

<i>Prítomní:</i>	Bc. Róbert Horváth	<i>Vedúci:</i>	Ing. Dušan Zeleník
	Bc. Peter Jurčík	<i>Miesto:</i>	Softvérové štúdio
	Bc. Peter Macko	<i>Dátum a čas:</i>	20.02.2012 16:00
	Bc. Vladimír Ruman (zápis)	<i>Trvanie:</i>	120 minút
	Bc. Peter Sládeček (diskusia)		
	Bc. Maroš Ubreži		
	Bc. Matúš Vacula		

Prehľad splnenia úloh z predchádzajúceho stretnutia

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
2253	Automaticke pridávanie zmien do Elastic Search-u a MySQL	Maroš Ubreži	21.02.2012	rozpracovaná	
2254	Google API – automatický tester	Peter Sládeček	21.02.2012	ukončená	18.02.2012
2255	Úprava korpusu-číslovky a pomenovania	Peter Jurčík Vladimír Ruman	21.02.2012	rozpracovaná	
2256	Skupina slov v korpuse	Peter Macko	21.02.2012	rozpracovaná	
2750	Vytvorenie krátkeho abstraktu	Róbert Horváth, Matúš Vacula	15.02.2012	ukončená	15.02.2012
2814	Vytvorenie prihlášky na TP Cup	Róbert Horváth, Matúš Vacula	22.02.2012	rozpracovaná	

Základné body stretnutia

Na stretnutí sa prediskutovali tieto body:

- Vyhodnotenie úloh z minulého stretnutia
- TP Cup
- Zhodnotenie zimného semestra
- Diskusia o ďalšom napredovaní v projekte

Vyhodnotenie úloh z minulého stretnutia

Bc. Peter Sládeček úspešne vyriešil problém s Google API pre automatický tester. Bc. Róbert Horváth a Bc. Matúš Vacula vytvorili abstrakt, ktorý včas odovzdali do systému Easy Chair.

Diskutovalo sa o úprave čísloviek a názvov v korpuse, čo majú na starosti Bc. Vladimír Ruman a Bc. Peter Jurčík. Vedúci Ing. Dušan Zeleník nás informoval o existencii programu, ktorý v texte rozlišuje názvy. Program by nám mohol poslúžiť hlavne pri tagovaní názvov. Rozhodli sme sa, že nie je dôležité rozlišovať medzi rôznymi typmi číselných údajov. Súčasnú riešenie dokáže objaviť číslovky v texte ako aj dátumy a čas. V tejto úlohe sa bude pokračovať aj v nasledujúcom šprinte.

Vedúci nám doplnil informácie o TP Cupe, hlavne sme zistili, ktorú šablónu treba použiť. Tvorba abstraktu bude prebiehať aj v nasledujúcom šprinte.

Zhodnotenie zimného semestra

Na základe pripomienok, ktoré sme dostali, vedúci navrhol tieto zmeny do letného semestra:

Git – je potreba omnoho viac commitovať zdrojový kód a umiestňovať do spoločného repozitáru všetko čo vytvoríme.

Zdrojový kód – treba začať lepšie kontrolovať vytvorený kód. Člen tímu, ktorý vyrieši task, ktorý má pridelený, neuzavrie ho ale nastaví stav feedback. Potom niekto poverený kontrolou kódu skontroluje riešenie a prípadne upraví nedostatky. Až potom sa task v redmine zmení na uzavretý. Na základe tohto vznikol návrh na kontrolu už vytvorené kódu a jeho upravenie.

Tiež sme diskutovali nad pripomienkou Prof. Bielikovej o analýze zložitosti nášho algoritmu pre preklad textu. Túto pripomienku sme zhodnotili ako opodstatnenú.

Smerovanie projektu v letnom semestri

Treba zvýšiť efektívnosť prvotného prekladu textu tak, aby sa znížil počet vygenerovaných viet, ktoré sa potom hľadajú v korpuse.

Ako základný problém do letného semestra sme identifikovali potrebu zaviesť Term frequency slov v našom korpuse.

Pri vyhľadávaní podobných viet v Elasticsearch potrebujeme skvalitniť analyzátor, ktorý sa používa na hľadanie podobnosti. Elasticsearch nám umožňuje navrhnúť a implementovať si vlastný analyzátor. Náš návrh spočíva vo vyhľadávaní podobných slov na základe prvých niekoľko písmen.

Prehľad úloh vyplývajúcich z diskusie

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
2847	Kontrola zdrojového kodu	Matúš Vacula	05.03.2012	nová	
2848	Iplementovat term frequency	Peter Sládeček	05.03.2012	nová	
2849	Zrychlenie restovej služby	Peter Macko	05.03.2012	nová	
2850	Analuzovat možnost zmeny indexu v elasticsearch	Maroš Ubreži	05.03.2012	nová	

Spracoval: Bc. Vladimír Ruman

Overil: Ing. Dušan Zeleník