

Ľuboš Demovič
Vedúci: Michal Holub
11.03.2014

Prepájanie dát na Webe za účelom obohacovania kontextu o nové informácie

Motivácia a cieľ

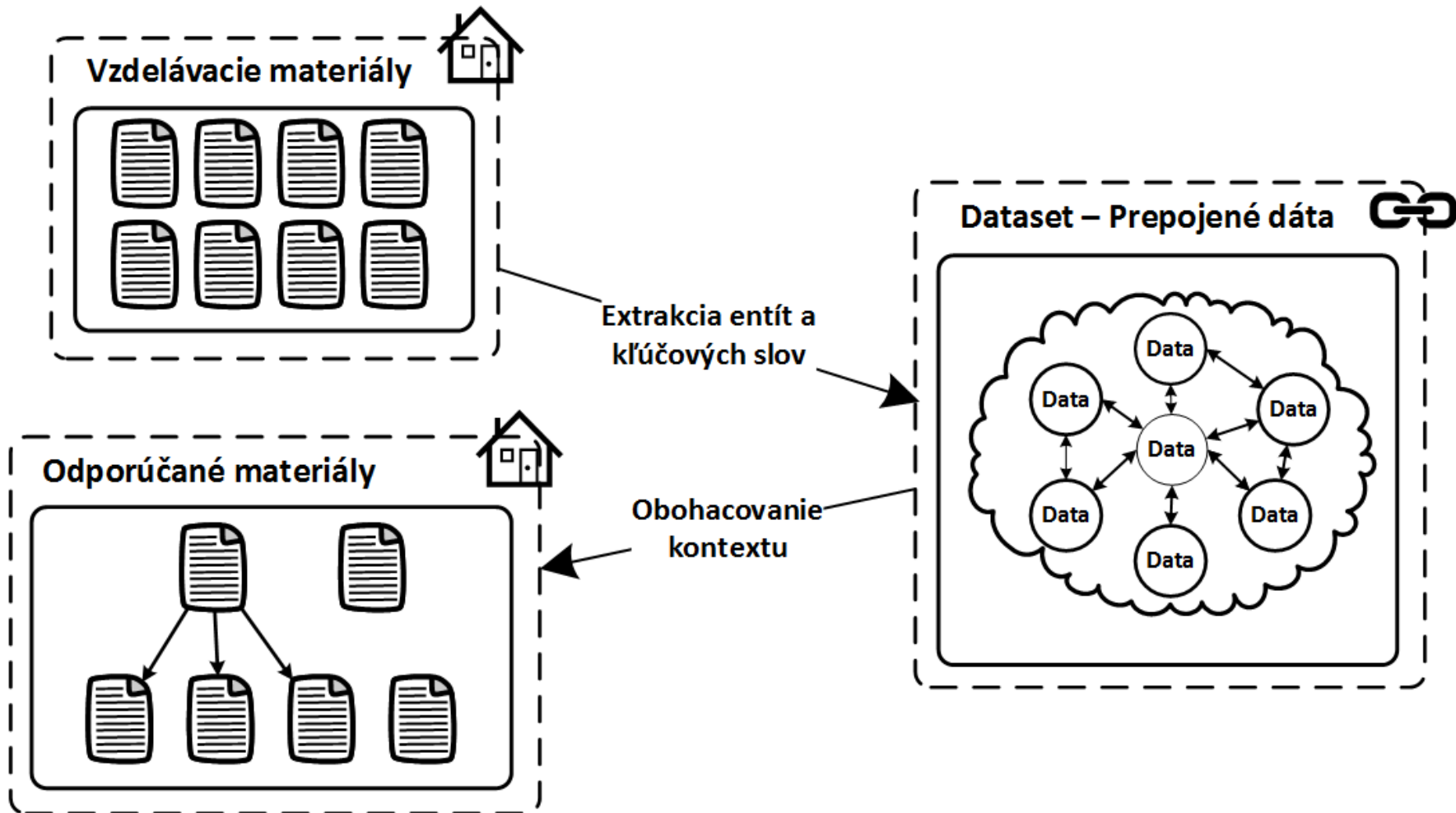
- Využiť potenciál Linked Data v učebných prácach
 - Automatizovane prepojiť entity v angličtine a v slovenčine s anglickou DBpédia
 - Obohatenie kontextu o nové informácie
 - Zlepšenie dostupnosti študijných materiálov

