

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

Jakub Adam, Monika Filipčíková, Andrej Švec, Filip Vozár

moderateIT

Dokumentácia k riadeniu z predmetu Tímový projekt

Študijný program: Informačné systémy / Softvérové inžinierstvo

Vedúci tímu: Ing. Jakub Šimko, PhD.

Zimný semester šk. r. 2016 / 2017

Big picture	6
Úvod	6
Roly členov tímu a podiel práce	6
Roly členov tímu	6
Podiel práce	6
Aplikácie manažmentov	6
Manažment kvality a podpory vývoja	6
Manažment plánovania	7
Manažment dokumentácie	7
Manažment komunikácie	7
Sumarizácie šprintov	9
Zápis zo stretnutia #1	9
Priebeh stretnutia	9
Nové úlohy	9
Zápis zo stretnutia #2	10
Priebeh stretnutia	10
Nové úlohy	10
Zápis zo stretnutia #4	11
Priebeh stretnutia	11
Nové úlohy	11
Aktuálny backlog	12
Aktuálny stav úloh	12
Zápis zo stretnutia #5	13
Priebeh stretnutia	13
Nové úlohy	13
Aktuálny backlog	14
Aktuálny stav úloh	14
Zápis zo stretnutia #6	15
Priebeh stretnutia	15
Úlohy na ďalší týždeň	15
Aktuálny backlog	16
Aktuálny stav úloh	16
Retrospektíva šprintu "ARA"	17

Čo sa podarilo	17
Čo sa nepodarilo	17
Čo plánujeme zlepšiť	17
Review šprintu "ARA"	18
Šprint review	18
Zápis zo stretnutia #7	20
Priebeh stretnutia	20
Úlohy na ďalší týždeň	20
Aktuálny backlog	21
Aktuálny stav úloh	21
Zápis zo stretnutia #8	22
Priebeh stretnutia	22
Úlohy na ďalší týždeň	22
Aktuálny backlog	23
Aktuálny stav úloh	23
Retrospektíva šprintu "Bubo bubo"	24
Zhrnutie	24
V čom pokračujeme?	24
Čo prestať robiť?	24
Čo začať robiť?	25
Review šprintu "Bubo bubo"	26
Šprint review	26
Zápis zo stretnutia #9	28
Priebeh stretnutia	28
Úlohy na ďalší týždeň	28
Aktuálny backlog	29
Aktuálny stav úloh	29
Používané metodiky	30
Metodika pre tvorbu dokumentácie	30
Zápisnice z tímových stretnutí	30
Zápis retrospektívy	30
Zápis sprint review	31
Web stránka projektu	31
Dokumentácia k moderateIT API	32
Jira exporthy	32

Globálna retrospektíva ZS/LS	32
Čo robíme dobre?	32
Čo robíme zle?	32
Návrhy na zlepšenie	33
Príloha A: Motivačný dokument	35
Big picture	41
Úvod	41
Globálne ciele pre ZS/LS	41
Nasadenie do SME.sk	41
Ďalší vývoj	42
Celkový pohľad na systém	43
Úvod do architektúry	43
Rozmiestnenie a nasadenie komponentov	44
Dátový model	45
Moduly	45
Moduly systému	46
Autentifikačný server	46
Analýza	46
Návrh	47
Implementácia	47
Testovanie	47
Klasifikátor	48
Analýza	48
Návrh	48
Implementácia	48
Testovanie	49
Diakritikovač	49
Analýza	49
Návrh	49
Implementácia	49
Testovanie	50
Portál	50
Analýza	50
Návrh	50
Implementácia	50

Testovanie	51
API	51
Analýza	51
Návrh	51
Implementácia	51
Testovanie	52

Dokumentácia k riadeniu

Big picture

Úvod

Riadenie projektu je veľmi dôležitý proces, na ktorý je potrebné vynaložiť nemalé úsilie. Je potrebné, aby boli pokryté všetky časti projektu. Či už ide o samotný vývoj projektu, testovanie ba dokonca aj dokumentáciu. Je potrebné, aby boli na začiatku stanovené jasné pravidlá podľa, ktorých sa budú riadiť členovia projektu, v závislosti od pozícií a roli, ktoré budú zastávať. Bez správneho manažmentu nie je možné efektívne robiť pokroky či už ide o veľký projekt alebo o projekt, na ktorom sa podieľa malé množstvo ľudí.

Aj napriek tomu, že náš tím je tvorený len štyrmi členmi, je užitočné, aby sme si niektoré veci jasne zadefinovali a robili ich takmer automaticky. Práve pomocou zlepšovania sa v procesoch a zadefinovaní si postupov pri vyvíjaní projektu vylepšíme problémy, ktoré predtým nastali.

Roly členov tímu a podiel práce

Roly členov tímu

Scrummaster - Monika Filipčíková

Zapisovateľ - Jakub Adam

Zapisovateľ do systému JIRA - Andrej Švec

Podiel práce

Každý sa podieľal rovnakým dielom na obidvoch častiach dokumentácie.

Aplikácie manažmentov

Manažment kvality a podpory vývoja

Keďže na tomto projekte ako tím pracujeme už takmer 2 roky, už pred začatím tímového projektu sme mali zaužívané niektoré postupy, ktorých aplikácia nám mala zjednodušiť vývoj a orientáciu v cudzích kódach v rámci tímu a pomôcť nám udržať kvalitu vyprodukovaného kódu na čo najvyššej úrovni.

Išlo napríklad o dohodu o štýle pomenovania jednotlivých detektorov, premenných, metód ... Aj keď sa niekedy môžu podobné dohody zdať nepodstatné, my sme ich prínos ocenili.

Manažment plánovania

Keďže sme pred začatím tímového projektu mali len veľmi malé skúsenosti so scrumom, neuvedomovali sme si, aké dôležité je kvalitné plánovanie úloh na projekte.

Počas celej práce na projekte sme sa snažili využívať praxou overené SCRUM metódy. Z počiatku pre nás bolo náročné ohodnocovanie taskov s využitím prístupu "Planning poker". To bolo spôsobené tým, že nie každý člen tímu vždy vedel, čo všetko sa daného tasku týka, aj napriek tomu, že kvalita opisov úloh bola na veľmi dobrej úrovni. Postupom času sme sa však aj v tomto zdokonalili.

Ďalším kritickým bodom bolo plánovanie toho, koľko story-pointov stihneme za šprint dodať. To sa ukázalo hlavne na nestihnutí všetkých úloh v šprinte "Bubo budo". Pri plánovaní sme totiž nezohľadnili plánované zápočty a odovzdávania projektov na iných predmetoch. S týmto je spojený aj správny výber úloh do šprintu. Bolo potrebné vyberať úlohy tak, aby mal každý člen tímu počas šprintu podobný počet story-pointov. To bolo kritické hlavne v prípadoch, že sme počas šprintu potrebovali urobiť viacero úloh konkrétneho typu, ktoré boli svojou povahou určené pre konkrétneho člena tímu, keďže nie každý sa rozumieme všetkému.

Manažment dokumentácie

Aj keď to nie je najobľúbenejšia časť práce na projekte, uvedomujeme si, že tvorba dokumentácie je v našom prípade veľmi dôležitá. Aj keď zápisnice zo stretnutí, sprint review a retrospektívy šprintov, ktorých spisovanie mal na starosti manažér dokumentácie pre nás malo viacero výhod, tak najdôležitejšou časťou dokumentácie bola dokumentáciu k moderateIT API, ktoré je hlavnou časťou nášho projektu.

