

<i>Prítomní:</i>	Bc. Róbert Horváth (zápis)	<i>Vedúci:</i>	Ing. Dušan Zeleník
	Bc. Peter Jurčík	<i>Miesto:</i>	Softvérové štúdio
	Bc. Peter Macko (vedenie)	<i>Dátum a čas:</i>	12.3.2012 16:00
	Bc. Vladimír Ruman	<i>Trvanie:</i>	180 minút
	Bc. Peter Sládeček		
	Bc. Maroš Ubreží		
	Bc. Matúš Vacula		

## Prehľad splnenia úloh z predchádzajúceho stretnutia

ID	Popis	Pridelená	Predpokladané ukončenie	stav	Skutočné ukončenie
2943	Konfigurovateľnosť prekladača	Peter Sládeček	5.3.2012	ukončená	
2944	Nyhľadávanie cez N-gramy	Róbert Horváth Matúš Vacula	5.3.2012	rozpracovaná	
3045	Úpratať REST-ovej služby	Peter Macko	11.3.2012	ukončená	
3044	Osekanie počtu prekladov slov získaných z prekladača	Maroš Ubreží	11.3.2012	ukončená	
2043	Využitie možnosti zoznamu synonym v Elastic-Search	Maroš Ubreží	11.3.2012	ukončená	

## Základné body stretnutia

- Duplicity v synonymickom slovníku
- Prehľad verzií, ktoré budú dostupné po šprinte
- Vyhodnocovanie metódy, ktorá pracuje s n-gramami
- TP Cup analýza chýb v odovzdanom abstrakte
- Príprava testovania
- Príprava na TP Cup

## Duplicity v synonymickom slovníku

Pri využití synonymického slovníka v ElasticSearch sme narazili na problém s duplicitným výskytom slov v synonymickom slovníku. ElasticSearch však dokázal bez problémov vypočítať index, a preto sme nijak slovník nečistili. Podľa výsledkov testov uvidíme, či je potrebné prerábať databázu synonym, ale môže ostať nezmenená.

## Prehľad verzií, ktoré budú dostupné po šprinte

V aktuálnom šprinte sa dokončuje súčasne niekoľko alternatívnych stratégií pre preklad textu, a preto sme sa pozreli, ktoré všetky by mali byť po šprinte funkčné. Identifikovali sme nasledujúcich 7 stratégií:

- MySQL.
- ElasticSearch.
- ElasticSearch s limitom.
- ElasticSearch s obmedzeným počtom dopytov podľa Term-frequency.
- ElasticSearch s doplnením synonym z slovníka do ElasticSearch.
- ElasticSearch s doplnením synonym z synonymického slovníka do ElasticSearch.
- ElasticSearch vyhľadávanie použitím n-gramov.

## Vyhodnocovanie metódy, ktorá pracuje s n-gramami

Pôvodný zámer bol vyhodnotiť správnosť prekladu n-tíc vzájomným porovnaním všetkých prekladov. Počas implementácie sme si však uvedomili, že by išlo o veľký počet iterácií (exponenciálna návštevnosť), a preto sme sa vyhodnocovanie zjednodušili. Metóda bude hľadať len presné zhody a v prípade, že sa nenájde použije sa skóre z ElasticSearch.

## TP Cup analýza chýb v odovzdanom abstrakte

Dostali sme spätnú väzbu k odovzdanému abstraktu, bol akceptovaný, avšak obsahoval zopár chýb. Prešli sme si spolu nedostatky a keďže to neboli závažné nedostatky, nebolo ich potrebné opravovať priamo na stretnutí. Opravenie a odovzdanie finálnej verzie na tlač dostali na starosť Matúš Vacula a Róbert Horváth.

## Príprava testovania

Keďže na konci šprintu budeme mať viacero fungujúcich stratégií prekladu, dôležité je pre nás kvalitu týchto stratégií vyhodnotiť. Použijeme už existujúci automatický testovač, ktorý sme mali pripravený na konci zimného semestra. Potrebné je ho trochu upraviť, aby dokázal vyhodnocovať všetky naše verzie. Aby bolo vyhodnotenie objektívne, pripraví sa na začiatku testovacia množina niekoľkých viet, ktoré sa použijú pre porovnanie našich a konkurenčných riešení.

## Príprava na TP Cup

Okrem odovzdania rozšíreného abstraktu na tlač je potrebné pripraviť aj obhajobu nášho projektu na súťaž TP Cup. Dôležité je začať pracovať na plagáte A0, ktorý bude na súťaži vystavený a pokúsiť sa zhromaždiť výsledky, ktoré by nás dokázali na súťaži prezentovať.

**Prehľad úloh vyplývajúcich z diskusie**

<b>ID</b>	<b>Popis</b>	<b>Pridelená</b>	<b>Predpokladané ukončenie</b>	<b>stav</b>	<b>Skutočné ukončenie</b>
<b>3206</b>	Navrhnutie a upravenie logovania		19.3.2012	nová	
<b>3207</b>	Analýza skórovania ES	Maroš Ubreží	19.3.2012	nová	
<b>3235</b>	Upraviť automatický tester	Peter Sládeček	19.3.2012	nová	
<b>3241</b>	Upravenie UI	Peter Jurčík	19.3.2012	nová	

---

**Spracoval:** Bc. Róbert Horváth

---

**Overil:** Ing. Dušan Zeleník