

BVCC General Meeting

July 12, 2021

Internet Searching

Joel Ewing, BVCC

Internet Searching

- **Finding content has been an issue since the early days of networking**
 - **Word of mouth**
 - **References from one site to another**
 - **Batch Unix tools created to search for sites with useful servers & search for content with specific keywords – could take hours or days and results served one person or group**
 - **Early search engines in 1990 before the "web" (Archie, Gopher, Veronica, Jughead)**
 - **early "web" indexes manually maintained or based on info from web site creators. Explosive growth of the web quickly made indexes using manual updating techniques impractical.**

Search Engines

- **Search engines now taken for granted – a Google search on "Search Engines" yields among others : Google, Bing, DuckDuckGo, CC Search**
 - **1st successful SE design/implementation: AltaVista in 1995 – created as demo of DEC AlphaServer computer**
 - **They do the exhaustive searches ("crawling" web sites with "spider" programs) and make results generally available to all**
 - **Supported by ad revenue from untargeted and targeted ads based on tracking user interests, and on advertising charges to boost ranking of a website**

Google Search

- **Not the first SE (at least 5 SEs of note in 1996), but this was the first (1998) to adopt a business model of selling search terms (very profitable).**
- **By 2000 Google became dominant with algorithms that achieved better results.**
- **In a few years "Google" became a verb. May 2021, usage: 92%**
- **While Google does display some results at the top based on paid advertising, they generally make clear that is why they are first.**

Microsoft Bing

- **Evolved from MSN Search (1998) using results from earlier search engines.**
- **MS developed their own web crawler (msnbot 2004)**
- **Rebranded as "Bing" (2009), Bing technology adopted by Yahoo Search**
- **Default for MS browsers**
- **May 2021 – 2.27% usage (StatCounter GlobalStats)**

DuckDuckGo

- **Unlike Google & Bing, does not return customized results based on user's activity history – lack of tracking user history is also one reason some are attracted to this SE**
- **May 2021 – 0.59% usage**

Other Search Engines

- **Yahoo – 1.5% (uses Bing for actual search)**
- **Baidu – 1.46% (heavily used in China)**
- **Yandex – 0.69% (62% usage in Russia)**
- **Others with significant usage in specific countries**
- **Some SEs deal with specialized searches (CC Search returns public domain content permitting unrestricted use), or allows indirect access to another SE to prevent personal tracking**

How Search Engines Work

- **A "spider" that does web crawling of pages on a web site**
 - **Rescan of previously known website**
 - **Scan of a website encountered in a link on a known website**
 - **Scan of a website formally "submitted" to the SE**
 - **New sites could be found checking defined domain names to see if a web server responds at that domain. (Don't know how often this is done)**

How Search Engines Work

- **Content of web pages are indexed**
 - Algorithms used are proprietary
 - Website designer can influence what pages are scanned and possibly influence their ranking by choice of words or suggested keywords (can be a source of useful information, or a source of mis-information)
- **Search of database uses proprietary ranking algorithms**
 - may be partially random and can be influenced by advertising or by user's prior searching history.

Bad-Actor Internet Crawlers

- **Scans of IP addresses for open webserver ports – looking for devices with security exposures to exploit**
 - **Could involve any device that uses a webserver for control if device exposed to Internet, including a router with remote config enabled**
 - **Hackers attempt to access specific files or directories to determine if dealing with a device with known exposures or a website with applications installed that have exposures.**
 - **40 – 80 daily probes on my home web server. Easy to detect potential abusers – I use unique directory names with no "home" page or links to content – one must know specific directory/file names to see content – attempts to access commonly used file names or directories associated with a web app will fail – any attempt to request website content from the Internet using just my IP address rather than my domain name will fail, as that technique would typically only be used by a hacker.**

Search Engine Access

- **Don't have to use browser default: Can always go directly to the SE website:**
 - **Google: www.google.com**
 - **Bing: www.bing.com**
 - **DuckDuckgo: www.duckduckgo.com**

Google Search Operators

- **All criteria influence ranking – not absolute**
- **Exact match: "some phrase"**
- **OR or | : x OR y, matching x, y, or both**
- **AND: x AND y (default, so same as x y)**
- **- : exclude, jobs -apple (jobs but not apple)**
- *** : wildcard match any word/phrase: ruehle * center**
- **() : grouping (ipad OR iphone) apple**
- **\$: search for prices: ipad \$329**

Google Search Operators

- **define:** – display word meaning `define:computer`
- **cache:** – return most recent cached web page if indexed `cache:bvcomputerclub.org`
- **filetype:** – restrict results to specific filetype
`paradise lost filetype:pdf`
- **site:** – show only results from specified website –
`site:bvcomputerclub.org officers`
- **related:** – find sites related to a given domain (not sure what this does)