(DBpédia nie je v slovenčine, už ani nie v češtine)



- Dáta na Webe sú neštruktúrované
- Doména: Študentský webový portál

Návrh nášho riešenia



Ako prepojiť slovenčinu s DBpédia



Ľudovít Štúr (*1815 - †1856)

Hypotéza:

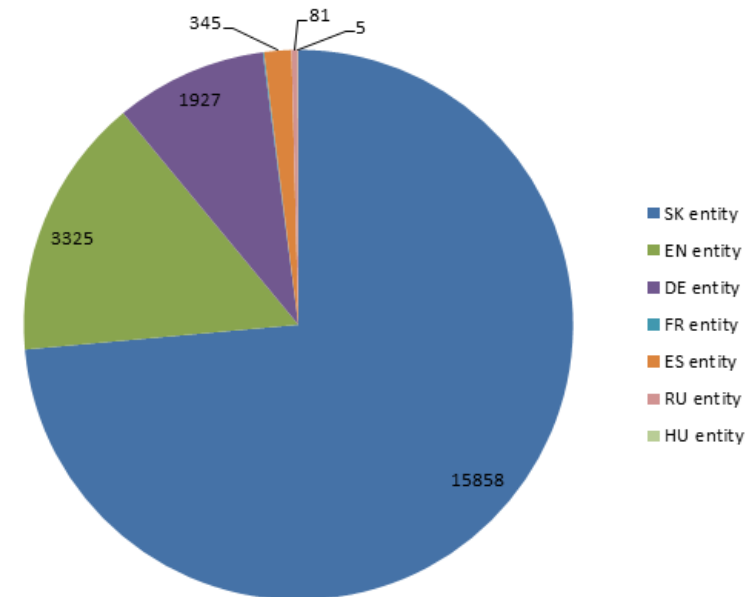
Dokážeme s vysokou presnosťou automatizovane prepojiť slovenskú entitu s anglickou DBpédiou.

Overenie - Experiment 1

- Prepojenia vyhľadávacích dopytov
- Počet všetkých dopytov: 66 180
- Počet všetkých jedinečných dopytov: 21 556

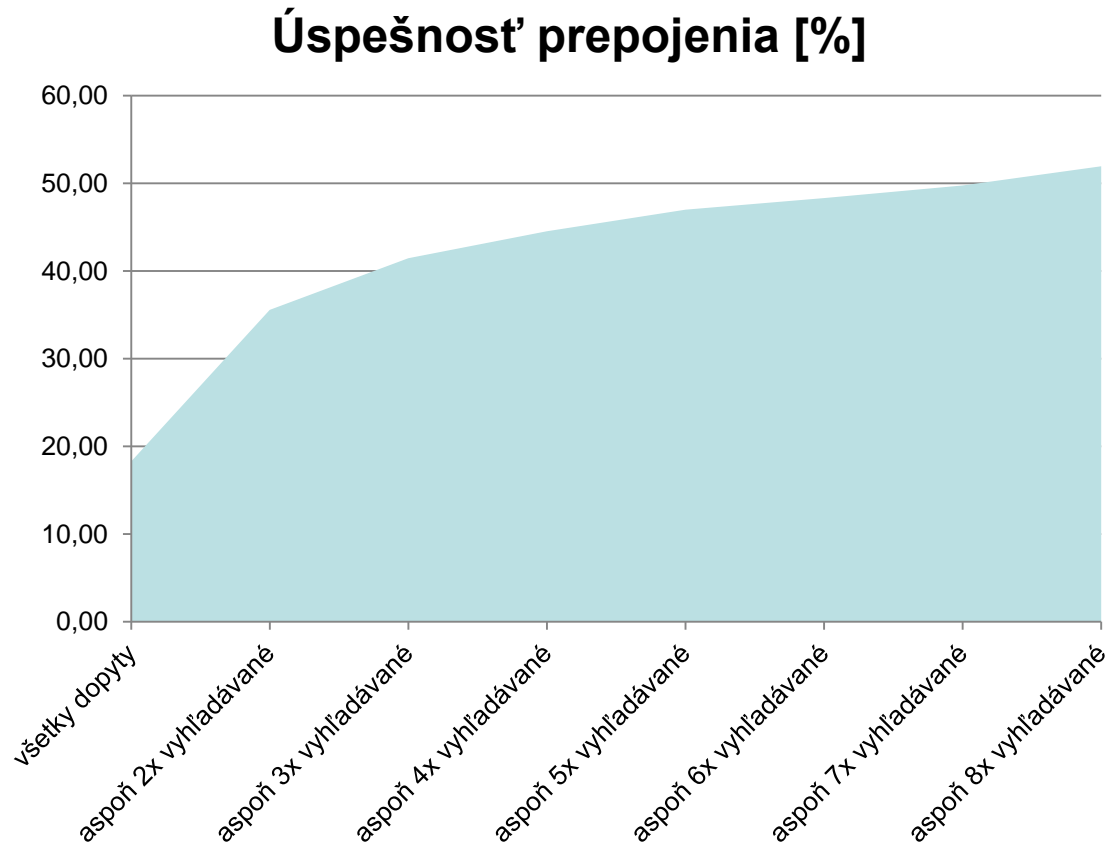
- Úspešnosť prepojenia:
 - Rozlišujeme až 9 rôznych jazykov
 - Záleži od populárnosti vyhľadávania (Čím viac hľadané, tým lepšie prepojenia)
 - Úspešnosť od 18% až po 52%

