

# Basics of searching in free sources

## Internet - global information network

- worldwide system of interconnected computer networks (the so-called "network of networks")
- has no owner, no control centre that decides what, where and when is published on the network
- the operation and further development of the Internet is handled by non-governmental organizations (ISOC - Internet Society, W3C - WWW Consortium)
- it is free

## History of the Internet

- first vision of a computer network - 1946 in Murray Leinster's short story "A Logic named Joe"
- originally developed for government and military purposes
- 1969 - ARPANET - the first precursor - interconnecting 4 computer networks (USA)
- 1972 - first e-mail program
- 1973 - TCP (Transmission Control Protocol) - allows data transfer between networked computers (today - TCP/IP - a set of protocols for computer network communication)
- 1987 - The Internet was born
- 1989 - design of the development of the www and hypertext
- 1994 - commercialization of the Internet
- 2006 - 1 billion users

## Basic Internet services

- WWW
- electronic mail
- Instant messaging
- FTP
- VoIP (Voice over Internet Protocol)

## Internet as an information resource

- the selection of published information is not professionally edited or controlled
  - information is outdated, lacking dates of origin
  - poor quality, original source cannot be verified
  - no information on the origin of the source (original, adopted)

## Web - visible x invisible

- **visible web** (visible web, surface web) - freely searchable resources
- **invisible web** (invisible web, deep web) - resources whose content is not accessible by standard search engines (they are invisible to them) - databases, password protected sites, unlinked websites

## Visible Web Tools

- **subject catalogues'** (browsing) - logically manually sorted links
  - List (<https://www.seznam.cz/>), Yahoo (<https://www.yahoo.com/>)
- **search engines'** (browsing, searching) - components robot, index, search engine
  - Google (<http://www.google.cz>), ask.com (<https://www.ask.com/>), AltaVista (<https://www.altavista.com/>), now [www.cuil.com](http://www.cuil.com) (<http://www.cuil.com>)
  - **metasearch engines** - search multiple search engines in parallel
- Metacrawler ([http://omblockedips.com/?honeypot&params=QjGCa85uLILh52RB2yz1CM\\_BL349\\_5Bf0JNiUUqXazRPJ41IZt4TSqz2kaNH-RvYWDyA7hXo4CkWZSjHxJv-ngL-1vbakX7AZcKOOVjel-vgFYIx7P3McmKLizjeINISWnvJZOj-xi37csnnYkagZv9KqcckKO7Hp9UIRjLUnAUFXWaZgtUKYa0ukfWKZM5Z7GRDoP\\_8c1SJ1NptFHsACg](http://omblockedips.com/?honeypot&params=QjGCa85uLILh52RB2yz1CM_BL349_5Bf0JNiUUqXazRPJ41IZt4TSqz2kaNH-RvYWDyA7hXo4CkWZSjHxJv-ngL-1vbakX7AZcKOOVjel-vgFYIx7P3McmKLizjeINISWnvJZOj-xi37csnnYkagZv9KqcckKO7Hp9UIRjLUnAUFXWaZgtUKYa0ukfWKZM5Z7GRDoP_8c1SJ1NptFHsACg)), Dogpile (<http://www.dogpile.com/>)
- [www.vyhledavace.net](http://www.vyhledavace.net) (<http://www.vyhledavace.net>)

## Search Strategy Basics - Tips and Tricks

### Search phrases, wildcards

- quotes create a phrase
- wild card symbols - ?, \*, \$

### Boolean operators

- **AND'** (= all of the words) - narrows the query

- **OR** (= any of the words) - expands the query
- **NOT** (= removes unwanted references)

Ex. Locks *NOT* ((security *AND* device) *OR* locking *OR* FAB *OR* entry *OR* door *OR* door *OR* keys)'

### Proximity Operators

- Specify the sequence or distance between search terms
- **NEAR'** - specifies the maximum distance between words (e.g. x near/3 y → x and y will be max 3 words apart)
- **ADJ'** (adjacent) - words next to each other, but no matter the order
- **Followed'** by - words next to each other in the exact order

## Free Information Resources - Medicine

- [www.nlk.cz/nlkcz/infozdroj/infozdroj.php](http://www.nlk.cz/nlkcz/infozdroj/infozdroj.php) (<https://nlk.cz>)
- [pez.cuni.cz](http://pez.cuni.cz) (<http://pez.cuni.cz>)
- MediClub
- Zdrav.cz

## Internet today - Web 2.0

- The term describes the trends and new possibilities of the Internet
- Term first used 2004 by Tim O'Reilly
- Ross Mayfield - "Web 1.0 was commerce, web 2.0 is people"
- Users can do more than just get information

### Earlier

- web content is created by the owner
- updates as the owner allows
- there is no community - the visitor is a passive recipient of information

### Internet today - Web 2.0

- Visitors actively participate in content creation
- Interaction required - through discussions, chat, personal profile
- Web is a living organism - millions of content creators
- Emergence of communities - interest, professional - personal profile creation
- blogs, wiki systems
- AJAX - changes the content of their pages without having to reload them (maps - zoom)
- tagging, folksonomy
- RSS - on-demand notifications of news on selected pages

### Web 2.0 search engines

- [www.kartoo.com](http://www.kartoo.com) (<http://www.kartoo.com>)

## Links

### Related articles

- Information and Information Institutions
- Catalogues (1st LF UK, NT)
- Fundamentals of open source search
- Orientation to available e-resources
- Information resources at the UK
- Full-text electronic information resources
- Citing literature used
- Digital portfolio from an R&D perspective
- Written scholarly or professional communication

### Resource

- KRAJÍČKOVÁ, Jitka. *Basics of searching in free resources* [online]. [cit. 2012-03-15]. <<https://el.lf1.cuni.cz/p97169685/>>.