Google Search Operators

- **intitle:** – find pages with certain word(s) in title –
intitle:bvcc **intitle:"bvcc history"**
- **allintitle:** – all words after ":" must be in title
allintitle:bvcc history
- **inurl:** & **allinurl:** – similar to above but words must be in URL for the page
allinurl:bvcomputerclub b
- **AROUND(x)** – find pages containing two words or phrases within x words of each other

Combining Search Operators

- **site:bvcomputerclub.org -inurl:https**
show any indexed pages found that are not secure
- **Checking for duplicated content within a website:**
site:bvcomputerclub.org "2nd Monday"
- **Checking for some phrase or sentence on your website being used elsewhere**
"Clinics, classes, the Genealogy" -site:bvcomputerclub.org
- **"Joel Ewing" OR "Joel C Ewing" OR "Joel C. Ewing"**
(try add -teacher -theater -actor -manager)
(try add arkansas

Bing Search Operators

- **term1 +term2** – look for term1 and term2 "must" be present
- **term1+term2** – exact phrase, both term1 and term2 must be present (same as quotation marks)
- **"word1 word2 etc"** – match exact phrase
- **term1 -term2** – exclude: look for term1, but must not have term2 – can also be written as **term1 NOT term2**
- **()** – grouping of other operators
- **term1 AND term2** – both terms present (the default) – can also use **&**
- **term1 OR term2** – either term1 or term2 or both present – can also use **|**
- **filetype:<type>** as in **filetype:pdf** – looks for documents based on specified file type (which could be different than file extension if file inconsistently created)
- **intitle: search-terms** – looks for terms in web page title

Bing Search Operators

- **ip: ip-address (Bing only)** searches websites at that address
- **language:xx (Bing only)** – searches pages with specific language code (en=English)
- **prefer:term (Bing only)** – give higher priority in listing order to sites with this key term
- **site:<domain name>** – limits search to pages from that specific domain and subdomains
- **url:<domain/subdomain name>** – limits search to a specific domain or subdomain, excluding other sub-domains
- **info:<domain name>** – return info about specified website
- **imagesize:<size>**, where size is "small", "medium", or "large – return images with different size ranges.
- **define:term** – give a word definition
- ***** – match any phrase
- **Others (but don't always work)**

DuckDuckGo Search Parameters

- **term1 term2** – results relevant to term1 or term2
- **"several words"** – exact match
- **term1 -term2** – fewer instances of term2 in results
- **term1 +term2** – more instances of term2 in results
- **filetype:xxx** (e.g. **filetype:pdf**) restrict results to file type (pdf, doc(x), xls(x), ppt(x), html)
- **site:example.com** – restrict to pages from example.com
- **-site:example.com** – exclude any pages from example.com

DuckDuckGo Search Parameters

- **intitle:word** – include pages with "word" in page title
- **inurl:word** – include pages with "word" in page URL
- **!xxxx word** – Search site designated by "xxxx" for "word". Have to consult "bang" shortcut table. Common sites have short alias (!a = search Amazon; !w = search Wikipedia)
- Can enable "safe search" by setting option or using "safe.duckduckgo.com" to remove adult content from search results.

Using a non-default Search Engine

- **Can always just go to Search Engine website:
www.google.com, www.bing.com
www.duckduckgo.com, etc.**

Changing Firefox SE Default

- **Menu → Settings → Search → (or click on search magnifier & Change Search Settings) and select a new default under "Default Search Engine"**
- **For a temp change, when you start to enter keywords in the search bar, select a different SE from alternative selections in pull-down list (can change which ones appear from "Search Settings")**

Changing Chrome SE Default

- **Right-click on address/search bar and select "Manage Search Engines" – can click on a new different SE and make it the default. Can also add additional search engines to the choices.**



- Undo Ctrl+Z
- Cut Ctrl+X
- Copy Ctrl+C
- Paste Ctrl+V
- Paste and search for "www.google.com, www.bing.com www.duckduckgo..."
- Delete
- Select all Ctrl+A
- Manage search engines... ●
- Always show full URLs

Search engine	Keyword	Query URL
Google (Default)	google.com	{google:baseURL}search?q=%s&{go...
Yahoo!	yahoo.com	https://search.yahoo.com/search(g...
Bing	bing.com	https://www.bing.com/search?q=%s...
DuckDuckGo	duckduckgo.com	https://duckduckgo.com
Ecosia	ecosia.org	https://www.ecosia.org

Other search engines Add

- Make default ●
- Edit
- Remove from list

Changing Edge SE Default

- **Go to ... (Settings & More) → Settings → Privacy, Search, and services → Services → Address bar and search >**
 - **Select new default for "Search engine used in the address bar". If desired SE not listed, new ones can be added by selecting "Manage search engines >". If you leave "Search on new tabs..." set to "Search box", the search box on a new tab will still go to Bing!**
- **Doesn't appear to be any way to change one search to a different SE other than going to the SE's website**

The image shows a Microsoft Edge browser window. The address bar is empty. The main content area displays a news article with a photo of two flags (Chinese and American) and a weather widget for Bentonville, Arkansas. The weather widget shows a current temperature of 71°F and a forecast for the next four days. The browser's menu is open, showing various options like 'New tab', 'History', 'Downloads', and 'Settings'. A red arrow points to the 'Settings' option in the menu.

