

Zápisnica - Stretnutie č.15

Tím: 10 - Research RANK

Téma stretnutia: Konzultácia problémov pedagogickému vedúcemu tímu (zákazníkovi), prezentácia funkcionality front endu, naplánovanie ďalšieho šprintu

Dátum stretnutia: 3.4.2014

Čas stretnutia: 10:00-13:00

Miesto stretnutia: Jobsovo softvérovo štúdio

Stretnutie viedol: Bc. Tomáš Jánošík

Prítomní:

Bc. Michael Gloger
Bc. Tomáš Jánošík
Bc. Daniel Kíč
Bc. Šimon Kompas
Bc. Rastislav Kostrab
Bc. Stanislav Kubica

Nepřítomní:

Zapisovateľ: Bc. Daniel Kíč

Obsah stretnutia:

- Ukončenie šprintu:
 - V šprinte Heineken sa podarilo zmeniť rozhranie ručného párovania publikácií tak, že informácie o publikácií, ktorej duplikát hľadáme, sa pri prezeraní rozhrania a rolovaní udržiujú v hornej časti lišty. Takto sa docielilo to, že pri párovaní sú informácie publikácie stále k dispozícii bez rolovania stránky.
 - Pri samotnom napájaní ručného párovania sa však vyskytol systémový problém v zjednotení duplikátov v databáze. Táto chyba bude odstránená v ďalšom šprinte.
 - Počas šprintu sa taktiež skúmala nestabilita servera, ktorá sa nakoniec však nepotvrdila.
 - Taktiež sa upravoval a optimalizoval dátový model, a komponent ORM(objektovo relačný mapovač). Zmeny boli zamerané na zvýšenie rýchlosti komunikácie s databázou.
 - V databáze pribudli nové informácie zo zdroja WoS. Bol pripravený a odladený automatizovaný parser a loader do databázy, ktorý komunikuje s webovou službou WoS a získava z nej informácie. Parser bol doladený, keďže v riešení z minulého šprintu sa objavila chyba pri práci s WSDL.
 - Do databázy boli pridané publikácie z EUBA a SPU zo zdroja WoS.

- Boli preskúmané možnosti zrýchlenia databázy. Došlo sa k záveru, že ďalšie optimalizácie pri povahe práce s databázou by neboli vhodné.
- Boli preskúmané údaje stiahnuté zo zdroja Scopus. Pri publikáciách z TUKE boli zistené nedostatky následkom ktorých boli publikácie z TUKE stiahnuté ešte raz. Problém bola drobná chyba v kóde.
- Do systému na sťahovanie údajov zo služby Scopus sa spravil kontrolný systém na zisťovanie Scopus ID. Toto je optimalizácia systému, ktorý tak sťahuje relevantnejšie údaje. Údaje zo zdroja Scopus boli pomocou tohto systému následne obohatené.
- Boli preskúmané možnosti optimalizácie výberu kandidátov na párovanie. Skúmali sa optimalizácie na strane databázy ako je indexácia ale aj na strane samotného algoritmu. Na základe pokusov bolo rozhodnuté optimalizovať samotný algoritmus na základe písomného návrhu. Tento návrh bol implementovaný, avšak k jeho testovaniu nedošlo následkom chyby prostredníka s databázou, ktorej riešenie je naplánované na ďalší šprint.
- Taktiež sa dokončil import zo zdroja Crepč, ktorý by teraz mal byť kompletný.
- Bol implementovaný systém do komponentu C Komparátor, ktorý meria jeho výkon a má výstup do súboru. Na základe údajov následne zozbieraných z C komparátora bolo rozhodnuté, že je potrebné optimalizovať J Komparátor, lebo výkon C komparátora je dostačujúci.
- Bola vykonaná formálna inšpekcia zdrojových kódov. Na základe tejto inšpekcie bol vypracovaný dokument s pripomienkami pre jednotlivé zdrojové súbory.
- Bola vykonaná refaktorizácia dokumentácie. Nová štruktúra viacej vyhovuje všeobecným štandardom. Vzhľadom na to, že pribudla povinnosť uviesť pomerne veľkú množinu informácií, nie je dokumentácia plne doplnená podľa novej štruktúry. Dokončenie tejto úlohy bolo presunuté na ďalší šprint.
- Celkovo boli úlohy šprintu splnené, avšak jedna z hlavných úloh a to je manuálne párovanie zostala nedokončená. To kvôli neočakávanému problému v nástroji na spájaní publikácií (Join publikácií). Dokončenie tejto úlohy je naplánované na ďalší šprint.
- Podarilo sa nám dosiahnuť lepšiu úroveň vo vykonávaní našich manažérskych úloh.
- Celkovo šprint nebol splnený na 100%, ale i tak si myslíme, že bol úspešný v mnohých aspektoch, či už idem o manažovanie, tímovú prácu alebo aj v mnohých úlohách.
- IIT.SRC –TP cup;
 - Vzhľadom na blížiaci sa termín IIT.SRC sa na stretnutí diskutovalo a dokumentoch a náležitostiach, ktoré je potrebné odovzdať. Dohodlo sa, že sa fotka tímu, ktorá je vyžadovaná urobí v pondelok, 7.4.2014 pri neformálnom tímovom stretnutí.
 - V súvislosti s IIT.SRC sa po diskusií prijalo rozhodnutie, že členovia tímu prídu na ďalšie tímové stretnutie vo štvrtok 10.4.2014 s nápadmi na poster, spoločný znak, ktorý budú členovia tímu na IIT.SRC nosiť a s nápadmi na prezentáciu na IIT.SRC.
 - Dotazník, ktorý je potrebné vyplniť do 10.4.2014 bude vyplnení po neformálnom stretnutí 7.4.2014. Spolu s fotkou budú zaslané na revíziu pedagogickému vedúcemu tímu.

