

# Web a očekávání budoucího

Vilém Sklenák – [sklenak@vse.cz](mailto:sklenak@vse.cz)

Vysoká škola ekonomická, fakulta informatiky a statistiky,  
katedra informačního a znalostního inženýrství

Inforum2013, 21. 5. 2013

# Kontrola od minule

# Kontrola od minule



# Obsah

- 1 Jak se vyvíjel web
- 2 Budoucnost webu
- 3 Big Data
- 4 Sémantizace obsahu webu
- 5 Další faktory ovlivňující budoucí web

# Obsah

- 1 Jak se vyvíjel web
- 2 Budoucnost webu
- 3 Big Data
- 4 Sémantizace obsahu webu
- 5 Další faktory ovlivňující budoucí web

# Z dávné prehistorie

**1962** – **idea „galaktické sítě“** jako možnosti sociální interakce, přístup k datům či programům v sadě navzájem propojených počítačů (prof. J. R. R. Licklider z MIT)

**1969** – **ARPANET** – první globální technologie s využitím přepínání paketů (používá se do současnosti)

**1971** – **první e-mail**, zavedení znaku @ jako oddělovače mezi jmény uživatele a počítače

**1971** – **první e-book** – počátek projektu **Gutenberg**

**1974** – **„Internet“** – protokoly TCP/IP jako základ komunikace mezi počítači z různých sítí, proto byl systém nazván „Internet“

# Z méně dávné prehistorie

**1978** – **první e-mailový spam** – odeslána nevyžádaná zpráva k 500 kalifornských uživatelům sítě ARPANET

**1979** – **první diskusní skupiny** – Usenet News jako systém s možností vyjadřovat se k různým tématům

**1979** – **první hra pro více hráčů** – textově orientované prostředí s rolemi a chatem, předchůdce dnešních online her

**1982** – **první emoce :-)**

**1983/84** – **DNS** – zavedení systému doménových jmen jako alternativního způsobu pojmenování počítačů k číselným IP adresám

**1988** – **první malware** – první virus šířící se počítačovou sítí – „Morrisův červ“

# Zrození webu

**1989/90** – návrh služby **World Wide Web** – CERN, Tim Berners-Lee, globální hypertextový systém založený na: HTML, HTTP a URL

**1991** – první webová stránka – 6. 8. 1991, **web o samotném webu**

**1993** – uvolnění zdrojových kódů, World Wide Web zdarma, vývoj serverů a prohlížečů třetími stranami bez licenčních poplatků

**1993** – **Mosaic**, první grafický prohlížeč

**1994** – založení **W3C** – konsorcium pro vývoj standardů pro WWW



# Expanze

**1994/95** – **komeracionalizace**, využití Internetu i pro komerční subjekty, zabezpečení SSL, vzniká Amazon.com

**1998** – **Google, Napster** – revoluce ve vyhledávání a ve sdílení souborů

**2000** – **splasknutí dotcom bubliny** – vlna krachů firem

**2001** – **Wikipedia** – počátek vlny služeb v duchu Web 2.0 a rovněž odstartování revoluce v sociálních médiích

**2003** – **MySpace** – svého času nejpopulárnější sociální síť

**2004** – **Web 2.0 a sociální média** – první konference Web 2.0, web jako prostředí pro rozšiřování vztahů a interakce, růst aplikací pro sdílení, pro zapojení uživatelů do tvorby obsahu, pro vznik sociálních sítí

# Expanze pokračuje

**2004 – Facebook**

**2004 – iPhone** – mobilní web jako platforma pro dostupnost webu na přenosných zařízeních (chytré telefony, později tablety)

**2009 – změny ICANN**, korporace *The Internet Corporation for Assigned Names and Numbers* je plně nezávislá na vládě USA

**2013–** – co bude dál?

# Obsah

- 1 Jak se vyvíjel web
- 2 Budoucnost webu**
- 3 Big Data
- 4 Sémantizace obsahu webu
- 5 Další faktory ovlivňující budoucí web

# Vývoj webu a verze

# Vývoj webu a verze

**HTTP** – 0.9, 1.0, 1.1

**HTML** – 2.0, 3.2, 4.0, 4.01, XHTML (1.0, 1.1, 2.0), HTML5

**CSS** – 1, 2, 2.1, 3, 4

Pod gescí konsorcia W3C.

# Uvažované verze webu

**Web 1.0** se obvykle označuje „read-only“ web, zhruba první desetiletí existence webu, hlavní „hráči“ – Netscape, Yahoo, AOL, eBay.