News article text: "ia says its military 'drove y' US warship"

Weather: Bentonville, Arkansas, 71°F

Today	Tue	Wed	Thu
79° 61°	88° 71°	89° 72°	88° 73°

- New tab **Ctrl+T**
- New window **Ctrl+N**
- New InPrivate window **Ctrl+Shift+N**
- Zoom: - 100% + ↗
- Favorites **Ctrl+Shift+O**
- History **Ctrl+H**
- Downloads **Ctrl+J**
- Apps >
- Extensions
- Collections **Ctrl+Shift+Y**
- Print **Ctrl+P**
- Web capture **Ctrl+Shift+S**
- Share
- Find on page **Ctrl+F**
- Read aloud **Ctrl+Shift+U**
- More tools >
- Settings** ←
- Help and feedback >
- Close Microsoft Edge

Settings

Search settings

- Profiles
- Privacy, search, and services**
- Appearance
- On startup
- New tab page
- Share, copy, and paste
- Cookies and site permissions
- Default browser
- Downloads
- Family
- Languages
- Printers
- System
- Reset settings
- Phone and other devices
- About Microsoft Edge



Manage certificates

Manage HTTPS/SSL certificates and settings



Microsoft Defender SmartScreen

Help protect me from malicious sites and downloads with Microsoft Defender SmartScreen.



Block potentially unwanted apps

Blocks downloads of low-reputation apps that might cause unexpected behaviors.



Use secure DNS to specify how to lookup the network address for websites

By default, Microsoft Edge uses your current service provider. Alternate DNS providers may cause some sites to not be reachable.



Use current service provider

Your current service provider may not provide secure DNS

Choose a service provider

Select a provider from the list or enter a custom provider

Enter custom provider

Services

Microsoft Edge may use web services to improve your browsing experience. You can always choose to turn this off.

Use a web service to help resolve navigation errors



Suggest similar sites when a website can't be found

If a website can't be found, the web address will be sent to Microsoft to try to find the correct site.



Show suggestions from Pinterest in Collections ?

Search for curated suggestions from Pinterest that may be related to your collection.



Save time and money with Shopping in Microsoft Edge

We'll automatically find you the best prices, coupons and rebates from across the web as you shop.



Address bar and search

Manage search suggestions and search engine used in the address bar



← Privacy, search, and services / Address bar and search

Show me search and site suggestions using my typed characters

If you turn this off, you'll only see suggestions from your favorites and history. Typed characters won't be sent to the search engine used in the address bar.

Show me suggestions from history, favorites and other data on this device using my typed characters

Benefits, available only with Bing Search from Microsoft. [Learn more](#)



InPrivate search

Gain privacy without compromising on search results or relevancy with Bing InPrivate Search



Microsoft Rewards

Earn Microsoft Rewards points towards gift cards or nonprofit donations by searching with Bing



Enterprise search

Find sites and people quickly, even within your organization, with Microsoft's Bing search

Search engine used in the address bar

Search on new tabs uses search box or address bar

Manage search engines

Google (Default) ▾

Bing (Recommended)

Yahoo!

Google (Default)

DuckDuckGo

Google Quicksearch

Dictionary.com Quicksearch

Stock Symbol Quicksearch










Wikipedia Quicksearch

Urban Dictionary

Address bar search engines

Add

To see more search engines here, open a new tab, go to the search engine you'd like to add, and search for something

Search engine	Keyword	URL	
 Bing (Recommended)	bing.com	{bing:baseURL}search?q=%s&{...	...
 Yahoo!	yahoo.com	https://search.yahoo.com/sear...	...
 Google (Default)	google.com	{google:baseURL}search?q=%...	...
 DuckDuckGo	duckduckgo.com	https://duckduckgo.com/?q=%s	...
 Google Quicksearch	google	http://www.google.com/search...	...
 Dictionary.com Quicksearch	dict	http://dictionary.reference.com...	...
 Stock Symbol Quicksearch	quote	http://www.google.com/search...	...
 Wikipedia Quicksearch	wp	http://en.wikipedia.org/wiki/Sp...	...
 Urban Dictionary	slang	http://www.urbandictionary.co...	...

