



FAKULTA SOCIÁLNÍCH VĚD  
*Univerzita Karlova v Praze*

# Internetové vyhledávače

Michala Sošková  
knihovna FSV



internetové vyhledávače  
google  
neviditelný web  
trendy



# jak fungují vyhledávače

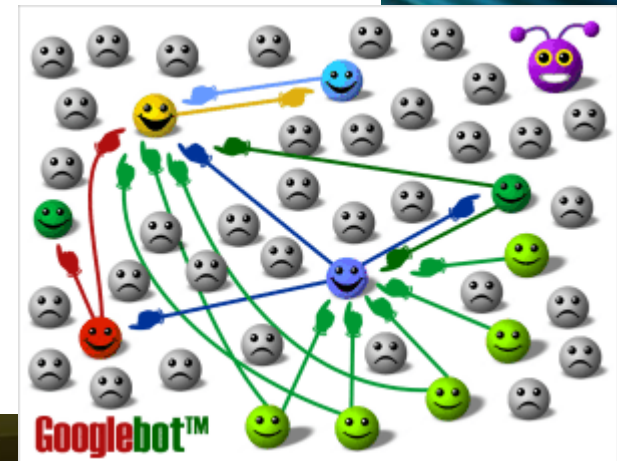
/ robot – spider/crawler

/ index



# robot

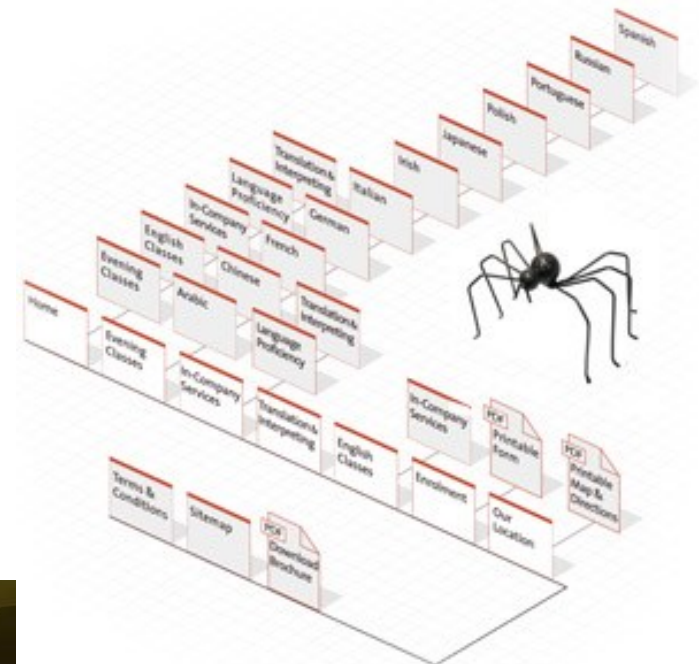
- / prochází stránky a **odkazy** na nich
  - > „mapa webu“
- / stáhne klíčová slova + informace o nich
  - > seznam všech slov = **index**
- / vrací se zpět od začátku





# vyhledání dotazu

- / při zadání dotazu vyhledávač neprohledává stránky, ale prohledává **index = seznam slov s jejich adresami + informací, zda jde o**
  - / **název**
  - / **podnázev**
  - / **titulek obrázku**
  - / **četnost v textu**





# index

## „jak správně vyhledávat na internetu“

- |  |                          |
|--|--------------------------|
| 1 internet-u (webu)                          | URL1, <b>URL2</b> , URL3 |
| <del>jak</del>                               |                          |
| <del>na</del>                                |                          |
| 2 správn- ě (-ý, -á,-í... vhodný, přesný...) | <b>URL2</b> , URL4, URL5 |
| 3 vyhled-ávat (-at, -ám...)                  | URL6, <b>URL2</b> , URL3 |



# index

## řazení výsledků ovlivňuje:

URL

titulek

podnadpis

četnost slov

začátek vět

.....



google...





# google

- / od r. 1998
- / 2 doktorandi Kalifornské univerzity - **Larry Page a Sergej Brin**
- / vycházeli z akademické sféry, citovanost odborných prací
- / googol - číslo, které lze napsat jako jedničku se sto nulami  
> překlep google
  
- / **PAGE RANK = RELEVANCE!!!**
- / kolik a jak důležitých stránek na určitou stránku odkazuje
- / 1-10 (10 má google, 7 idnes, 8 seznam)
- / <http://ranky.cz>
  
