



nes  fit

# Sociální sítě

## Softwarové nástroje – návrh

Radek Burget  
2017-05-24



# Co máme

- Profily na sociálních sítích
  - S použitelným API, stačí účet (Twitter, Flickr)
  - API pouze pro vlastní účet (Facebook)
  - Bez API (Linkedin)
- Profily prohlížečů na lokálním počítači
  - Které sítě dotyčný využívá (historie)
  - Možný přístup k účtu (cookies)
- *(Data z monitorování síťového provozu)*

# Cíle

- Vytvoření kopie dostupných příspěvků napříč soc. sítěmi (texty, obrázky, místa, časy)
- Porovnání různých profilů napříč soc. sítěmi
- Korelace s historií prohlížení, příp. s obsahem disku
- Příklady použití
  - Patří profil na soc. síti uživateli počítače?
  - Patří dva profily stejnému uživateli?
  - Známe-li jméno, který z vyhledaných profilů nás zajímá?

# Prostředky

- Jednotné úložiště dat (RDF)
  - Chronologie příspěvků, text, metadata (časy, místa)
- Získání dat ze sociálních sítí
  - API, kde je to možné
  - Extrakce dat – simulace prohlížeče
- Korelace profilů
  - Extrakce metadat z textu (témata, klíčová slova, ...)
  - Porovnání obrázků, geo tagů
  - Časy přístupů z lokálního zařízení

# Přístup k účtům

- Znalost přístupových údajů
- Využití uložených cookies
- Využití jiného účtu
  - Přístup ke „spřátelenému“ účtu
  - Často alespoň nějaký účet je lepší, než žádný

# Stav řešení

- Prototypy dílčích částí
  - Stahování dat přes API (Twitter)
  - Využití cookies pro přihlašování (FB, LinkedIn)
- Experimenty se simulací prohlížeče
  - Základní funkční infrastruktura (webdriver)
  - Práce na extrakci dat
- Výhled
  - Návrh jednotného rozšiřitelného úložiště (RDF?)



Questions?