**Web 2.0** znamená socializaci webu, lidé mohou spoluvytvářet a sdílet obsah – MySpace, Flickr, Facebook, Twitter, YouTube, . . .

**Web 3.0** je obtížné definovat, protože jednak je to určitá vize, jednak každý má svůj názor, co ano a co ne. Obecně se předpokládá, že půjde o evoluci Webu 2.0 s větší inteligencí

Přirozený vývoj.

# Faktory ovlivňující budoucnost webu

- 1 „mobilizace“ – přístupy z mobilních zařízení získávají stále větší podíl ve statistikách, výhoda přístupu v reálném čase (pasivně i aktivně), a proto je potřeba uvažovat způsoby interakce, podobu uživatelského rozhraní, dále pak platí silnější vazby na kontakty a zájmy – proto větší tlak na sociálně orientované webové aplikace apod. „Re-thinking“ všeho, co vás napadne ve spojení s webem a jeho používáním.

# Faktory ovlivňující budoucnost webu

- 1 „mobilizace“ – přístupy z mobilních zařízení získávají stále větší podíl ve statistikách, výhoda přístupu v reálném čase (pasivně i aktivně), a proto je potřeba uvažovat způsoby interakce, podobu uživatelského rozhraní, dále pak platí silnější vazby na kontakty a zájmy – proto větší tlak na sociálně orientované webové aplikace apod. „Re-thinking“ všeho, co vás napadne ve spojení s webem a jeho používáním.
- 2 Big Data



# Faktory ovlivňující budoucnost webu

- 1 „mobilizace“ – přístupy z mobilních zařízení získávají stále větší podíl ve statistikách, výhoda přístupu v reálném čase (pasivně i aktivně), a proto je potřeba uvažovat způsoby interakce, podobu uživatelského rozhraní, dále pak platí silnější vazby na kontakty a zájmy – proto větší tlak na sociálně orientované webové aplikace apod. „Re-thinking“ všeho, co vás napadne ve spojení s webem a jeho používáním.
- 2 Big Data
- 3 sémantický web

# Obsah

- 1 Jak se vyvíjel web
- 2 Budoucnost webu
- 3 Big Data**
- 4 Sémantizace obsahu webu
- 5 Další faktory ovlivňující budoucí web

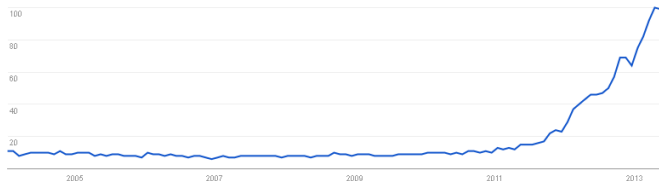
# Co je to

- relativně nový fenomén

## Interest over time ?

The number 100 represents the peak search interest

News headlines  Forecast ?



- Big Data = datové objemy rostou do té míry, že moderní databázové systémy se potenciálně blíží k mezi svých možností z hlediska jejich zpracování

# Kolik je moc

kilobyte	kB	$2^{10}$	$10^3$
megabyte	MB	$2^{20}$	$10^6$
gigabyte	GB	$2^{30}$	$10^9$
terabyte	TB	$2^{40}$	$10^{12}$
petabyte	PB	$2^{50}$	$10^{15}$
exabyte	EB	$2^{60}$	$10^{18}$
zettabyte	ZB	$2^{70}$	$10^{21}$
yottabyte	YB	$2^{80}$	$10^{24}$

# Co si představit, co je opravdu velké

- V roce 2012 žádné datové úložiště neobsahovalo 1 ZB.
- Jedna z firem prodávající externí paměti např. jen v roce 2011 dodala zákazníkům více než 330 EB.
- Podle odhadů International Data Corporation (IDC) globální objem dat dosáhl na konci roku 2012 2,7 ZB, což je o 48 % více než v 2011.
- Podle odhadů firmy Cisco dosáhne objem datových přenosů v Internetu v roce 2016 více než 1 ZB, což je tolik jako za všechny roky až do roku 2012 včetně. Počet zařízení připojených k Internetu dosáhne téměř 19 mld. . .

# Jaké jsou důsledky

Problémy různého druhu:

- uchovávání,
- zálohování,
- ochrana,
- organizace
- ...
- dostupnost kritických dat vždy a všude.

vše s ohledem na náklady.

# Jaké jsou důsledky

Problémy různého druhu:

- uchovávání,
- zálohování,
- ochrana,
- organizace
- ...
- dostupnost kritických dat vždy a všude.

vše s ohledem na náklady.