Keďže paralelne s prácou na našom projekte sa pracuje aj na zmene diskusnej platformi zákazníka, ktorý má záujem do nej integrovať nami vytvorenú službu, bolo pre nás veľmi dôležité kvalitne zdokumentovať používanie tohto API. Táto úloha sa s vysokou prioritou objavila hneď v prvom šprinte a myslíme si že jej kvalita je na veľmi vysokej úrovni. Táto dokumentácia je zverejnená online, takže zákazník má k dispozícii vždy aktuálnu verziu dokumentácie s podrobne spracovanými informáciami.

Manažment komunikácie

Komunikácia je pri tímových projektoch veľmi dôležitá. Keďže vieme, že nie je dostatočné stretávať sa len raz do týždňa, zaviedli sme viacero spôsobov komunikácie.

V rámci tímu používame na neformálnu komunikáciu aplikáciu Slack. Tento nástroj sa nám počas práce na projekte veľmi osvedčil. V tomto nástroji máme vytvorených viacero kanálov, z ktorých každý je prioritne určený na inú časť práce na projekte. Tento nástroj máme integrovaný s nástrojom JIRA, ktorý používame na riadenie úloh. Takže hneď ako sa vytvorí nová úloha, alebo nejakým spôsobom upraví staršia, každý člen tímu je na to hneď upozornený. Rovnako je to aj pri zmene statusu úlohy.

Taktiež máme so slackom integrovaný aj projekt na Bitbuckete, takže dostávame informácie aj novo vzniknutých commit-och, takže členovia tímu sa hneď môžu pustiť do code review k tomuto commitu.

Na oficiálnejšiu komunikáciu k taskom využívame možnosti pridávania komentárov k úloham priamo v nástroji JIRA.

Na komunikáciu so zákazníkom využívame prioritne mailovú komunikáciu.

Už počas prvého šprintu sme si uvedomili, že sme produktívnejší keď sa spolu stretávame a programujeme spoločne. Preto sme sa rozhodli okrem pravidelného týždňového stretnutia zaviesť aj pravidelné stretnutia určené práve na spoločné programovanie, čo sa nám veľmi odvedčilo.

Sumarizácie šprintov

Zápis zo stretnutia #1

Dátum: 21.09.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Jakub Šimko
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Jakub Šimko
Andrej Švec
Filip Vozár

Priebeh stretnutia

- Diskusia o SCRUM-e
- Diskusia o smerovaní projektu v zimnom semestri

Nové úlohy

- Vybrať systém na management projektu, ktorý bude podporovať definované typy user stories, úlohy, fidelity, odhad v story pointoch, vytváranie šprintov
- Naplniť backlog

Zápis zo stretnutia #2

Dátum: 28.09.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Jakub Šimko
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Jakub Šimko
Filip Vozár

Priebeh stretnutia

- Diskusia o SCRUM-e

Nové úlohy

- Vybrať systém na management projektu, ktorý bude podporovať definované typy user stories, úlohy, fidelity, odhad v story pointoch, vytváranie šprintov
- Naplniť backlog

Zápis zo stretnutia #4

Dátum: 12.10.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Jakub Šimko
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Jakub Šimko
Filip Vozár

Priebeh stretnutia

- Diskusia o začatí šprintu
- Zadefinovanie "Definition of done" pre programovacie a analytické úlohy
- Zadefinovanie spôsobu určovania názvov šprintov
- Začiatok šprintu
 - Zoradenie user stories podľa priority
 - Ohodnotenie user stories (Planning poker)
 - Výber user stories do šprintu

Nové úlohy

- Pracovať na úlohách z aktuálneho šprintu • Pridávať nové úlohy do backlogu v prípade potreby

Aktuálny backlog

T	Key	Summary	Reporter	P	Created
	MD-16	Spraviť web na tlač	Filip Vozar	↑	11. Oct 2016
	MD-15	Webhooky na vrátenie výsledkov	Andrej Svec	↑	09. Oct 2016
	MD-14	Filtrovanie a zoraďovanie v zozname komentárov	Andrej Svec	↑	09. Oct 2016
	MD-12	Upraviť farby na webe	Andrej Svec	↓	09. Oct 2016
	MD-11	Opraviť existujúce testy	Andrej Svec	↑	05. Oct 2016
	MD-9	Autentifikácia voči portálu pomocou access tokenu uloženého v kúčke :)	Andrej Svec	↑	05. Oct 2016
	MD-8	Vytvoriť v databáze entitu autora pokiaľ neexistuje pri vytváraní jeho prvého príspevku	Andrej Svec	↑	05. Oct 2016
	MD-7	Vytvoriť endpointy pre zápis akcií Approve, Reject, Ban, Report	Andrej Svec	↑	05. Oct 2016
	MD-6	Dokumentácia k základným API volaniam	Andrej Svec	↑	05. Oct 2016
	MD-5	Zobraziť podrobné štatistiky autora na novej karte	Andrej Svec	↑	05. Oct 2016
	MD-4	Zobraziť podrobné štatistiky komentáru na zvlášť stránke + Refaktor javascriptu	Andrej Svec	↑	05. Oct 2016
	MD-3	Mini iframe so zobrazením celkového hodnotenia komentáru a autora	Andrej Svec	↑	05. Oct 2016
	MD-2	Zobraziť 100 najhorších príspevkov za posledných 24-72 hodín (posuvník s časom)	Andrej Svec	↑	04. Oct 2016
	MD-1	Zobraziť 100 najhorších príspevkov za posledných 24 hodín	Andrej Svec	↑	04. Oct 2016

Aktuálny stav úloh

T	Key	Summary	Reporter	P	Created	Assignee	Status
	MD-29	MD-16 / Urobiť web	Andrej Svec	↑	12. Oct 2016	Jakub Adam	TO DO
	MD-28	MD-16 / Pripraviť server	Andrej Svec	↑	12. Oct 2016	Filip Vozar	TO DO
	MD-26	MD-4 / Zobrazit štatistiky komentara v novom okne	Andrej Svec	↑	12. Oct 2016	Monika Filipcikova	TO DO
	MD-25	MD-4 / Refaktor javascriptu pre zobrazovanie statistik a grafov	Andrej Svec	↑	12. Oct 2016	Andrej Svec	TO DO
	MD-16	Spraviť web na tlač	Filip Vozar	↑	11. Oct 2016	Jakub Adam	TO DO
	MD-9	Autentifikácia voči portálu pomocou access tokenu uloženého v kúčke :)	Andrej Svec	↑	05. Oct 2016	Jakub Adam	TO DO
	MD-8	Vytvoriť v databáze entitu autora pokiaľ neexistuje pri vytváraní jeho prvého príspevku	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	TO DO
	MD-7	Vytvoriť endpointy pre zápis akcií Approve, Reject, Ban, Report	Andrej Svec	↑	05. Oct 2016	Filip Vozar	TO DO
	MD-6	Dokumentácia k základným API volaniam	Andrej Svec	↑	05. Oct 2016	Filip Vozar	TO DO
	MD-5	Zobraziť podrobné štatistiky autora na novej karte	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	TO DO
	MD-4	Zobraziť podrobné štatistiky komentáru na zvlášť stránke + Refaktor javascriptu	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	TO DO
	MD-3	Mini iframe so zobrazením celkového hodnotenia komentáru a autora	Andrej Svec	↑	05. Oct 2016	Andrej Svec	TO DO

