

Nástroje pro léčbu informačního zahlcení

Jiří Šmerda Vladimír Dosoudil Marek Winkler

Masarykova univerzita, Ústav výpočetní techniky

Masarykova univerzita, Fakulta informatiky,
Laboratoř znalostních a informačních robotů

22. ledna 2008



Osnova

- 1 Úvod
- 2 Motivace
 - Motivační příklad
 - Proces získávání informací
 - Řešení dnešními nástroji
 - Identifikované problémy
- 3 Představení vyvíjených technologií
 - Východiska v servisních systémech
 - Schéma užitečnosti
 - Aplikační oblasti
- 4 Konkrétní užití v oblasti digitálních knihoven
 - Cíl v oblasti digitálních knihoven
 - Současný stav vývoje
- 5 Shrnutí

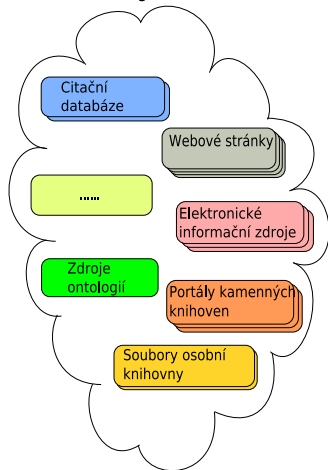
Motivační příklady z oblasti digitálních knihoven

Uznávání autoři a renomované časopisy

- Jsem lékař a zajímám se o infarkt mozku.
- Největší problém vidím ve faktu, že existuje v různých informačních zdrojích příliš mnoho článků o tomto tématu a nemám čas je všechny procházet a pročítat.
- Chci vědět, co publikovali uznávané vědecké kapacity o infarktu mozku v renomovaných časopisech. Zajímají mě relevantní články jak v elektronickém, tak v papírovém vydání.

