

PVo70 – Digitální knihovny

**Technorati – searching the  
blog space**

**Bc. Petr Škoda  
2008**

# Obsah

---

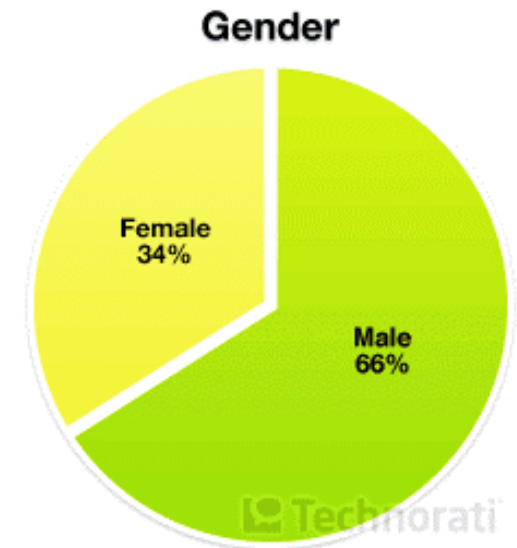
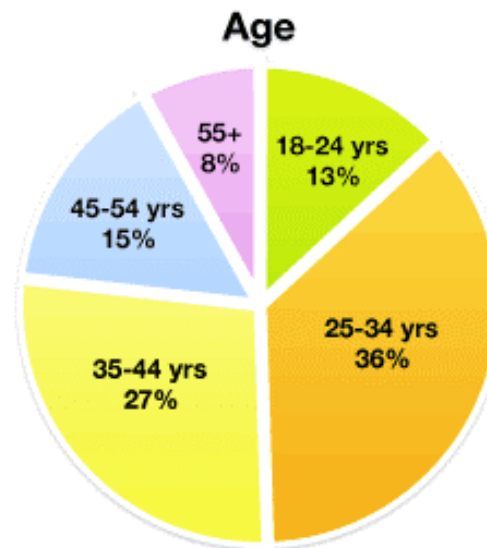
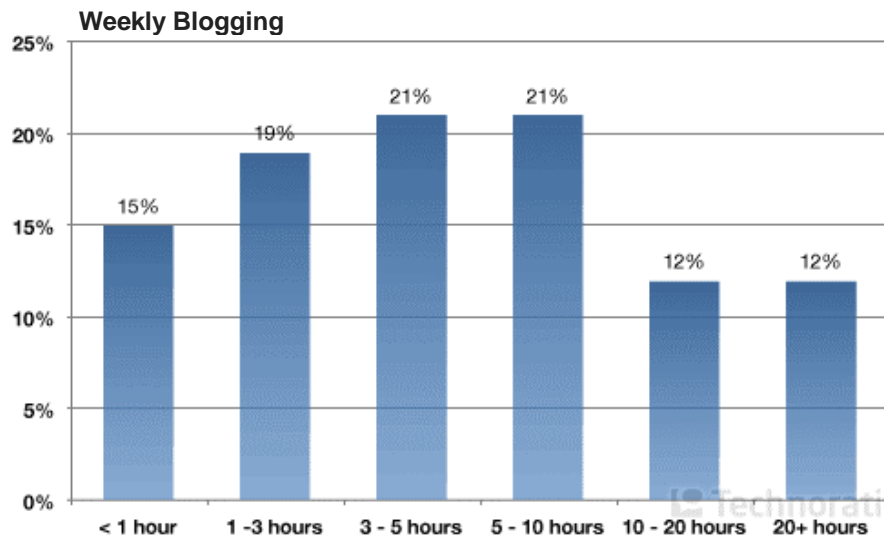
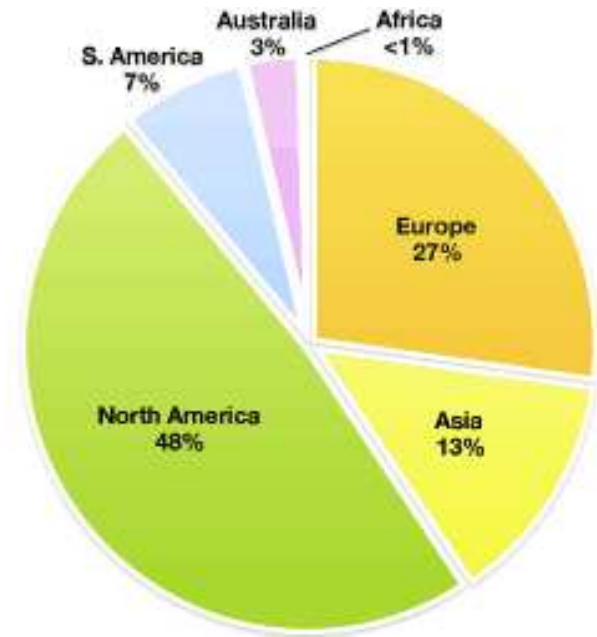
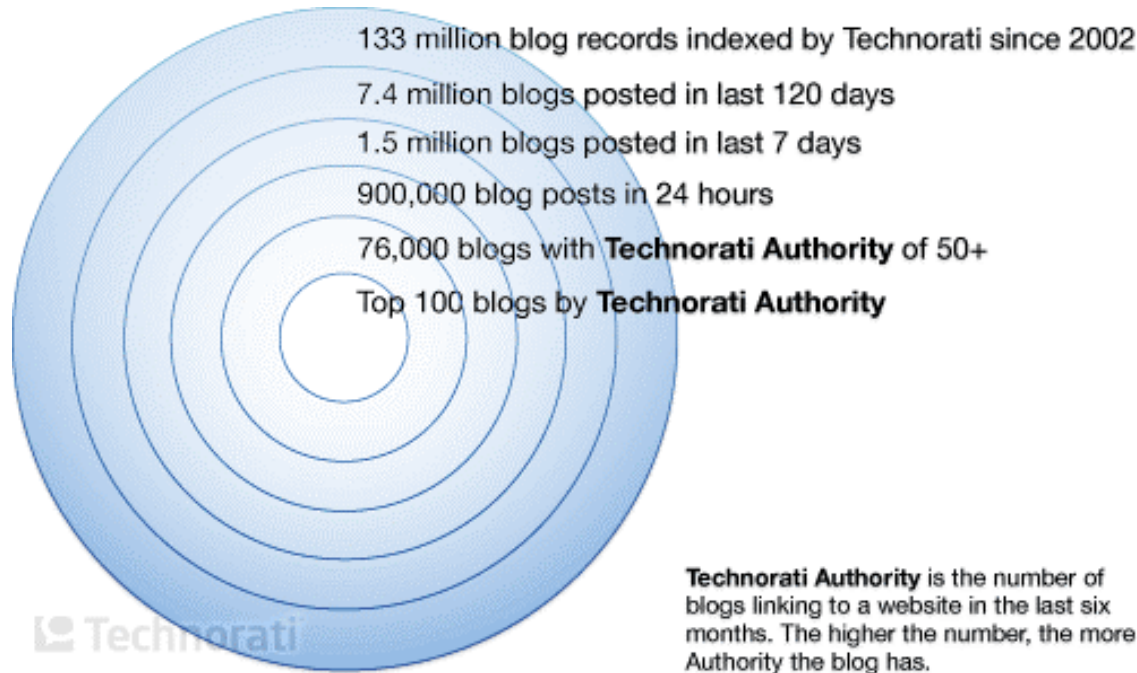
- **Technorati**
- **Blogosféra**
- **Proč Yahoo! nestačí?**
- **Jak to dělat lépe**
- **Technologie**
- **Problémy Technorati**

