

<i>Prítomní:</i>	Bc. Róbert Horváth (vedenie)	<i>Vedúci:</i>	Ing. Dušan Zeleník
	Bc. Peter Jurčík	<i>Miesto:</i>	Softvérové štúdio
	Bc. Peter Macko	<i>Dátum a čas:</i>	19.3.2012 16:00
	Bc. Vladimír Ruman	<i>Trvanie:</i>	180minút
	Bc. Peter Sládeček		
	Bc. Maroš Ubreži		
	Bc. Matúš Vacula (zápis)		

Prehľad splnenia úloh z predchádzajúceho stretnutia

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
3206	Navrhnutie a upravenie logovania	Maroš Ubreži	19.3.2012	Ukončená	
3207	Analýza skórovania ES	Vladimír Ruman	19.3.2012	Ukončená	
3235	Upraviť automatický tester	Peter Sládeček	19.3.2012	Ukončená	
3241	Upravenie UI	Peter Jurčík	19.3.2012	Vo vývoji	

Základné body stretnutia

- Skórovanie v ElasticSearch
- Automatický test
- Synonymické indexy
- Vyhľadávanie n-gramov
- TP Cup

Skórovanie v ElasticSearch

Bol analyzovaný vzťah pre výpočet skóre vyhľadanej vety v ElasticSearch. Zakladá sa na počte nájdených slov a ich frekvencie výskytu (TF/IDF). Tento vzťah je možné upraviť podľa našich preferencií, aby sme čo najlepšie určili výslednú vetu prekladu.

Automatický test

Automatický test bol uvedený znovu do prevádzky. Všetky verzie boli otestované týmto testom. Vzorka bola však príliš nízka, aby sme získali štatisticky relevantné údaje.

Synonymické indexy

V ElasticSearch boli vytvorené indexy, ktoré indikujú zhodu aj keď je predmetom vyhľadávania synonymum slova nájdeného v ES. Týmto boli ušetrené dopyty na ES a značne sa zlepšil výkon. Zavedenie týchto indexov však znížilo presnosť prekladu na niektorých vetách a nenašli sa na-

příklad niektoré vety, ktoré boli nájdené v predchádzajúcich verziách prekladača. Zisťovanie príčiny bude predmetom jednej z vyplývajúcich úloh.

Vyhľadávanie n-gramov

Z diskusie vyplynulo, že sa budeme snažiť vyhľadávať text za pomoci n-gramov. Za týmto účelom sa vygenerujú nové indexy nad ES. Ďalej bude potrebné vytvoriť funkcionality pre samotné vyhľadávanie a tiež pre spájanie jednotlivých n-gramov do vety, ktorá bude výsledným prekladom.

TP Cup

Jednou z úloh je vytvoriť prezentáciu a projekt prezentovať na PeWe seminári. Projekt bude prezentovať Peter Macko. Zahájila sa aj práca na plagáte, ktorý bude prezentovať výsledný produkt projektu v súťaži TP Cup.

Prehľad úloh vyplývajúcich z diskusie

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
3310	Vytvorenie prezentácie na PeWe	Peter Macko	21.3.2012	nová	
3311	Vytvoriť index pre ES Shingle	Maroš Ubreži	26.3.2012	nová	
3312	Preskúmať problém so synonymickými indexami	Valdimír Ruman	26.3.2012	nová	
3313	Logovanie	Peter Sládeček	26.3.2012	nová	
3314	Spájanie n-gramov	Matuš Vacula, Róbert Horváth	26.3.2012	nová	