- Uznesenia tímu:
 - Stand up meeting v piatok a sobotu sú od tohto dátumu dobrovoľné. Stand up stretnutia budú povinné v nedeľu, pondelok, utorok a stredu. Štvrtok je nepovinný, pretože vo štvrtky bývajú oficiálne stretnutia tímu.
- Kontrola úloh a review zo strany zákazníka. Nasledujúce pripomienky je treba zapracovať:
 - Je nutné upraviť obsah databázy, kvôli dodržaniu formálnej zrozumiteľnosti jednotlivých názvu polí, keďže to v momentálnej verzii je sémanticky nejasné.

Otvorené úlohy & backlog:

ID	Popis úlohy	Zodpovednosť	Stav	Založené
TP-114	Preložiť chybové hlasy validácie formulárov do slovenčiny		Backlog	10.2.2014
TP-122	Pridať autentifikáciu na všetky podstranky - len prihlásený a schválený používateľ môže prístupovať do DB		Backlog	20.2.2014
TP-124	Pridať možnosť konfigurácie zobrazovaných stĺpcov pri zoznamoch		Backlog	20.2.2014
TP-126	Pridať fazety na filtrovanie		Backlog	20.2.2014
TP-129	Pridať zoznam citujúcich autorov		Backlog	20.2.2014
TP-130	Pridať možnosti vyhľadávania ohlasov podľa percenta zhody		Backlog	20.2.2014
TP-131	Používať externé authority viaf.org, dbpedia.org		Backlog	20.2.2014
TP-185	Opraviť možnosti filtrovania bez diakritiky		Backlog	26.3.2014
TP-203	Autocitácie		Backlog	3.4.2014
TP-207	Napojenie vyberu atributov na DB		Backlog	3.4.2014

Úlohy pre ďalší šprint:

ID	Popis úlohy	Zodpovednosť
TP-69	Vytvoriť package v java pre Scheduler	Bc. Daniel Kíč