Zápis zo stretnutia #5

Dátum: 19.10.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Andrej Švec
Filip Vozár

Priebeh stretnutia

- Diskusia o priebehu šprintu
- Diskusia o problémoch zistených pri code review
- Diskusia o veľkosti user stories a ich delení na task-y
- Revíza aktuálneho backlogu
- Diskusia o nasadení online dokumentácie
- Diskusia o požiadavkách od sme.sk týkajúcich sa úloh v aktuálnom šprinte
- Riešenie problému s mergovaním vetiev

Nové úlohy

- Pracovať na úlohách z aktuálneho šprintu
- Spojiť sa s ľuďmi zo sme.sk a informovať ich o aktuálnom prograse

Aktuálny backlog

T	Key	Summary	Reporter	P	Created
	MD-34	Lepsie handlovenie couchbase (a redis) connection - po timeoute failuju aj dalsie pokusy, aj ked sa problem s timeoutom vyriesil a couchbase je plne dostupny	Filip Vozar	↑	18. Oct 2016
	MD-32	Odoslat formular "Try our service" Enterom, SHIFT + Enter je newline	Andrej Svec	↑	17. Oct 2016
	MD-31	ANALYZA Co ma vratit off-topic detektor, ked nie je parent	Andrej Svec	↑	17. Oct 2016
	MD-30	Premysliet a implementovat monitoring a logovanie	Filip Vozar	↑	15. Oct 2016
	MD-24	Nahrat SME dáta z roku 2016 do couchbase, aby sme mohli trénovať	Andrej Svec	↑	12. Oct 2016
	MD-22	Pridanie knižnice fuzzy do projektu	Monika Filipcikova	↑	12. Oct 2016
	MD-21	Integracia dvoch modelov fuzzy + klasifikator	Monika Filipcikova	↑	12. Oct 2016
	MD-20	Implementácia pravidiel pre fuzzy	Monika Filipcikova	↑	12. Oct 2016
	MD-19	SLIDO - analýza detektorov	Monika Filipcikova	↑	12. Oct 2016
	MD-18	ANALYZA najdenie fuzzy kniznice, vytvorenie simple prototypu	Monika Filipcikova	↑	12. Oct 2016
	MD-17	Admin rozhranie pre zakaznikov	Filip Vozar	↑	12. Oct 2016
	MD-15	Webhooky na vrátenie výsledkov	Andrej Svec	↑	09. Oct 2016
	MD-14	Filtrovanie a zoraďovanie v zozname komentárov	Andrej Svec	↑	09. Oct 2016
	MD-12	Upraviť farby na webe	Andrej Svec	↓	09. Oct 2016
	MD-11	Opraviť existujúce testy	Andrej Svec	↑	05. Oct 2016
	MD-2	Zobraziť 100 najhorších príspevkov za posledných 24-72 hodín (posuvník s časom)	Andrej Svec	↑	04. Oct 2016
	MD-1	Zobraziť 100 najhorších príspevkov za posledných 24 hodín	Andrej Svec	↑	04. Oct 2016

Aktuálny stav úloh

T	Key	Summary	Reporter	P	Created	Assignee	Status
	MD-29	MD-16 / Urobiť web	Andrej Svec	↑	12. Oct 2016	Jakub Adam	IN PROGRESS
	MD-28	MD-16 / Pripraviť server	Andrej Svec	↑	12. Oct 2016	Filip Vozar	DONE
	MD-26	MD-4 / Zobrazit statistiky komentara v novom okne	Andrej Svec	↑	12. Oct 2016	Monika Filipcikova	TO DO
	MD-25	MD-4 / Refaktor javascriptu pre zobrazovanie statistik a grafov	Andrej Svec	↑	12. Oct 2016	Andrej Svec	IN PROGRESS
	MD-16	Spraviť web na timak	Filip Vozar	↑	11. Oct 2016	Jakub Adam	TO DO
	MD-9	Autentifikácia voči portálu pomocou access tokenu uloženého v kúikine :)	Andrej Svec	↑	05. Oct 2016	Jakub Adam	IN REVIEW
	MD-8	Vytvoriť v databáze entitu autora pokiaľ neexistuje pri vytváraní jeho prvého príspevku	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	IN PROGRESS
	MD-7	Vytvoriť endpointy pre zápis akcií Approve, Reject, Ban, Report	Andrej Svec	↑	05. Oct 2016	Filip Vozar	TO DO
	MD-6	Dokumentácia k základným API volaniam	Andrej Svec	↑	05. Oct 2016	Filip Vozar	IN REVIEW
	MD-5	Zobraziť podrobné štatistiky autora na novej karte	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	TO DO
	MD-4	Zobraziť podrobné štatistiky komentáru na zvlášť stránke + Refaktor javascriptu	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	DONE
	MD-3	Mini iframe so zobrazením celkového hodnotenia komentáru a autora	Andrej Svec	↑	05. Oct 2016	Andrej Svec	IN REVIEW

Zápis zo stretnutia #6

Dátum: 26.10.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Andrej Švec
Filip Vozár

Priebeh stretnutia

- Diskusia o riešení výnimiek v kóde
- Review
- Slávnostné uzavretie šprintu "ARA"
- Retrospektíva šprintu "ARA"
- Diskusia o stretávaní počas týždňa
- Otváranie šprintu
 - Pridávanie taskov
 - Zoradzovanie taskov
 - Ohodnocovanie taskov (Planning poker)
 - Vyberanie taskov do šprintu
- Slávnosné spustenie šprintu "Bubo bubo"

Úlohy na ďalší týždeň

- Pracovať na úlohách z aktuálneho šprintu.

Aktuálny backlog

T	Key	Summary	Reporter	P	Created
	MD-38	Spustiť detektory na SME dátach z roku 2016	Andrej Svec	↑	21. Oct 2016
	MD-37	Nastaviť zalohovanie databaz a konfigurácii clustra	Filip Vozar	↑	21. Oct 2016
	MD-34	Lepšie handlovanie couchbase (a redis) connection - po timeoute failuju aj dalsie pokusy, aj ked sa problem s timeoutom vyriesil a couchbase je plne dostupny	Filip Vozar	↑	18. Oct 2016
	MD-32	Odoslat formular "Try our service" Enterom, SHIFT + Enter je newline	Andrej Svec	↑	17. Oct 2016
	MD-31	ANALYZA Co ma vratit off-topic detektor, ked nie je parent	Andrej Svec	↑	17. Oct 2016
	MD-30	Premysliet a implementovat monitoring a logovanie	Filip Vozar	↑	15. Oct 2016
	MD-24	Nahrat SME dáta z roku 2016 do couchbase, aby sme mohli trénovať	Andrej Svec	↑	12. Oct 2016
	MD-22	Pridanie knižnice fuzzy do projektu	Monika Filipcikova	↑	12. Oct 2016
	MD-21	Integracia dvoch modelov fuzzy + klasifikator	Monika Filipcikova	↑	12. Oct 2016
	MD-20	Implementácia pravidiel pre fuzzy	Monika Filipcikova	↑	12. Oct 2016
	MD-19	SLIDO - analýza detektorov	Monika Filipcikova	↑	12. Oct 2016
	MD-18	ANALYZA najdenie fuzzy kniznice, vytvorenie simple prototypu	Monika Filipcikova	↑	12. Oct 2016
	MD-17	Admin rozhranie pre zakaznikov	Filip Vozar	↑	12. Oct 2016
	MD-15	Webhooky na vrátenie výsledkov	Andrej Svec	↑	09. Oct 2016
	MD-14	Filtrovanie a zoraďovanie v zozname komentárov	Andrej Svec	↑	09. Oct 2016
	MD-12	Upraviť farby na webe	Andrej Svec	↓	09. Oct 2016
	MD-11	Opraviť existujúce testy	Andrej Svec	↑	05. Oct 2016
	MD-2	Zobraziť 100 najhorších príspevkov za posledných 24-72 hodín (posuvník s časom)	Andrej Svec	↑	04. Oct 2016
	MD-1	Zobraziť 100 najhorších príspevkov za posledných 24 hodín	Andrej Svec	↑	04. Oct 2016

