** People can work from home instead of going to the office, take fewer business trips, and drive less which in effect is saving the environment.
- 4.** It is a library at your fingertips. A person is bound to learn something new when surfing the web.
- 5.** Like-minded people can gather within online communities in order to either feel a part of and to share their beliefs.
- 6.** Artists gain a larger following by the circulation of their music.
- 7.** Anonymity creates a need to speak your mind. This can create social and political change.



Thank you