A co navíc:

- vyhledávání,
- analýzy – hledání nových souvislostí alias *data mining*,
- vizualizace.

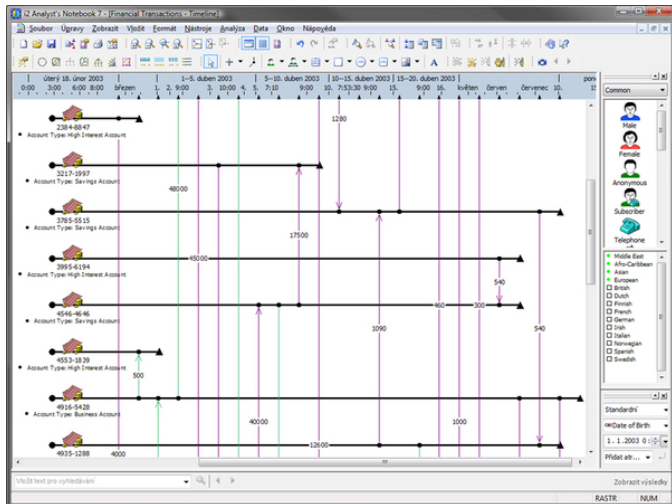
# Co je lepší?



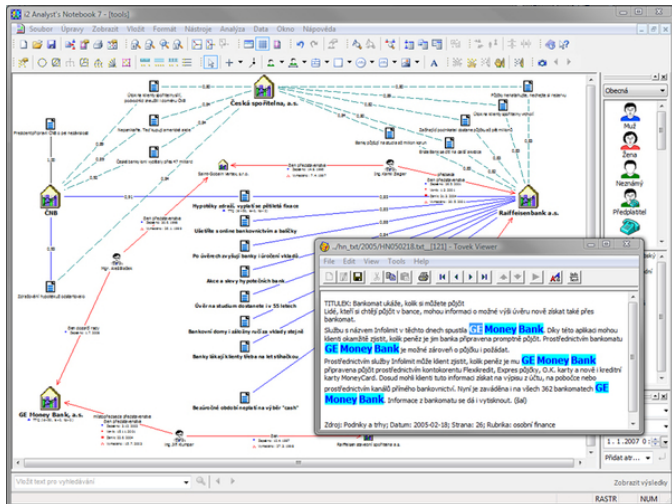




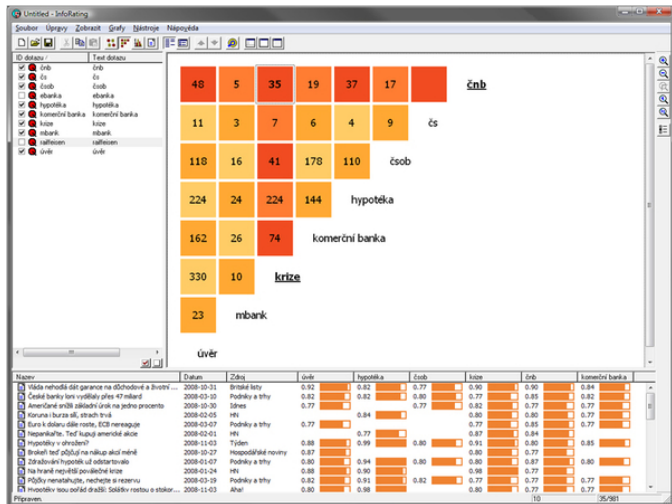
# Co je lepší?



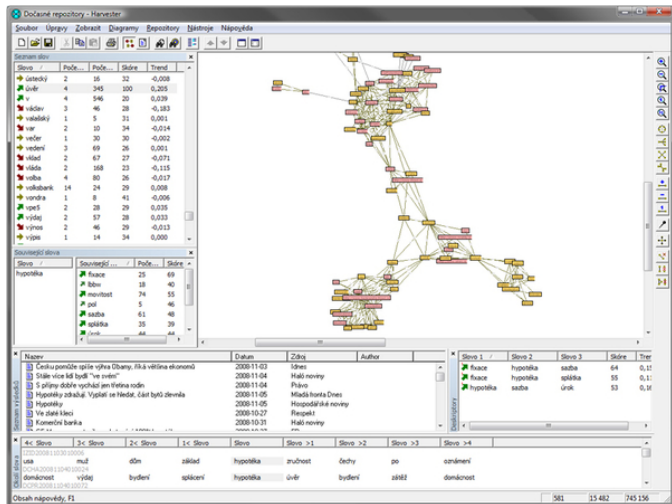
# Co je lepší?