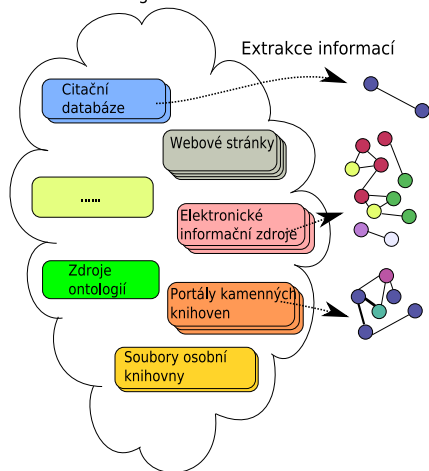
Ilustrace procesu získávání informací

Prostor digitálních knihoven

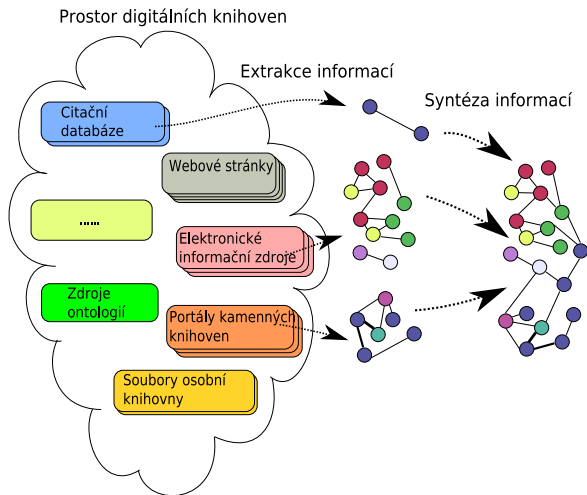


Ilustrace procesu získávání informací

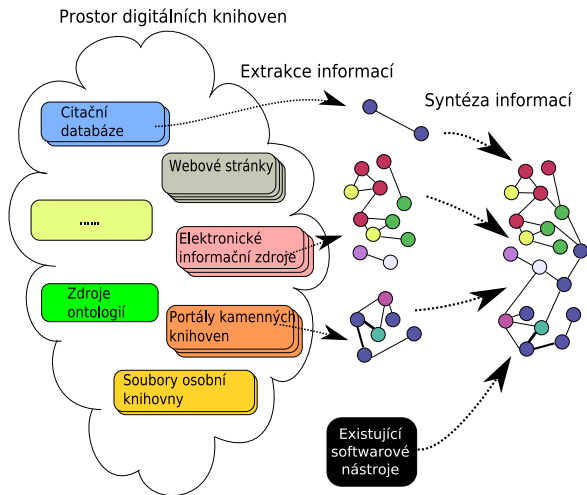
Prostor digitálních knihoven



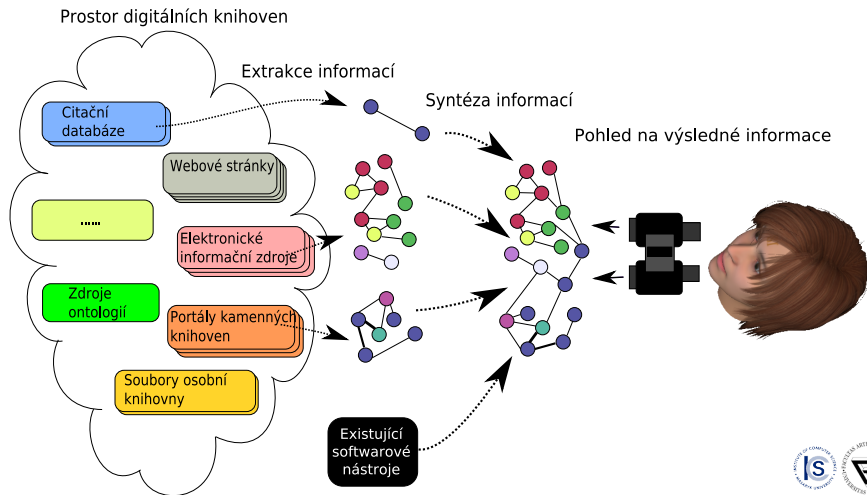
Ilustrace procesu získávání informací



Ilustrace procesu získávání informací



Ilustrace procesu získávání informací



Další motivační příklad

Dostupnost časopisu nebo plného textu článku

- Je pro mě dostupný fulltext daného časopisu/článku?

Dnešní nástroje

- pomáhají řešit dílčí úlohy
- informace jsou dostupné v izolovaných celcích
- celkový pohled chybí

⇒ v současné době nejsme schopni uspokojivě zmíněné příklady řešit.

Jednoduchost přístupu a informační schopnost

Jednoduchost přístupu

- informace sice dostupné jsou, ale nejsou snadno dostupné
- v důsledku vede k tomu, že nejsou využívány

Informační schopnost

- informační schopnost sjednocení všech informačních zdrojů je větší než sjednocení informačních schopností jednotlivých informačních zdrojů

Informační nedostatek a informační zahlcení

Informační nedostatek

- absence povědomí o dostupnosti užitečných informací
- odkud máme vlastně chtěné informace získat?

Informační zahlcení

- při práci s vybranými zdroji postupně docházíme k tomu, že obsahují příliš veliké množství informací
- „oddělit zrna od plev“ je náročné
- **informační zahlcení lze ovšem mírnit za pomoci dalších informací, které jsou nám dostupné**

Obecné představení vyvíjených technologií

Technologie znalostních a informačních robotů

servisní systém

- komplexní adaptivní systém, kde lidé a technologie navzájem spolupracují k tomu, aby vytvořili přidanou hodnotu
- teoretická východiska ve směru SSME (Service, Science Management and Engineering), který je iniciativou IBM



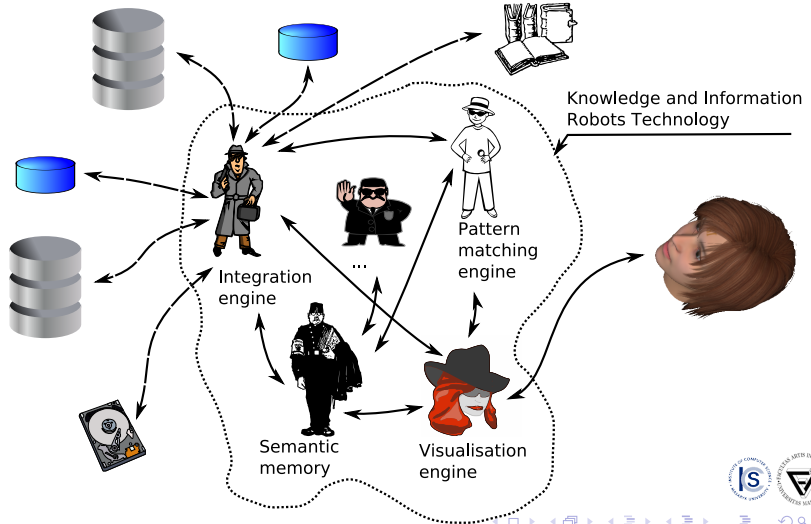
Technologie znalostních a informačních robotů