Aktuálny stav úloh

T	Key	Summary	Reporter	P	Created	Assignee	Status
	MD-29	MD-16 / Urobiť web	Andrej Svec	↑	12. Oct 2016	Jakub Adam	DONE
	MD-28	MD-16 / Pripraviť server	Andrej Svec	↑	12. Oct 2016	Filip Vozar	DONE
	MD-26	MD-4 / Zobrazit statistiky komentara v novom okne	Andrej Svec	↑	12. Oct 2016	Monika Filipcikova	DONE
	MD-25	MD-4 / Refaktor javascriptu pre zobrazovanie statistik a grafov	Andrej Svec	↑	12. Oct 2016	Andrej Svec	DONE
	MD-16	Spraviť web na timak	Filip Vozar	↑	11. Oct 2016	Jakub Adam	DONE
	MD-9	Autentifikácia voči portálu pomocou access tokenu uloženého v kúkinie :)	Andrej Svec	↑	05. Oct 2016	Jakub Adam	DONE
	MD-8	Vytvorit v databáze entitu autora pokiaľ neexistuje pri vytváraní jeho prvého príspevku	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	DONE
	MD-7	Vytvorit endpointy pre zápis akcií Approve, Reject, Ban, Report	Andrej Svec	↑	05. Oct 2016	Filip Vozar	DONE
	MD-6	Dokumentácia k základným API volaniam	Andrej Svec	↑	05. Oct 2016	Filip Vozar	DONE
	MD-5	Zobraziť podrobné štatistiky autora na novej karte	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	DONE
	MD-4	Zobraziť podrobné štatistiky komentáru na zvlášť stránke + Refaktor javascriptu	Andrej Svec	↑	05. Oct 2016	Monika Filipcikova	DONE
	MD-3	Mini iframe so zobrazením celkového hodnotenia komentáru a autora	Andrej Svec	↑	05. Oct 2016	Andrej Svec	DONE

Retrospektíva šprintu “ARA”

Dátum: 26.10.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Andrej Švec
Filip Vozár

Čo sa podarilo

- dobre fungujúca komunikácia
- splnenie všetkých úloh v šprinte

Čo sa nepodarilo

- stretávali sme sa len raz v týždni, čo je málo
- neskore schvaľovanie pull requestov
- menší prehľad o stave projektu, keďže dochádzalo vývoju na vetvách, o ktorých neboli nikde záznamy

Čo plánujeme zlepšiť

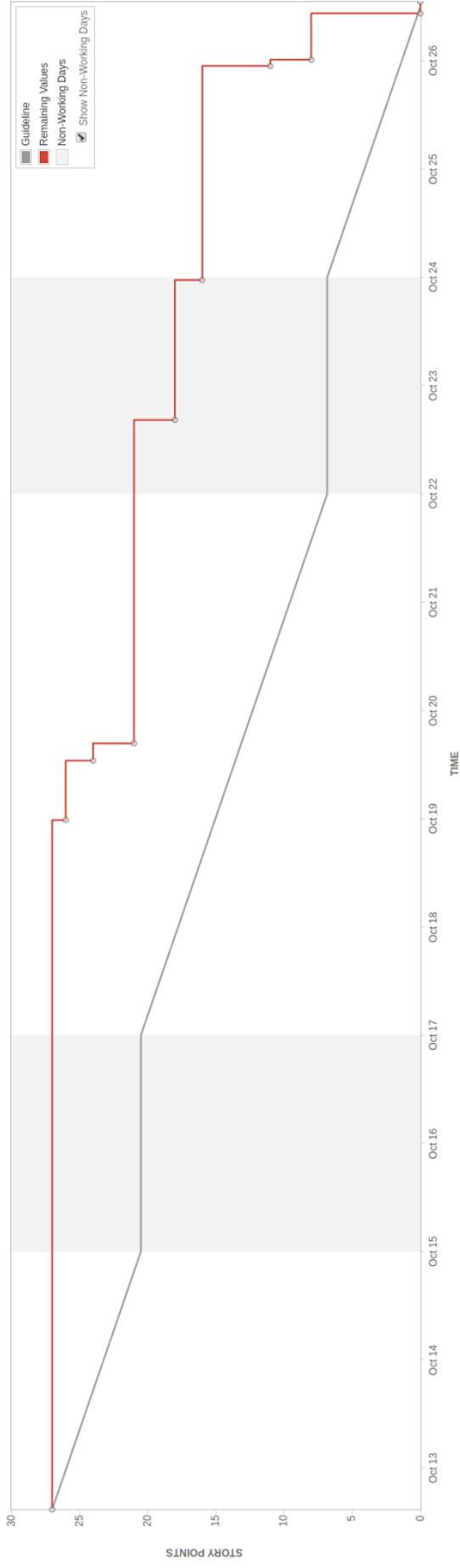
- approve pri pull requeste treba dať hneď ako s tým súhlasíme
- nemergovať do developu ak nastáva nekonzistentný stav medzi API a Portálom
- nevyvíjať na vlastných branch-ach bez toho aby o tom ostatní vedeli
- nepracovať na úlohách, ktoré nie sú v aktuálnom šprinte.
- nepracovať na úlohách, ktoré majú sú naviazané na nedokončených úlohách

Review šprintu “ARA”

Dátum: 26.10.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Andrej Švec
Filip Vozár

Šprint review

- Všetky úlohy sa nám podarilo splniť
- Problematický vývoj, keďže nie je oddelene produkčné prostredie od vývojového
 - Riešenie: Zjednodušenie develop prostredia (autentifikačné tokeny, ENV premenné)
- Nemergovať do developu ak nastáva nekonzistentný stav medzi API a Portálom
- Problematické testovanie
 - Riešenie: Analýza dôvodu, prečo testy nebežia
 - Analýza najvhodnejších testovacích frameworkov



Date	Issue	Event Type	Event Detail	Inc.	Dec.	Remaining
12. Oct 2016 14:53	MD-16 MD-3 MD-4 MD-5 MD-6 MD-7 MD-8 MD-9	Sprint start		3		
18. Oct 2016 23:41	MD-9	Burndown	Issue completed		1	26
19. Oct 2016 12:53	MD-3	Burndown	Issue completed		2	24
19. Oct 2016 16:42	MD-6	Burndown	Issue completed		3	21
22. Oct 2016 16:23	MD-16	Burndown	Issue completed		3	18
23. Oct 2016 23:23	MD-8	Burndown	Issue completed		2	16
25. Oct 2016 22:49	MD-7	Burndown	Issue completed		5	11
26. Oct 2016 0:12	MD-5	Burndown	Issue completed		3	8
26. Oct 2016 10:30	MD-4	Burndown	Issue completed		8	0
26. Oct 2016 13:02	MD-16 MD-3 MD-4 MD-5 MD-6 MD-7 MD-8 MD-9	Sprint ended by Andrej Svec				0

