🔎

Search engine investigation

List of search engines

- Baidu, China’s most popular search engine.
- Bing, a search engine owned and operated by Microsoft, and the second most popular worldwide. Supports advanced search keywords.
- binsearch.info, a search engine for binary Usenet newsgroups.
- Common Crawl, “an open repository of web crawl data that can be accessed and analyzed by anyone.”
- DuckDuckGo, a privacy-focused search engine that compiles results from many different sources. Supports search syntax.

- Google, which offers the world’s most popular search engine, and uses a ranking system to attempt to return the most relevant results. Supports search operators.
- Internet Archive Wayback Machine, “building a digital library of Internet sites and other cultural artifacts in digital form.”
- Startpage, a search engine that uses Google’s results without collecting personal information through trackers and logs. Supports search operators.
- Shodan, a service for searching Internet-connected devices and services. Usage options include a limited free plan as well as paid subscription plans.

Search operators

- `site:` will limit the search to the provided domain.
- `inurl:` will only return results that include the keyword in the URL.
- `intitle:` will only return results that have the keyword in the page title.
- `intext:` or `inbody:` will only search for the keyword in the body of pages.
- `filetype:` will match only a specific filetype, i.e. png, or php.
- `cache:` Search the cache
- `OR,AND, -, *, IN`
- `( )` Group terms
- `Related:` Find sites related to a given domain.

What are we looking for?

- Footholds
- Files containing usernames
- Sensitive Directories
- Web Server Detection
- Vulnerable Files
- Vulnerable Servers

- Error Messages
- Files containing juicy info
- Files containing passwords
- Sensitive Online Shopping Info
- New subdomains

Resources

- Google hacking database (https://www.exploit-db.com/google-hacking-database)
- SANS Cheat sheet: https://www.sans.org/security-resources/GoogleCheatSheet.pdf
- gbhackers: https://gbhackers.com/latest-google-dorks-list/
- pentest-tools.com: https://pentest-tools.com/information-gathering/google-hacking#
- https://owasp.org/www-project-web-security-testing-guide/latest/4-Web_Application_Security_Testing/01-Information_Gathering/01-Conduct_Search_Engine_Discovery_Reconnaissance_for_Information_Leakage

Techniques

Subdomain enumeration manual

- Site: google.com
- Site: google.com -www
- Site: google.com -www -blog
- Site: google.com -www -blog -mail -maps....

Finding exploits

- Go to GHDB
- Find dork that fits your targets

Look for files

- Site: google.com filetype:pdf
- Site: google.com filetype:csv
- Site: google.com filetype:...

Looking for backup files

- Site: google.com filetype:BAK
- Google Dork Query: *intitle:"index of" "backup.bak"*

Look for login panels

site:google.com inurl:login | inurl:signin | intitle:Login | intitle:"sign in" | inurl:aut

Directory listing

site:google.com intitle:index.of

Database files

site:google.com ext:sql | ext:dbf | ext:mdb

SQL errors

site:google.com intext:"sql syntax near" | intext:"syntax error has occurred" | intext:"incorrect syntax near" | intext:"unexpected end of SQL command" | intext:"Warning: mysql_connect()" | intext:"Warning: mysql_query()" | intext:"Warning: pg_connect()"

PHP errors and alerts

site:google.com "PHP Parse error" | "PHP Warning" | "PHP Error"

Signup pages

site:google.com inurl:signup | inurl:register | intitle:Signup

Fetch new endpoints

Site:www.google.com -inurl:index.php

Site:www.google.com -inurl:index.php -inurl:login.php

https://github.com/tomnomnom/waybackurls

Possibly interesting files

- PDF
- DOCX, DOC
- XLSX, XLS
- PPT, PPTX
- BAK
- SQL,DBF,MDB
- CONF
- YAML
- XML
- LOG