- / + další metody pro zpřesnění



## a co peníze?

- / **NEAGRESIVNÍ REKLAMA** – pro každého
- / **ŠIROKÉ PORTFOLIO** – e-mail, kancelářské aplikace, fotky, zprávy, prohlížeč chrome, operační systém android, youtube, google+, hangout, mapy...
- / **část vždy zdarma**



google

<http://www.youtube.com/watch?v=BNHR6IQJGZs>



# vyhledávání na googlu

- / google.com x google.cz (prohledává přednostně domény.cz)
- / př: singapur
- / zpřesňování
- / vyhledávání různých formátů
  - obrázky
    - př. magnetismus
    - dle souboru
  - videa



# vyhledávání na googlu

## / rozšířené vyhledávání

- novinář + žurnalista
- novinář -žurnalista
- novinář OR žurnalista
- „novinář“
- ~žurnalista
  
- link:fsv.cuni.cz – které stránky odkazují na fsv
  
- [http://technet.idnes.cz/triky-google-c9w-/sw\\_internet.aspx?c=A140217\\_152500\\_sw\\_internet\\_pka](http://technet.idnes.cz/triky-google-c9w-/sw_internet.aspx?c=A140217_152500_sw_internet_pka)



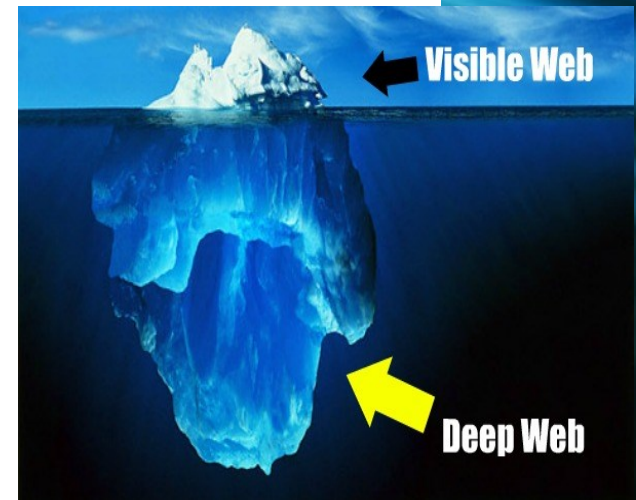
# SEO

- / **SEO – search engine optimization**
- / **x**
- / **Google Panda a Google Pinguine**
  - / [http://seopedia.cz/wiki/Google Panda](http://seopedia.cz/wiki/Google_Panda)



# neviditelný web

- / neprolinkované stránky
- / placené databáze x nepřístupné databáze
- / zaheslované stránky (is)
- / netextové informace (obrázky, videa, software... bez popisků)
- / dynamicky tvořené stránky (java script - letenky)
- / sociální sítě (?)
- / zakázané autorem nebo vyhledávačem
- / změněné stránky - archive.org, webarchiv.cz





# další vyhledávací služby

seznam.cz

bing.com (facebook)

<http://www.ebizmba.com/articles/search-engines> 15 top vyhledávačů

metavyhledávače

<http://www.dogpile.com/> (google, yahoo)

<http://www.metacrawler.com/>

vizuální vyhledávače

<http://taggalaxy.de>

přirozené vyhledávače

<http://www.wolframalpha.com>, <http://ask.com/>





# portály, rozcestníky (katalogy)

## / rozcestník / katalog / brána

- / strukturovaný obsah, kategorie
- / když nevíte, kde začít
  - / <http://www.dmoz.org/>

## / portál

- / komplexní služby
- / vyhledávání + katalog + zprávy + e-mail
- / seznam.cz



# současné trendy

## / sociální sítě

### / 10 SOCIAL NETWORK SEARCH ENGINES FOR JOURNALISTS

<http://davidhiggerson.wordpress.com/2011/07/05/10-social-network-search-engines-for-journalists/>

/ [topsy.com](http://topsy.com)

/ <http://www.socialmention.com/>

/ <http://search.twitter.com/>

/ <http://www.klaboseni.cz>

- / mobilní přístup, práce s jazykem/hlasem
- / přirozené vyhledávání – co nejmenší nároky na uživatele
- / personalizace – sledování historie, geolokační služby, soc. sítě
- / augmented reality
- / reklama + prodej!



# zajímavosti

/ <http://www.google.com/trends>



# google books

## /books.google.cz (2004)

- 129,864,880 (srpen 2010) mil. knih na světě,
- knihy s volným právem užití, ke stažení zdarma v pdf (dle licencí)
- vyhledávání uvnitř knihy



# google scholar

## / scholar.google.cz

/ Google Scholar je specializovaný vyhledávač odborných textů (např. články z časopisů, preprinty, vysokoškolské kvalifikační práce). ScholarSFX je speciální verze link serveru SFX, která slouží k propojení výsledků vyhledávání v Google Scholar s plnými texty v databázích (např. EBSCO, Proquest, JSTOR)

/ citovanost, podobné články, verze

/ plné texty

/ vyhledávají i v některých licencovaných databázích – sfx – nastavit knihovnu!