# Co je lepší?



# Co je lepší?



# Big Data vs. web

Data na webu v posledních letech přibývají nepoměrně rychleji než dříve. Jeden z důsledků služeb na principu Webu 2.0:

- upload videí nebo fotografií pro účely sdílení,
- blogy a uživatelské komentáře,
- ...

1995 – více než  $10^6$  indexovaných dokumentů

2000 – více než  $10^9$  indexovaných dokumentů

2013 – více než  $50 \cdot 10^9$  indexovaných dokumentů

nebo také: na konci roku 2009 odhad, že obsah webu je asi 500 EB.

vybrané statistiky: [YouTube](#), [Facebook](#), [Twitter](#)

# Obsah

- 1 Jak se vyvíjel web
- 2 Budoucnost webu
- 3 Big Data
- 4 Sémantizace obsahu webu**
- 5 Další faktory ovlivňující budoucí web

# Sémantický web

- Vznik myšlenky a rozvoj základních principů sémantického webu není věcí posledních 10 let.
- Ale – 2001 = Berners-Lee, Tim, Hendler, James, Lassilla, Ora. The Semantic Web. *Scientific American*, 2001. vol. 284, no. May. str. 35–43. URL: <http://www.sciam.com/2001/0501issue/0501berners-lee.html>.
- V té době však práce na sémantickém webu trvaly již několik let.
- Od počátku platí, že *„sémantický web je rozšířením současného webu, jež datům přiřazuje přesný význam, díky kterému bude možná kooperace jak lidí, tak softwaru“*.



# Linked Data

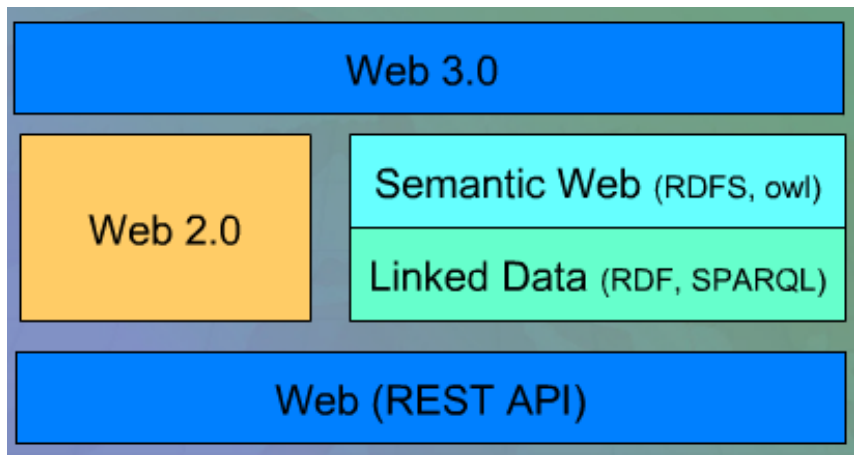
- jeden z aktuálních trendů směřujících k budování sémantického webu
- Tim Berners-Lee: „*The value of your own information is very much a function of what it links to, as well as the inherent value of the information within the web page*“
- propojování strukturovaných dat na webu pomocí odkazů
- styl publikování a propojování, soubor doporučení i data samotná
- *web dat*

# Linked Data a sémantický web

Linked data i sémantický web se shodují v mnoha předpokladech: snaží se o rozšíření současného webu, propagují použití rámce RDF, identifikátorů URI a efektivnější zpracování dat webovými aplikacemi. Také propojování dat je společným bodem obou iniciativ.

*„Vize sémantického webu je založena na budování globálního webu strojově zpracovatelných dat. Zatímco sémantický web, nebo web dat je cílem či výsledkem, linked data představuje prostředek nebo způsob jeho dosažení.“*

# Linked Data a sémantický web vs. Web 3.0



# Obsah

- 1 Jak se vyvíjel web
- 2 Budoucnost webu
- 3 Big Data
- 4 Sémantizace obsahu webu
- 5 **Další faktory ovlivňující budoucí web**

# Web jako komunikační platforma

- **IP mobilita**

# Web jako komunikační platforma

- IP mobilita
- Cloud (virtualizace zdrojů)

# Web jako komunikační platforma

- **IP mobilita**
- **Cloud (virtualizace zdrojů)**
- **Internet věcí**

Kdyby web neexistoval,



Kdyby web neexistoval,  
museli bychom ho vymyslet.

Kdyby web neexistoval,  
museli bychom ho vymyslet.

Děkuji za pozornost.