

<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>	13.12.2011 11:00
	Bc. Peter Sládeček (zápis)	<i>Trvanie:</i>	180 minút
	Bc. Maroš Ubreži (diskusia)		
	Bc. Matúš Vacula		
<i>Absencia:</i>	Bc. Vladimír Ruman		

Prehľad splnenia úloh z predchádzajúceho stretnutia

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
2232	Integrovať Elastic Search do prekladača	Peter Macko Maroš Ubreži Matúš Vacula	06.12.2011	Ukončená	
2255	Integrovať dokumentáciu 3. a 4. šprintu	Matúš Vacula	29.11.2011	Ukončená	
2261	Vytvoriť backlog k projektu	Matúš Vacula	29.11.2011	Ukončená	

Základné body stretnutia

Na stretnutí boli prediskutované nasledujúce body:

- Absencia
- Stav úloh z minulého stretnutia
- Prezentácia prekladača aDictIT vedúcemu
- Problém s konektorom Elastic Searchu.
- Prezentácia projektu druhému tímu
- Rozdelenie úloh počas prázdnin

Absencia

Člen tímu Bc. Vladimír Ruman sa nezúčastnil tímového stretnutia z dôvodu osobných povinností, ktoré nebolo možné odložiť. O tomto fakte nás i vedúceho včas informoval prostredníctvom mailu. Daná absencia je považovaná za ospravedlnenú.

Stav úloh z minulého stretnutia

Elastic Search bol úspešne integrovaný. Korpus je načínaný, prekladač v ňom vyhľadáva prekladané vety.

Členom tímu sa podarilo vytvoriť jednotlivé časti dokumentácie k inžinierskemu dielu a k riadeniu. Tie boli tematicky pospájané do jedného celku. Tlačená forma bola odovzdaná vedúcemu. Elektronická verzia riadenia sa vložila do AIS v dvoch častiach, keďže jej veľkosť presahovala povolenú kvótu.

Backlog projektu bol úspešne vytvorený. Identifikuje požiadavky, či rozšírenia na vyvíjaný prekladač. Súčasťou sú aj odhadované časy potrebné na vývoja danej funkcionality.

Prezentácia prekladača aDictIT vedúcemu

Vytvorené riešenie bolo prezentované vedúcemu. Používateľ si môže zvoliť verziu, na ktorej sa daný preklad bude vykonávať. Verzia číslo jeden pracuje s MySQL databázou, druhé riešenie už s integrovaným Elastic Searchom. Výsledky boli otestované na vete: „*It was coincidence.*“. Prvá verzia vrátila ako najpravdepodobnejší preklad vetu „*To bola náhoda.*“. Ide o presné vyhľadanie vety v korpuse alebo preklad slovo po slove. Druhé riešenie vyhľadá pomocou množiny slov v korpuse aj také vety, ktoré neobsahujú presný preklad. Výsledkom prekladu bola veta „*Bola to iba zhoda okolností.*“.

Problém s konektorom Elastic Searchu

Na stretnutí bol identifikovaný problém s konektorom. Ten je zatiaľ nutné vytvárať pri každom dopyte, čo výrazne spomaľuje náš prekladač. Členovia tímu analyzovali možnosti jeho odstránenia. Riešením by mohlo byť zapracovanie návrhového vzoru Singleton. Peter Macko prisľúbil, že sa pozrie na možnosti Tomacatu a pokúsi sa ho v rámci nasledujúcej identifikovanej úlohy odstrániť.

Prezentácia projektu druhému tímu

Dňa 15.12.2011 o 11:00 bude prebiehať prezentácia vytvoreného riešenia druhému tímu. Vytvorili sme nasledujúci kostru prezentácie:

- **motivácia** – preklad textu je viacjazyčný, chceme mať rýchlu a presnú službu jazykovo nezávislú od heuristik a pravidiel, Google API už nie je podporované (iba v platenej verzii)
- **metóda**
 - *jednorazová príprava*
 - *korpus* – príprava sme.sk a tituliek z Lase, možnosť rozšírenia na web
 - *indexovanie* – v korpuse nahradiť číslovky a mená za tagy, vytvoriť štatistický stemmer
 - *pokroky prekladača*
 - *preklad pri každom requeste* - preklad po slovách, zapracovaný slovník PC Translator
 - *vyhľadávanie viet* – neskôr spraviť vyhľadávanie pomocou N-gramov

- **ukážka** – ukázať vyššie spomínanú vetu vo verzii 1 a 2, vybrať dáke cherrypick preklady
- **kvalita testovania a overenie** – spomenúť code review (programovanie v pároch), unit testovanie, automatické porovnávanie výsledkov s Google a Bing (identifikovaný problém alternatívnych prekladov)
- **plán na letný semester**
 - *zrýchlenie prekladu* – Hadoop, zúženie kandidátov cez frekvencie výskytov (slova na vstupe, vety v korpuse)
 - *kontext prekladu* – muž / žena, domény (korpus)
 - *proxy plugin* – preklad označeného textu na stránke
 - *N-gramy* - hľadanie častí vety, ak nie je kompletná
- **možnosti rozšírenia o ďalšie korpuse a jazyky**

Rozdelenie úloh počas prázdnin

Na stretnutí boli prediskutované okruhy tém, ktoré by bolo vhodné implementovať do nášho riešenia počas skúškového obdobia. Vychádzali z predchádzajúceho stretnutia, kde boli identifikované. Členovia tímu vyjadrili názor, v ktorom vyjadrili som súhlas alebo nesúhlas s danou technológiou / rozšírením. Následná diskusia viedla k prijatiu alebo zamietnutiu danej funkcionality.

Prehľad úloh vyplývajúcich z diskusie

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
2252	Analýza Adaptive proxy	Róbert Horváth	21.02.2012	Nová	
2253	Automatické pridávanie zmien do Elastic Searchu a MySQL	Maroš Úbreži Matúš Vacula	21.02.2012	Nová	
2254	Google API - automatický tester	Peter Sládeček	21.02.2012	Nová	
2255	Úprava korpusu - číslovky, pomenovania	Peter Jurčík Vladimír Ruman	21.02.2012	Nová	
2256	Skupina slov v korpuse	Peter Macko	21.02.2012	Nová	