

Získavanie metadát k osobným multimédiám

Efektívne spracovanie dát získaných
pomocou hry s účelom



Motivácia a ciele

- ▶ Nedostatok kľúčových slov a opisných dát v personálnych fotoalbumoch
- ▶ Získavanie opisných dát
 - Návrh metód na efektívne spracovanie anotácií k fotografiám
 - Využitie hry s účelom na získavanie potrebných anotácií



Metódy získavania metadát

- ▶ Manuálne prístupy
- ▶ Automatické prístupy
- ▶ Čerpanie z davu (angl. crowdsourcing)
- ▶ Hry s účelom



Výber vhodnej metódy

- ▶ Potrebné typy metadát
 - Mená osôb
 - Miesto
 - Udalosť
 - Sviatok
- ▶ Čiastočné zlyhanie predošlých metód
- ▶ Kombinácia automatického prístupu s hrou s účelom



Metóda extrakcie kľúčových slov

- ▶ Vlastné predspracovanie anotácií
 - Tokenizácia
 - Lematizácia
 - Preklad do angličtiny
 - Získanie základného tvaru slova
 - Vylúčenie stop slov
- ▶ Využitie externých nástrojov (Alchemy API, Tagthe.net, Zemanta API)



Metóda extrakcie menných entít

- ▶ Využitie zoznamov menných entít
- ▶ Hľadanie výskytu týchto entít v anotáciách
- ▶ Použitie rôznych porovnávacích metód
 - Hammingova vzdialenosť
 - Levenshteinova vzdialenosť
 - Jaro–Winkler vzdialenosť
- ▶ Použitie externých nástrojov



Vzorové priebežné výsledky

Entita (meno osoby)	Miera zhody	N-gram z anotácie
Michal Tvarozek	1	tvarozek
Jozef Tvarozek	1	tvarozek
Jozef Tvarozek	0,8571429	jozko tvarozek
Jozo Vasko	0,8	jozko

Entita (meno osoby)	Miera zhody	N-gram z anotácie
Fruhvald Gabor	0,8	gabo
Gabor Nagy	0,8	gabo
Belucz Gabor	0,8	gabo
Gabor Koller	0,8	gabo

