

<i>Prítomní:</i>	Bc. Róbert Horváth (zápis)	<i>Vedúci:</i>	Ing. Dušan Zeleník
	Bc. Peter Jurčík	<i>Miesto:</i>	Softvérové štúdio
	Bc. Peter Macko (diskusia)	<i>Dátum a čas:</i>	22.11.2011 11:00
	Bc. Vladimír Ruman	<i>Trvanie:</i>	120 minút
	Bc. Peter Sládeček		
	Bc. Maroš Ubreži		
	Bc. Matúš Vacula		

Prehľad splnenia úloh z predchádzajúceho stretnutia

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
1687	Restová služba	Peter Macko	22.11.2011	Ukončená	22.11.2011
1688	Stiahnutie a spracovanie filmových titulkov	Vlado Ruman Peter Jurčík	22.11.2011	Ukončená	21.11.2011
1689	Vyhľadávanie spojení pomocou Elastic search	Peter Macko Maroš Ubreži Matúš Vacula	22.11.2011	Ukončená	22.11.2011
1691	Tvorba dokumentácie pre TP Cup	Matúš Vacula	19.11.2011	Ukončená	18.11.2011
1693	Pripraviť test case pre porovnanie prekladu prekladačov	Peter Sládeček Róbert Horváth	22.11.2011	Ukončená	22.11.2011

Základné body stretnutia

Na stretnutí sme skontrolovali plnenie úloh z predchádzajúceho šprintu. Potom sme vybrali úlohy pre nasledujúci šprint a diskutovali o nasledujúcich témach:

- Získavanie viet do korpusu zo slovenských titulkov
- Zmena testovania na automatické
- Odovzdanie TP Cup prihlášky
- Cieľ vývoja v nasledujúcom šprinte
- Spracovanie čísel a vlastných názvov z textu
- Nahradiť MySQL databázu Elastic Searchom
- Pridanie prekladača a restovej služby do Git repozitára

Získavanie viet do korpusu zo slovenských titulkov

Pri spracovaní titulkov boli problémy s encodingom, ktoré boli úspešne vyriešené. Po vyfiltrovaní databázy titulkov o veľkosti 10000 ostalo len 700 súborov slovenských titulkov. Titulky budú pridané do korpusu.

Zmena testovania na automatické

Testovanie, ktoré bolo vytvorené na predchádzajúcom stretnutí je dobré, avšak chceli by sme mať možnosť zaznamenávať vývoj vo forme zvyšovania úspešnosti. Je preto potrebné vytvoriť automatické testovanie, kde už používateľ nebude zadávať text. Po spustení testovania sa overí úspešnosť prekladov na vopred definovaných vetách. Aby bolo možné zaznamenať postup vývoja a zlepšovanie prototypu, je potrebné výsledky testov uchovávať.

Odovzdanie TP Cup prihlášky

TP cup prihláška bola vypracovaná, vedúci tímu nám dal súhlas k jej odovzdaniu.

Cieľ vývoja v nasledujúcom šprinte

Doterajším cieľom prototypu bolo, aby dokázal vygenerovať všetky možnosti prekladu, medzi ktorými sa pravdepodobne nachádzal aj správny preklad. Od teraz je hlavným cieľom výber správneho(najpravdepodobnejšieho) prekladu a jeho prezentácia.

Spracovanie čísiel a vlastných názvov z textu

Vety obsahujúce vlastné podstatné mená alebo číslice sa nemusia podariť preložiť prostredníctvom nášho prototypu v prípade, že sa v korpuse nenachádzajú presne so zadanými slovami. Zvýšiť úspešnosť prekladu vylepší, ak správne identifikujeme výskyt čísiel a vlastných podstatných mien, ktoré následne dokážeme nahrádzať. Podobným problémom sa zaoberal študent bakalárskeho štúdia, ktoré cieľom bola extrakcia faktov z textu. Úlohou pre nás je analyzovať možnosť využitia jeho práce pre vylepšenie nášho prototypu.

Nahradiť MySQL databázu Elastic Searchom

Jadrom aktuálne vytvoreného prototypu je MySQL databáza. Pre zvýšenie rýchlosti vyhľadávania a umožnenie vyhľadania množiny slov je potrebné, aby jadrom prototypu bol Elastic Search.

Pridanie prekladača a restovej služby do Git repozitára

Git repozitár musí obsahovať všetky súčasti, ktoré sa podieľajú na tvorbe prototypy. Je preto potrebné do neho vložiť zdrojové súbory:

- prekladača,
- restovej služby.

Prehľad úloh vyplývajúcich z diskusie

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
2217	Automatizovať testovanie prekladača	Peter Sládeček Róbert Horváth	06.12.2011	Vo vývoji	
2232	Integrovať Elastic search do prekladača	Peter Macko Maroš Ubreži Matúš Vacula	06.12.2011	Vo vývoji	
2255	Integrovať dokumentáciu 3. a 4. šprintu	Matúš Vacula	29.11.2011	Vo vývoji	
2256	Napísať zápisnicu z 10. Stretnutia a vložiť úlohy do Redmine	Róbert Horváth	29.11.2011	Vo vývoji	
2259	Vytvoriť plán na letný semester	Róbert Horváth	29.11.2011	Vo vývoji	
2261	Vytvoriť backlog projektu	Matúš Vacula	29.11.2011	Vo vývoji	
2265	Vytvoriť zhodnotenie plánov projektu	Vladimír Ruman	29.11.2011	Vo vývoji	
2266	Integrovať metodiky	Peter Jurčík	29.11.2011	Vo vývoji	
2268	Analyzovať možnosti extrakcie faktov z vety	Peter Jurčík Vladimír Ruman	06.12.2011	Vo vývoji	
2286	Pridanie prekladača a restovej služby do Git repozitára	Peter Macko Vladimír Ruman	06.11.2011	Vo vývoji	

Spracoval: Bc. Róbert Horváth

Overil: Ing. Dušan Zeleník