Zápis zo stretnutia #7

Dátum: 2.11.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Jakub Šimko
Andrej Švec
Filip Vozár

Priebeh stretnutia

- Diskusia o priebehu minulých stretnutí
 - Pripomienkovanie "Sprint review" (nový dokument, technické veci)
 - Pripomienky k retrospektíve (aké otázky si klásť)
 - Diskusia k deleniu user stories na subtasky (Deliť len dovedy, kým to ma prínos pre zákazníka)
 - Diskusia k aktuálnemu priebehu (Čo sa spravolo, aké sú problémy?)
- Diskusia k tomu, čo robiť ak sa vyskytne problém, ktorý treba riešiť urgentne.

Úlohy na ďalší týždeň

- Pracovať na úlohách z aktuálneho šprintu.

Aktuálny backlog

T	Key	Summary	Reporter	P	Created
	MD-54	Zistiť stav testov	Andrej Svec	↑	26. Oct 2016
	MD-53	ANALYZA Ako zobrazit zoznam komentarov	Andrej Svec	↑	26. Oct 2016
	MD-52	Vyrobit strankovanie na portali	Andrej Svec	↑	26. Oct 2016
	MD-48	Vytvorit skript na trenovanie	Andrej Svec	↑	26. Oct 2016
	MD-47	ANALYZA Aky framework na testovanie	Andrej Svec	↑	26. Oct 2016
	MD-46	Pravidlovy fuzzy klasifikator	Andrej Svec	↑	26. Oct 2016
	MD-45	Nasadenie v SME	Andrej Svec	↑	26. Oct 2016
	MD-43	Spisat vnutornu konfiguraciu, ako rozchodit projekt	Andrej Svec	↑	26. Oct 2016
	MD-38	Spustit detektory na SME dátach z roku 2016	Andrej Svec	↑	21. Oct 2016
	MD-37	Nastavit zalohovanie databaz a konfiguracii clustra	Filip Vozar	↑	21. Oct 2016
	MD-34	Lepsie handlovenie couchbase (a redis) connection - po timeoutu failuju aj dalsie pokusy, aj ked sa problem s timeoutom vyriesil a couchbase je plne dostupny	Filip Vozar	↑	18. Oct 2016
	MD-32	Odoslat formular "Try our service" Enterom, SHIFT + Enter je newline	Andrej Svec	↑	17. Oct 2016
	MD-31	ANALYZA Co ma vratit off-topic detektor, ked nie je parent	Andrej Svec	↑	17. Oct 2016
	MD-30	Premysliet a implementovat monitoring a logovanie	Filip Vozar	↑	15. Oct 2016
	MD-24	Nahrat SME dáta z roku 2016 do couchbase, aby sme mohli trénovať	Andrej Svec	↑	12. Oct 2016
	MD-22	Pridanie knižnice fuzzy do projektu	Monika Filipcikova	↑	12. Oct 2016
	MD-21	Integracia dvoch modelov fuzzy + klasifikator	Monika Filipcikova	↑	12. Oct 2016
	MD-20	Implementácia pravidiel pre fuzzy	Monika Filipcikova	↑	12. Oct 2016
	MD-19	SLIDO - analiza detektorov	Monika Filipcikova	↑	12. Oct 2016
	MD-18	ANALYZA najdenie fuzzy kniznice, vytvorenie simple prototypu	Monika Filipcikova	↑	12. Oct 2016
	MD-17	Admin rozhranie pre zakaznikov	Filip Vozar	↑	12. Oct 2016
	MD-15	Webhooky na vrátenie vysledkov	Andrej Svec	↑	09. Oct 2016
	MD-14	Filtrovanie a zoradovanie v zozname komentárov	Andrej Svec	↑	09. Oct 2016
	MD-12	Upravit farby na webe	Andrej Svec	↓	09. Oct 2016
	MD-11	Opravit existujúce testy	Andrej Svec	↑	05. Oct 2016
	MD-2	Zobrazit 100 najhorších príspevkov za posledných 24-72 hodin (posuvnik s časom)	Andrej Svec	↑	04. Oct 2016
	MD-1	Zobrazit 100 najhorších príspevkov za posledných 24 hodin	Andrej Svec	↑	04. Oct 2016

Aktuálny stav úloh

T	Key	Summary	Reporter	P	Created	Assignee	Status
	MD-56	Spustit analyzu clanku pri jeho pridani	Andrej Svec	↑	01. Nov 2016	Jakub Adam	TO DO
	MD-54	Zistiť stav testov	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	TO DO
	MD-53	ANALYZA Ako zobrazit zoznam komentarov	Andrej Svec	↑	26. Oct 2016	Jakub Adam	IN PROGRESS
	MD-47	ANALYZA Aky framework na testovanie	Andrej Svec	↑	26. Oct 2016	Filip Vozar	IN PROGRESS
	MD-44	Pripravit nasadenie API a Portal pre SME	Andrej Svec	↑	26. Oct 2016	Filip Vozar	IN PROGRESS
	MD-42	Prepisat approve a reject status ak pride nove volanie	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	TO DO
	MD-41	Vyzadovat ID komentara pri jeho vytvarani	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	IN PROGRESS
	MD-40	Upravit parametre, ktore prijimaju controllery, tak aby sedeli podla dokumentacie	Filip Vozar	↑	25. Oct 2016	Monika Filipcikova	TO DO
	MD-39	Upravit Portal model - vymazat token	Andrej Svec	↑	21. Oct 2016	Jakub Adam	IN PROGRESS
	MD-36	Natrenovat model pre SME, aby nebol pretrenovany	Andrej Svec	↑	21. Oct 2016	Andrej Svec	IN PROGRESS
	MD-35	Zmena tvorby primarnych klucov pre dokumenty v Couchbase, upratanie views	Andrej Svec	↑	19. Oct 2016	Filip Vozar	DONE
	MD-33	Zabezpecit defaultne nastavenia autentifikacie, aby sa jednoduchsie dal rozchodit projekt	Filip Vozar	↑	18. Oct 2016	Andrej Svec	TO DO

Zápis zo stretnutia #8

Dátum: 9.11.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Monika Filipčíková, Filip Vozár
Prítomní: Monika Filipčíková
Jakub Šimko
Andrej Švec
Filip Vozár

Priebeh stretnutia

- Diskusia o dôvodoch nestihnúť všetkých taskov v šprinte
- Každý člen tímu povedal, čo sa mu podarilo spraviť počas šprintu.
- Šprint review
- Diskusia o závislostiach medzi taskami
- Retrospektíva
- Otváranie šprintu
 - Zoradzovanie taskov
 - Ohodnocovanie taskov (Planning poker)
 - Vyberanie taskov do šprintu
- Slávnosť spustenia šprintu "Corvus corax"

Úlohy na ďalší týždeň

- Pracovať na úlohách z aktuálneho šprintu.
- Pripraviť dokumentáciu na kontrolné odovzdanie