# Projekt Technorati

---

- vyhledávač zaměřený na svět blogů
- první svého druhu
- „Jestliže Google je referenční knihovnou internetu, pak je Technorati jeho kavárnou.“ (*Time Magazine*)
- nová úroveň aktuálnosti
- **Vývoj**
  - prvotní nápad a implementace **David Sifry** 2002
  - iritován neschopností konvenčních vyhledávačů
  - cílem bylo vytvořit nástroj pro vyhledávání v blozích

# Blogosféra (*blogosphere*)



# Proč Yahoo! nestačí?

---

- konvenční vyhledávače hodnotí obsah podle počtu na něho směřujících odkazů
- osvědčené kritérium
- objevil Google
- **toto kritérium je vhodné i pro blogy**
- **Indexování – prohledávání – crawling**
  - Vyhledávače pro web se snaží zaindexovat co nejvíce obsahu
  - cowlery prolézají celý Internet
  - nové věci jsou zaindexovány v horizontu dnů (záleží na povaze webu a na tom, jak často se mění)
  - není možné indexovat blogosféru „denně“

# Jak vyhledávat lépe

---

- **Dejte o sobě vědět – ping**
  - obnovení zavrženého nápadu
  - pro blogy výhodné – dynamika
  - nové příspěvky indexujeme „ihned“ po jejich vložení
  - nástroj pro upozornění vyhledávačů na nový obsah součástí blogovacích aplikací
- **Kdo? Kdy? Kde?**
  - ping vymyslel Dave Winer pro *weblogs.com*
  - nevyužíval ping pro vyhledávání, jen informoval uživatele o novinkách na blozích
  - Sifry nejdříve prohledával blogy podle záznamů z *weblogs.com*, později ping přímo na Technorati

# Technologie

---

- **LAMP**

- Linux
- MySQL
- PHP + Perl

- **Potírání spamu – *sblogs***

- jakmile bylo možné vyhledávat v blozích, začaly se objevovat podvodné blogy s různými typy reklam
- snaha dostat se na přední místa vyhledávání odkazováním mezi sebou
- daří se poměrně dobře potírat, ale je třeba sofistikovaných technologií

# Problémy Technorati (Inc.)

---

- problematický postoj k MySpace
- v roce 2006 na MySpace 15 milionů blogů
- v roce 2007 110 milionů měsíčně aktivních uživatelů (obecně – ne nevybítně blogy)
- Technorati neindexuje MySpace blogy!!
- MySpace prý neobsahuje blogy, ale jen „příspěvky“, které nemají povahu blogu
- MySpace prý užívají převážně teenageři
- podle mého jsou to nesmyslné argumentace
- Technorati se zdá být příliš arogantní a komerční



# Závěrem

---

- přes problémy je Technorati uznávaný vyhledávač ve světě blogů
- konkurence neprohledává tolik zdrojů
- je třeba znát stinné stránky při používání Technorati
- zprávy o stavu světa blogu (*State of the Blogosphere*) brát napůl jako reklamu

*Děkuji za pozornost!*