

WEB MINING

Web mining je uplatnění technik z data miningu k vyhledávání informací na webu.

Typy web miningu

- **Web structure mining** – jak uživatel používá web (jaké linky)
- **Web content mining** – data posbírání přes vyhledávače
- **Usage mining** – data z on-line formulářů

Web structure mining

- Extrakce informací z hypertextových odkazů
- Extrakce informací ze struktur dokumentů
- Příklad – PageRank (Google) – analyzuje se počet linků na danou strukturu

Web content mining

- Analýza informací z obsahu
- Semi-strukturovaná data, web agents
- Vektorová reprezentace dokumentů

Dolování textů

- Předzpracování
- Dolování
- Upřesnění