Aktuálny backlog

T	Key	Summary	Reporter	P	Created
	MD-54	Zistiť stav testov	Andrej Svec	↑	26. Oct 2016
	MD-53	ANALYZA Ako zobrazit zoznam komentarov	Andrej Svec	↑	26. Oct 2016
	MD-52	Vyrobit strankovanie na portali	Andrej Svec	↑	26. Oct 2016
	MD-48	Vytvorit skript na trenovanie	Andrej Svec	↑	26. Oct 2016
	MD-47	ANALYZA Aky framework na testovanie	Andrej Svec	↑	26. Oct 2016
	MD-46	Pravidlovy fuzzy klasifikator	Andrej Svec	↑	26. Oct 2016
	MD-45	Nasadenie v SME	Andrej Svec	↑	26. Oct 2016
	MD-43	Spisat vnutornu konfiguraciu, ako rozchodit projekt	Andrej Svec	↑	26. Oct 2016
	MD-38	Spustiť detektory na SME dátach z roku 2016	Andrej Svec	↑	21. Oct 2016
	MD-37	Nastavit zalohovanie databaz a konfiguracii clustra	Filip Vozar	↑	21. Oct 2016
	MD-34	Lepsie handlovenie couchbase (a redis) connection - po timeoute failuju aj dalsie pokusy, aj ked sa problem s timeoutom vyriesil a couchbase je plne dostupny	Filip Vozar	↑	18. Oct 2016
	MD-32	Odoslat formular "Try our service" Enterom, SHIFT + Enter je newline	Andrej Svec	↑	17. Oct 2016
	MD-31	ANALYZA Co ma vratit off-topic detektor, ked nie je parent	Andrej Svec	↑	17. Oct 2016
	MD-30	Premysliet a implementovat monitoring a logovanie	Filip Vozar	↑	15. Oct 2016
	MD-24	Nahrať SME dáta z roku 2016 do couchbase, aby sme mohli trénovať	Andrej Svec	↑	12. Oct 2016
	MD-22	Pridanie knižnice fuzzy do projektu	Monika Filipcikova	↑	12. Oct 2016
	MD-21	Integracia dvoch modelov fuzzy + klasifikator	Monika Filipcikova	↑	12. Oct 2016
	MD-20	Implementácia pravidiel pre fuzzy	Monika Filipcikova	↑	12. Oct 2016
	MD-19	SLIDO - analýza detektorov	Monika Filipcikova	↑	12. Oct 2016
	MD-18	ANALYZA najdenie fuzzy kniznice, vytvorenie simple prototypu	Monika Filipcikova	↑	12. Oct 2016
	MD-17	Admin rozhranie pre zakaznikov	Filip Vozar	↑	12. Oct 2016
	MD-15	Webhooky na vrátenie výsledkov	Andrej Svec	↑	09. Oct 2016
	MD-14	Filtrovanie a zoradovanie v zozname komentárov	Andrej Svec	↑	09. Oct 2016
	MD-12	Upraviť farby na webe	Andrej Svec	↓	09. Oct 2016
	MD-11	Opraviť existujúce testy	Andrej Svec	↑	05. Oct 2016
	MD-2	Zobrazit 100 najhorších príspevkov za posledných 24-72 hodín (posuvník s časom)	Andrej Svec	↑	04. Oct 2016
	MD-1	Zobrazit 100 najhorších príspevkov za posledných 24 hodín	Andrej Svec	↑	04. Oct 2016

Aktuálny stav úloh

T	Key	Summary	Reporter	P	Created	Assignee	Status
	MD-56	Spustiť analýzu clanku pri jeho pridani	Andrej Svec	↑	01. Nov 2016	Jakub Adam	DONE
	MD-54	Zistiť stav testov	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	TO DO
	MD-53	ANALYZA Ako zobrazit zoznam komentarov	Andrej Svec	↑	26. Oct 2016	Jakub Adam	TO DO
	MD-47	ANALYZA Aky framework na testovanie	Andrej Svec	↑	26. Oct 2016	Filip Vozar	TO DO
	MD-44	Pripravit nasadenie API a Portal pre SME	Andrej Svec	↑	26. Oct 2016	Filip Vozar	IN PROGRESS
	MD-42	Prepisat approve a reject status ak pride nove volanie	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	DONE
	MD-41	Vyzadovat ID komentara pri jeho vytvarani	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	DONE
	MD-40	Upravit parametre, ktore prijimaju controllery, tak aby sedeli podla dokumentacie	Filip Vozar	↑	25. Oct 2016	Monika Filipcikova	DONE
	MD-39	Upravit Portal model - vymazat token	Andrej Svec	↑	21. Oct 2016	Jakub Adam	DONE
	MD-36	Natrenovat model pre SME, aby nebol pretrenovany	Andrej Svec	↑	21. Oct 2016	Andrej Svec	DONE
	MD-35	Zmena tvorby primarnych klucov pre dokumenty v Couchbase, upratanie views	Andrej Svec	↑	19. Oct 2016	Filip Vozar	DONE
	MD-33	Zabezpecit defaultne nastavenia autentifikacie, aby sa jednoduchšie dal rozchodit projekt	Filip Vozar	↑	18. Oct 2016	Andrej Svec	IN PROGRESS

Retrospektíva šprintu “Bubo bubo”

Dátum: 9.11.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Monika Filipčíková, Filip Vozár
Prítomní: Monika Filipčíková
Jakub Šimko
Andrej Švec
Filip Vozár

Zhrnutie

V tomto šprinte sa nám nepodarilo splniť všetky plánované úlohy. Podarilo sa splniť len 16 storypointov z plánovaných 42. Bolo to spôsobené hlavne zameraním sa členov tímu na iné študijné povinnosti. Ďalším dôvodom bolo zlé zvolenie taskov do šprintu, keďže počas šprintu sa objavili ďalšie úlohy, ktoré mali väčšiu prioritu ako tasky v šprinte. Kvôli tomu sme sa na ne viac zamerali a úlohy s nižšou prioritou, ktoré boli v šprinte plánované od začiatku sa nestihli spraviť. Taktiež nám prácu komplikoval fakt, že viaceré tasky boli medzi sebou závislé a nedokončenie jedného spôsobilo nedokončenie viacerých.

V čom pokračujeme?

- Dobre odhadovanie bodov pre tasky
- Reporty z Bitbucketu a JIRY sú v Slacku, takže sa k nim môžeme hneď vyjadrovať
- Dobre napísané descriptions k úlohám
- Dobré zaznamenávanie progresu v Jire, vďaka čomu máme dobrý prehľad o stave projektu

Čo prestať robiť?

- Nerobiť v dvoch taskoch to isté
- Nerobiť narýchlo kód a pull request, lebo kvalita klesá
- Necommitovať zbytočnosti, aby sa nestrácal prehľad medzi zmenami
- Nepridávať tasky do JIRY po začatí šprintu, radšej si vyhradiť viac času na plánovanie šprintu
- Nevytvárať “jednoriadkové” user stories
- Nezaradovať do šprintu tasky s prílišnými závislosťami medzi sebou

Čo začať robiť?