Základní rysy

- založeno na agentech, kteří jsou především
 - autonomní
 - kooperativní
 - skladební
 - distribuovatelní
 - interaktivní
 - adaptabilní
- výzkum v oblasti znalostních a informačních robotů je prováděn na Fakultě informatiky na Masarykově univerzitě v Laboratoři znalostních a informačních robotů
- konkrétní implementace se děje v rámci spin-offu Masarykovy univerzity, firmy Mycroft Mind, a.s.



Ilustrace architektury



Vyvíjené technologické prvky

- Vizualizační engine
 - kombinuje různé **vizualizační metody** v závislosti na povaze dat a preferencích uživatele
- Sémantická paměť
 - práce s **kontextově závislou, neurčitou a pozornostně ohodnocenou informací**
- Pattern engine
 - práce se vzory **struktur a chování**
- Integrovaná engine
 - **přiblížení dat** z místa jejich vzniku či původního uložení ke zpracování
- Organizační platforma
 - propojení předešlých technologií umožňující **škálovatelnost výpočetního výkonu** systému pomocí distribuovatelnosti

Schéma užitečnosti: Dalekohled do informačního prostoru

1 Ukaž

- zprostředkuje vhled a interaktivní procházení daty z různých pohledů
- interakce pomocí dvou základních pokynů: **obširněji**, **stručněji**

2 Poznej

- rozeznává a upozorňuje na definované vzory struktur v medicínských datech

3 Porad'

- navrhuje opatření pro případ, kdy se daný vzor objeví

4 Udělej

- podpora administrativní práce (např. generování reportů a jejich odesílání e-mailem)



Vývoj v aplikačních oblastech

- bezpečnost v oblasti počítačových sítí
- digitální knihovny
- inteligentní sensory
- project management
- zdravotnictví
- ...

Konkrétní použití vyvíjených technologií v oblasti digitálních knihoven



Náš cíl v oblasti digitálních knihoven

Vyvinout „dalekohled“ do oblasti digitálních knihoven, který podpoří proces získávání informací z prostoru digitálních knihoven.

- extrakce informací z heterogenních datových zdrojů
- adaptace informací použitím konceptuálního datového modelu oblasti digitálních knihoven
- informační syntéza (ne pouze integrace)
- práce s neurčitou a kontextově závislou informací
- použití metod pattern matchingu
- pokročilé metody vizualizace

Kombinací těchto metod lze uživateli umožnit soustředit se na důležité věci a odfiltrovat věci nedůležité.



Současný stav řešení v oblasti digitálních knihoven

- první konkrétní výsledky se očekávají v průběhu roku 2008
- zaměřujeme se na metadata
 - hlavně bibliografické záznamy
 - nesoustředíme se nyní na prohledávání plných textů
- užití technologie znalostních a informačních robotů v oblasti digitálních knihoven se děje v rámci Ústavu výpočetní techniky Masarykovy univerzity

Další témata zájmu do budoucna

- osobní knihovny
- adaptace na uživatele
- monitorování datových zdrojů v oblasti digitálních knihoven
- práce s ontologiemi
- práce s plnými texty

Shrnutí

- v první části jsme představili motivační příklady z oblasti digitálních knihoven, ilustrovali proces získávání informací z prostoru digitálních knihoven a poukázali jsme na problémy, na které jsme narazili při práci s informacemi
- v druhé části jsme představili hlavní principy technologie znalostních a informačních robotů
- ve třetí části jsme ukázali konkrétní použití představené technologie v oblasti digitálních knihoven

Diskuze

Děkuji za pozornost!

Otázky k diskuzi:

- Narazili jste ve vaší práci na některé problémy, které by mohly být řešeny užitím našeho přístupu?
- Napadá vás další případ použití?
- Zdá se vám náš přístup smysluplný?
- Kde vidíte potenciálně největší problémy?

Kontakt: Jiří Šmerda - smerda@ics.muni.cz