Rozloženie dopytov podľa jazyka



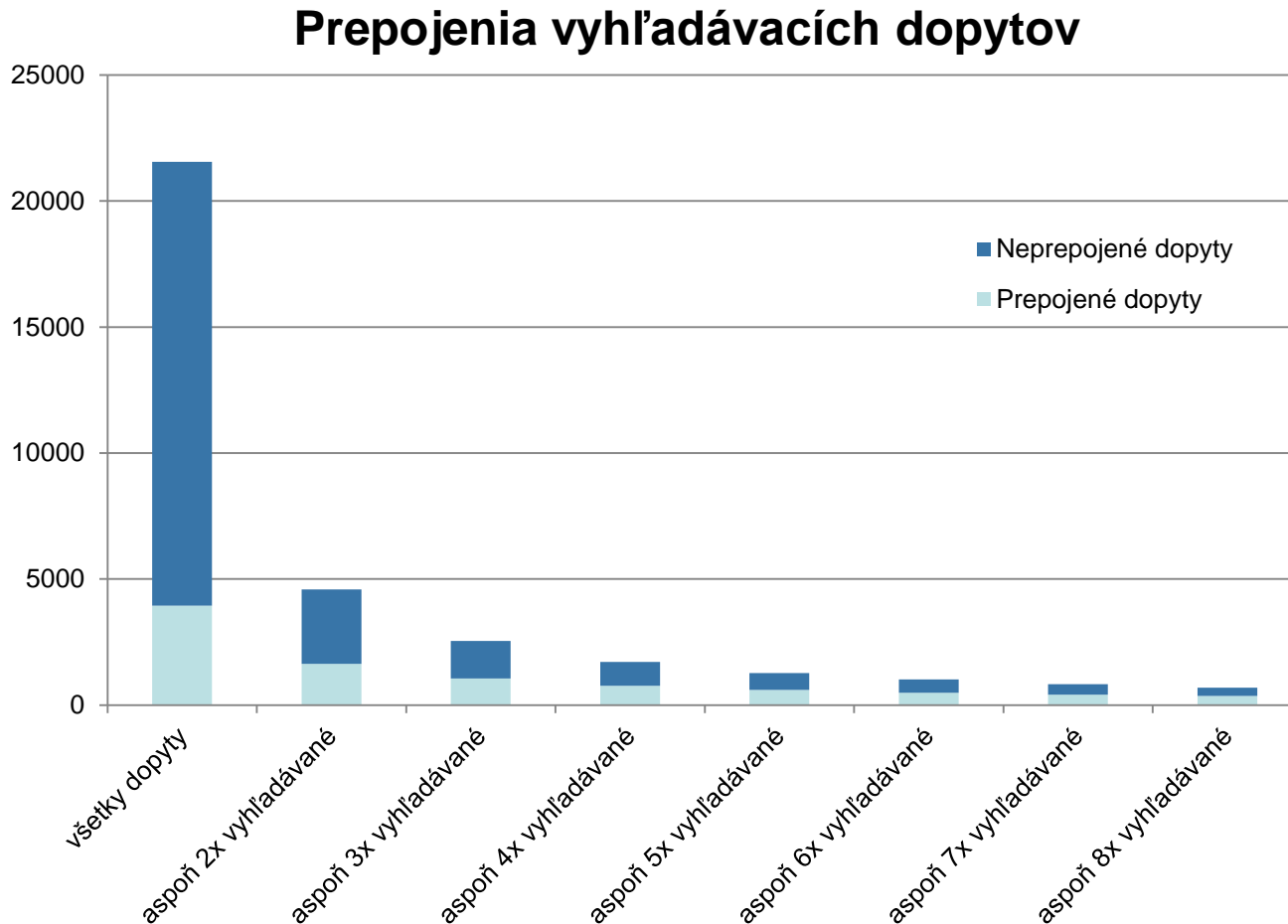
Overenie - Experiment 1

- Prepojenia vyhľadávacích dopytov podľa vyhľadávateľnosti



Overenie - Experiment 1

- Prepojenia vyhľadacích dopytov podľa vyhľadávanosti



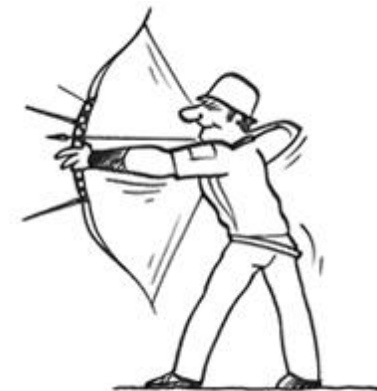
Overenie – experiment 2

- Určenie presnosti automatizovaného prepojenia **populárnych vyhľadávacích dopytov** s DBpédia
- Používateľovi zobrazíme:
 - Náhodne vygenerovanú študijnú prácu
 - Prepojené vyhľadávacie dopyty
 - Abstrakty daných entít z Wikipédia
(pre určenie správnosti prepojenia)
- Overíme na vybraných používateľoch
 - Explicitná spätná väzba: správne/nesprávne



Overenie – experiment 3

- Určenie presnosti automatizovaného prepojenia slovenskej entity s DBpédia
- Používateľovi zobrazíme:
 - Náhodne vygenerovanú študijnú prácu
 - Prepojené extrahované entity a kľúčové slová
 - Abstrakty daných entít z Wikipédia
(pre určenie správnosti prepojenia)
- Overíme na vybraných používateľoch
 - Explicitná spätná väzba: správne/nesprávne



Obohacovanie kontextu

Hypotéza:

Obohacovanie kontextu o nové informácie pri študijnom materiáli bude pre používateľov relevantné a vygenerované prepojenia im budú užitočné vo vzdelávacom procese.



Overenie – experiment 4

- Určenie relevantnosti obohateného kontextu o nové informácie využívajúce Linked Data
- Používateľovi zobrazíme:
 - Náhodne vygenerovanú študijnú prácu
 - Box s novými informáciami/prepojeniami na študijné materiály
- Overíme na reálnych používateľoch
 - Explicitná spätná väzba: relevantné/nerelevantné



Prepájanie dát na Webe za účelom obohacovania kontextu o nové informácie

- Motivácia a cieľ
- Návrh riešenia

- Overenie – experimenty
 1. Prepojenia vyhľadávacích dopytov
 2. Presnosť prepojenia vyhľadávacích dopytov
 3. Presnosť prepojenia s DBpédia
 4. Relevantnosť obohateného kontextu