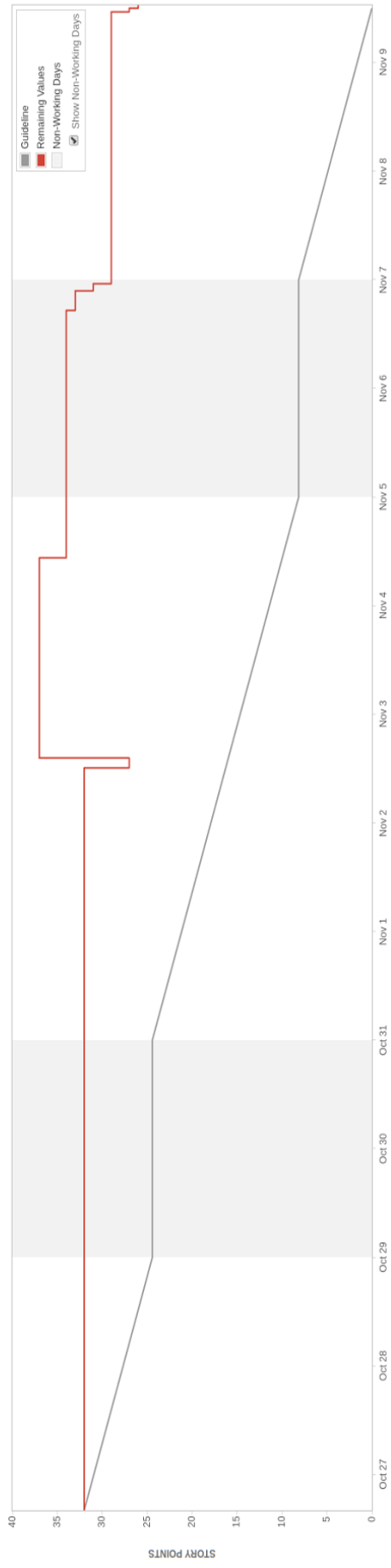
- Zlepšiť spôsob, akým si medzi sebou šírieme zásadné zmeny (napr. v databázovej schéme) - aby iný člen nestrácal čas hľadáním chyby, ktorá neexistuje.
- Zamerať sa na skvalitnenie testov
- Používať možnosť pridávania komentárov k taskom v JIRE pre zlepšenie prehľadu
- Po každotýždňovom stretnutí tímu spraviť export stavu projektu z JIRY
- Oboznámiť viacerých členov tímu s použitou technológiou (napr. Docker), aby sa nestávalo, že niektoré typy taskov budú stáť na jednom členovi.
- Treba si vyjasniť, akú funkcionality má portál poskytovať a neplánovať pridávanie funkcionality, ktorá nie je potrebná
- Hneď ako nás napadne user story, korektne zapísať ju do JIRY

Review Šprintu “Bubo bubo”

Dátum: 9.11.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Monika Filipčíková, Filip Vozár
Prítomní: Monika Filipčíková
Jakub Šimko
Andrej Švec
Filip Vozár

Šprint review

- Nepodarilo sa nám splniť všetky úlohy (16/42)
- Problematický vývoj, keďže nie je oddelene produkčné prostredie od vývojového
 - Riešenie: Zjednodušenie develop prostredia (autentifikačné tokeny, ENV premenné)
- Problematické testovanie
 - Riešenie: zameranie sa na sfunkčnenie testov, umožní nám to vyvíjať rýchlejšie, jednoduchšie refaktorovať, umožní nám to použiť continuous integration...
- Prílišná závislosť medzi taskami
 - Riešenie: Do šprintov sa snažiť vyberať tasky tak, aby od seba neboli príliš závislé, aby neboli jednotlivé tasky blokované.



Completed Issues

[View in Issue Navigator](#)

Key	Summary	Issue Type	Priority	Status	Story Points (16)
MD-35	Zmena tvorby primarnych klucov pre dokumenty v Couchbase, upratanie views	Story	↑ Highest	DONE	5
MD-36	Natrenovat model pre SME, aby nebol pretrateny	Story	↑ Medium	DONE	3
MD-39	Upravit Portal model - vymazat token	Story	↑ Medium	DONE	1
MD-40	Upravit parametre, ktoré prijimaju controllers, tak aby sedeli podľa dokumentacie	Story	↑ Medium	DONE	2
MD-41	Vyzadovat ID komentara pri jeho vytvarani	Story	↑ Medium	DONE	2
MD-42	Prepisat approve a reject status ak pride nove volanie	Story	↑ Medium	DONE	1
MD-56 *	Spustit analyzu clanku pri jeho pridani	Story	↑ High	DONE	2

Issues Not Completed

[View in Issue Navigator](#)

Key	Summary	Issue Type	Priority	Status	Story Points (26)
MD-33	Zabezpecit defaultne nastavenia autentifikacie, aby sa jednoduchšie dal rozchodit projekt	Story	↑ Medium	IN PROGRESS	5
MD-44	Pripavit nasadenie API a Portal pre SME	Story	↑ Medium	IN PROGRESS	5
MD-47	ANALYZA Aky framework na testovanie	Story	↑ Medium	TO DO	3
MD-53	ANALYZA Ako zobrazit zoznam komentarov	Story	↑ Medium	TO DO	3
MD-54	Zisit stav testov	Story	↑ Medium	TO DO	2
MD-57 *	Rozchodit diakritikovac lokalne	Story	↑ Medium	TO DO	8

Zápis zo stretnutia #9

Dátum: 16.11.2016
Miesto: Projektové laboratórium
Vedúci stretnutia: Monika Filipčíková
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková
Jakub Šimko
Andrej Švec
Filip Vozár

Priebeh stretnutia

- Standup o priebehu šprintu
- Diskusia o nasadení dvoch verzií portálu (demo, portál)
- Diskusia o zameraní nasledujúceho šprintu na skvalitnenie testov
- Diskusia o taskoch, ktorým sa treba prioritne venovať
- Diskusia o tom, ako sa ma správať portál, keď pri zavolaní iframu ešte nie sú dostupné výsledky analýz
- Diskusia o nasadení nástroja, ktorý by meral zložitosť systému a pokrytie testami.
- Diskusia o používaných metodikách
- Diskusia k dokumentácii
- Diskusia k prioritne testov v backlogu