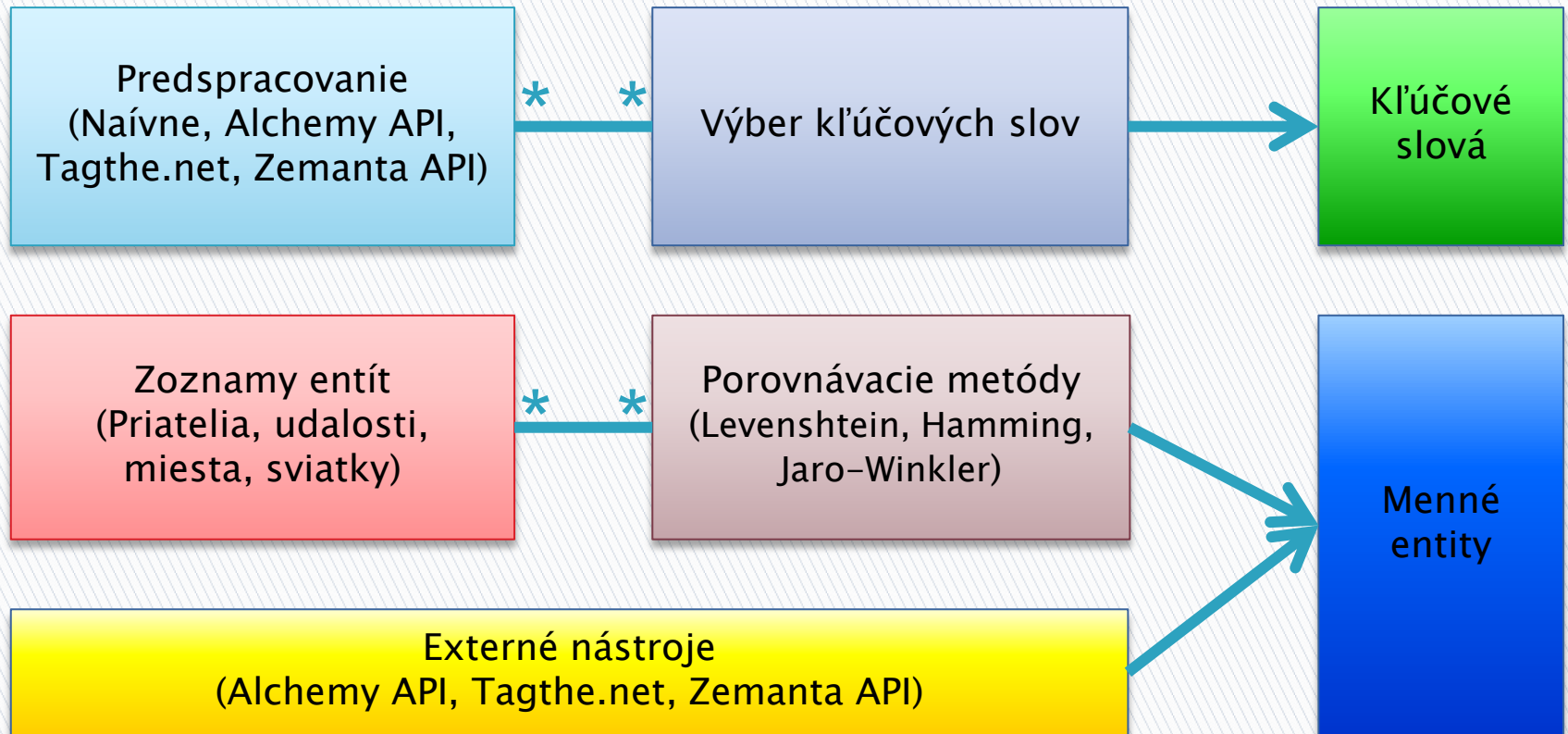


Riešenie nejasnosti

- ▶ Súčet hodnôt nad istou úrovňou
 - Postačuje len v prípade výskytu celého mena v anotácii
- ▶ Využitie doplňujúcich informácií
 - Zohľadnenie zoznamu priateľov hráča
 - Počet výskytu osoby s vlastníkom albumu na iných fotografiách
 - Rodinné a iné vzťahy získané zo sociálnych sietí



Model nezávislých extraktorov



PexAce

- ▶ Založené na hre pexeso
- ▶ Možnosť písania poznámok
- ▶ Vstup
 - Fotografie z úložísk
- ▶ Výstup
 - Fotografie s anotáciami vo forme voľného textu
 - Ďalšie logy (hráč, informácie o hre, vykonané akcie)



PexAce Personal

► Vylepšenia

- Zdieľanie verejných hier
- Možnosť prerušenia a pokračovania
- Vytvorenie hier podľa vlastných nastavení
- Zoznam vytvorených a začatých/ukončených hier
- Riešenie podvádžania

- Systém priateľstiev
- Pokročilý výber fotografií
- Automatické doplnenie anotácií



Fotografie

facebook.

Úložiská

flickr

PexAce Personal

Fotografie a
anotácie

Extraktory

Fotografie a k.
slová

Fotografie
a entity

Agregátor pre kľúčové
slová

Agregátor pre menné
entity

Galéria



Ukážka systému



Welcome


vm19.ucebne.fit.stuba.sk/Home/Welcome

Like You, Josef Buday and 39 others like this.

PexAce

Welcome About

Welcome to PexAce



Sign up now!

And keep your photos organized.

Feedback



Experiment spojený so súťažou

Počet registrácií (Počet súťažiacich)	17(13)
Počet pridaných fotografií	8924
Počet napísaných poznámok	2599
Počet fotografií s poznámkou	1040
Počet odohratých hier	243



Overenie použiteľnosti kľúčových slov – rozhranie



Viera Rozinajova 2,5



Viera Svobodova 2,5



Viera Rozinajova 1,666667



Viera Svobodova 1,666667



Maria Bielikova 1



Alexandra Bielekova 0,8888889



Overenie použiteľnosti kľúčových slov – výsledky

Používateľ	1	2	3	4
Počet priradených	122	179	94	9
Počet správne priradených	25	24	42	8
Úspešnosť	20%	13%	45%	88%
Na prvom mieste	20	24	40	8
Na druhom mieste	5	0	2	0
Priemerná pozícia správne priradenej osoby	1,2	1	1,05	1



Overenie pozorovaním a pomocou dotazníka

- ▶ Intuitívne rozhranie
- ▶ Motivácia kvôli zabudnutým fotoalbumom
- ▶ „Hra je nielen hrateľná ale aj zábavná“

- ▶ Otvorenosť na inovatívne riešenia
- ▶ Zdržanlivosť proti zdieľaniu fotografií
- ▶ Ochota pomáhať priateľom



Možné vylepšenia, plány

- ▶ Implementácia nových funkcionalít (zohľadnenie spätnej väzby)
 - Doplnenie ďalších úložísk (Instagram, Google+)
 - Pokročilý výber fotografií
 - Automatické pridanie anotácií ktoré už raz boli napísané



Sumarizácia

- ▶ Metóda na podporu získavania kľúčových slov k personálnym fotoalbumom
- ▶ Kombinácia rôznych metód a prístupov na získanie presnejších výsledkov
- ▶ Implicitná motivácia bez ponuky výhod alebo cien
- ▶ Rozšíriteľnosť
 - Úložiská
 - Iné typy multimédií (prípadnou výmenou hry)
 - Ďalšie extraktory



Otazka 1

- ▶ Prosím študenta, aby sa vyjadril k spôsobu výberu fotografií do hracej plochy, môže monotónnosť fotografií ovplyvniť kvalitu vytváraných anotácií a motiváciu hráča?

- ▶ TODO



Otazka 2

- ▶ Porovnali ste nové riešenie s pôvodným? Sú získané kľúčové slová kvalitnejšie (presnejšie, podrobnejšie)?

- ▶ TODO