TP-153	Premapovat source na original_id	Bc. Tomáš Jánošík
TP-162	Vytvorenie tried pre ciselniky	Bc. Rastislav Kostrab
TP-174	Pridat do publikacii parovanie merge join	Bc. Tomáš Jánošík
TP-175	Namapovanie autorov v ramci porovnaných publikacii	Bc. Tomáš Jánošík
TP-178	Vytvorenie ANT scriptu na zrychlenie deployovania na server	Bc. Rastislav Kostran
TP-196	Kompozicia navrhnu na optimalizaciu vyberu kandidatov	Bc. Daniel Klíč
TP-199	Napisat join tool na rucne parovanie vo frontende	Bc. Tomáš Jánošík
TP-200	Spustit parovanie publikacii	Bc. Tomáš Jánošík
TP-201	Stiahnut WoS ohlasy	Bc. Stanislav Kubica
TP-202	Scopus ohlasy	Bc. Rastislav Kostrab
TP-204	Zobrazovanie citacii vo frontende	Bc. Michael Gloger
TP-205	Odelenie publikacie a zdrojovej publikacie	Bc. Šimon Kompas
TP-206	Rozhranie na zobrazovanie atributov	Bc. Michael Gloger
TP-208	Pridat k rucnemu parovaniu potvrdenie	Bc. Michael Gloger

Retrospektíva

START DOING	STOP DOING	KEEP DOING
Písať do jiri pri log work koľko času odhadom ostáva	Pridávať nový task ktorý nie je v user story šprintu	denné stretnutia
Konsolidovať náš spôsob komunikácie		komunikácia
		Code-review
		Lepšia úroveň dokumentácie

User stories na ďalší šprint:

User story 1 – Pripraviť databázu, párovacie algoritmy a frontend na testovanie zákazníkom

Kedže nemáme normalizovanú testovaciu množinu, na ktorej by sme overili vlastnosti a správnosť párovacích algoritmov, túto činnosť musí vykonať zákazník (zástupca naparsovaných inštitúcií). My mu k tomu musíme pripraviť podklady a poskytnúť podporu v našej aplikácii. Zahŕňa to dokončiť parsovanie exportov minimálne z univerzít SPU, EU a TUKE, dokončiť párovacie algoritmy, otestovať ich správnu funkčnosť a následne vyčísliť hodnoty párovania pre všetky záznamy v databáze. Kedže zákazník nebude odborník na databázu, potrebujeme dokončiť zobrazovanie podobností na frontende tak, aby si to mohol jednoducho poprezerat' a zistiť, ktoré publikácie by mali byť napárované a nie sú. Užitočné bude zistiť aj štatistiky k jednotlivým algoritmom ohľadne ich trvania a efektivity.

User story 2 – Hľadanie ohlasov, aktualizácia údajov

Vzhľadom na to, že systém ma predovšetkým slúžiť výskumníkom na hľadanie ohlasov, je potrebné pripraviť zákazníkovi rozhranie na zobrazenie ohlasov, ktoré bude jednoduché a intuitívne na poskytnutie relevantnej informácie ohľadom ohlasov. Na zobrazovanie reálnych ohlasov je potrebné tieto stiahnuť zo zdrojov ako sú WoS a Scopus. Ohlasy je potrebné stiahnuť aspoň pre SPU, EU a TUKE. Používateľ chce tieto ohlasy mať bez autocitácií a taktiež bež duplikátov. Na to je potrebné pripraviť subsystém a pripraviť potrebné rutiny na deduplikáciu ohlasov ako aj rutiny na detekciu autocitácií. Je potrebné dohodnúť spôsob, akým sa budú ohlasy ukladať do databázy.

User story 3 – Personalizácia rozhrania

Zákazník by rád vedel ovplyvniť spôsob, akým sa mu zobrazujú údaje a zároveň chce, aby mu zobrazovalo len tie údaje, ktoré ho zaujímajú. Vzhľadom na to, že používateľ nebude technologický expert, tak mu musí byť poskytnuté jednoduché intuitívne rozhranie na selekciu toho, čo si praje mať zobrazené a čo nie. Na toto je potrebné pripraviť rozhranie a zároveň implementovať uloženie do databázy. Keď sa používateľ pripojí, chce, aby sa mu jeho nastavenia obnovili a preto je potrebné urobiť rutinu na znovu načítanie údajov z databázy.