Úlohy na ďalší týždeň

- Pracovať na úlohách z aktuálneho šprintu.
- Odovzdať dokumentáciu

Aktuálny backlog

T	Key	Summary	Reporter	P	Created
	MD-70	Testovanie cez aplary.io	Filip Vozar	↑	15. Nov 2016
	MD-69	Vymysliet a implementovať reportovanie chybových hlások v JSON API	Filip Vozar	↑	13. Nov 2016
	MD-68	Opraviť testy pre controllery	Monika Filipcikova	↑	12. Nov 2016
	MD-67	Osetrit iframe a statistiky keď ešte nie sú vypočítané výsledky	Andrej Svec	↑	11. Nov 2016
	MD-66	Preposlať chyby z API do portálu	Andrej Svec	↑	11. Nov 2016
	MD-65	Dokumentácia tímak	Andrej Svec	↑	09. Nov 2016
	MD-64	Odstranit z portálu to, čo už nepodporujeme	Andrej Svec	↑	09. Nov 2016
	MD-63	Upraviť user detektory, aby používali Map-Reduce	Andrej Svec	↑	09. Nov 2016
	MD-54	Zistiť stav testov	Andrej Svec	↑	26. Oct 2016
	MD-53	ANALYZA Ako zobrazit zoznam komentárov	Andrej Svec	↑	26. Oct 2016
	MD-52	Vyrobit strankovanie na portali	Andrej Svec	↑	26. Oct 2016
	MD-48	Vytvorit skript na trenovanie	Andrej Svec	↑	26. Oct 2016
	MD-47	ANALYZA Aky framework na testovanie	Andrej Svec	↑	26. Oct 2016
	MD-46	Pravidlovy fuzzy klasifikator	Andrej Svec	↑	26. Oct 2016
	MD-45	Nasadenie v SME	Andrej Svec	↑	26. Oct 2016
	MD-43	Spisat vnutornu konfiguraciu, ako rozchodit projekt	Andrej Svec	↑	26. Oct 2016
	MD-38	Spustit detektory na SME dátach z roku 2016	Andrej Svec	↑	21. Oct 2016
	MD-37	Nastavit zalohovanie databaz a konfiguracii clustra	Filip Vozar	↑	21. Oct 2016
	MD-34	Lepsie handlovenie couchbase (a redis) connection - po timeoute failuju aj dalsie pokusy, aj keď sa problém s timeoutom vyriesil a couchbase je plne dostupny	Filip Vozar	↑	18. Oct 2016
	MD-32	Odoslať formulár "Try our service" Enterom, SHIFT + Enter je newline	Andrej Svec	↑	17. Oct 2016
	MD-31	ANALYZA Co ma vratit off-topic detektor, keď nie je parent	Andrej Svec	↑	17. Oct 2016
	MD-30	Premysliet a implementovat monitoring a logovanie	Filip Vozar	↑	15. Oct 2016
	MD-24	Nahrat SME dáta z roku 2016 do couchbase, aby sme mohli trénovať	Andrej Svec	↑	12. Oct 2016
	MD-22	Pridanie knižnice fuzzy do projektu	Monika Filipcikova	↑	12. Oct 2016
	MD-21	Integracia dvoch modelov fuzzy + klasifikator	Monika Filipcikova	↑	12. Oct 2016
	MD-20	Implementácia pravidiel pre fuzzy	Monika Filipcikova	↑	12. Oct 2016
	MD-19	SLIDO - analýza detektorov	Monika Filipcikova	↑	12. Oct 2016
	MD-18	ANALYZA najdenie fuzzy knižnice, vytvorenie simple prototypu	Monika Filipcikova	↑	12. Oct 2016
	MD-17	Admin rozhranie pre zakaznikov	Filip Vozar	↑	12. Oct 2016
	MD-15	Webhooky na vrátenie výsledkov	Andrej Svec	↑	09. Oct 2016
	MD-14	Filtrovanie a zoradovanie v zozname komentárov	Andrej Svec	↑	09. Oct 2016
	MD-12	Upraviť farby na webe	Andrej Svec	↓	09. Oct 2016
	MD-11	Opraviť existujúce testy	Andrej Svec	↑	05. Oct 2016
	MD-2	Zobrazit 100 najhorších príspevkov za posledných 24-72 hodín (posuvník s časom)	Andrej Svec	↑	04. Oct 2016
	MD-1	Zobrazit 100 najhorších príspevkov za posledných 24 hodín	Andrej Svec	↑	04. Oct 2016

Aktuálny stav úloh

T	Key	Summary	Reporter	P	Created	Assignee	Status
	MD-65	Dokumentácia tímak	Andrej Svec	↑	09. Nov 2016	Jakub Adam	IN PROGRESS
	MD-62	MD-33 / Vytvorit rake task na seed demo portálu pre ucely vyvoja	Andrej Svec	↑	07. Nov 2016	Andrej Svec	IN PROGRESS
	MD-61	MD-33 / Defaultne nastavenia v yaml	Andrej Svec	↑	07. Nov 2016	Andrej Svec	DONE
	MD-60	MD-57 / Vytvorit Docker image diakritikovaca	Andrej Svec	↑	02. Nov 2016	Filip Vozar	DONE
	MD-59	MD-57 / Integrovat diakritikovac s API	Andrej Svec	↑	02. Nov 2016	Jakub Adam	DONE
	MD-58	MD-57 / Zistiť, ako rozchodit lokálne	Andrej Svec	↑	02. Nov 2016	Jakub Adam	DONE
	MD-57	Rozchodit diakritikovac lokálne	Andrej Svec	↑	02. Nov 2016	Filip Vozar	DONE
	MD-55	Osetrit pripad keď sa komentar uspesne vytvori, ale zlyha schedulovanie do Redisu	Filip Vozar	↑	31. Oct 2016	Filip Vozar	DONE
	MD-54	Zistiť stav testov	Andrej Svec	↑	26. Oct 2016	Monika Filipcikova	IN PROGRESS
	MD-49	Prerobit Sideiq tak aby používal Redis Sentinel	Andrej Svec	↑	26. Oct 2016	Filip Vozar	DONE
	MD-47	ANALYZA Aky framework na testovanie	Andrej Svec	↑	26. Oct 2016	Filip Vozar	IN PROGRESS
	MD-44	Pripravit nasadenie API a Portal pre SME	Andrej Svec	↑	26. Oct 2016	Filip Vozar	IN REVIEW
	MD-33	Zabezpecit defaultne nastavenia autentifikacie, aby sa jednoduchsie dal rozchodit projekt	Filip Vozar	↑	18. Oct 2016	Andrej Svec	TO DO

Používané metodiky

Metodika pre tvorbu dokumentácie

Za dodržiavanie a aktualizovanie tejto metodiky je zodpovedný manažér dokumentácie. Jeho úlohou je vypracovanie všetkých dokumentov zo stretnutí, medzi ktoré patria zápisnice zo stretnutí, retrospektívy a review jednotlivých šprintov. K zodpovednostiam manažéra dokumentácie patrí aj udržiavanie a aktualizácia webovej stránky projektu, kde je povinný pravidelne pridávať nové dokumenty, ktoré vznikli počas práce na projekte.

Ďalej budú opísané metodiky pre tvorbu jednotlivých typov dokumentov.

Zápisnice z tímových stretnutí

Zápisnicu vytvára manažér dokumentácie už počas stretnutia, kedy zapisuje priebeh stretnutia a najdôležitejšie informácie. Po skončení stretnutia sa tieto zápisky upravujú do vopred dohodnutej formálnej podoby, v akej bude zverejnená na stránkach projektu a finálnej dokumentácii. Každý člen tímu má právo zobrazit' a pripomienkovať / upraviť vytvorenú zápisnicu, aby bola dosiahnutá najvyššia možná kvalita.

Formálna zápisnica zo stretnutia sa skladá zo 4 častí:

1. Nadpis vo forme "Zápis zo stretnutia #X", kde znak X je nahradený poradovým číslom stretnutia, ku ktorému zápisnica patrí.
2. Hlavička dokumentu - obsahuje základné informácie o stretnutí
 - a. Dátum konania stretnutia
 - b. Miesto stretnutia
 - c. Vedúci stretnutia - člen tímu zodpovedný za priebeh stretnutia
 - d. Zapisovateľ - člen tímu poverený zápisom priebehu stretnutia
 - e. Prítomní - Menný zoznam členov tímu, ktorý boli prítomní na stretnutí
3. Priebeh stretnutia - niekoľko úrovňový zoznam odrážok, dokumentujúci ako prebiehalo dané stretnutie
4. Nové úlohy - zoznam odrážok, v ktorých sú zapísané úlohy, na ktoré sa má tím zamerať do nasledujúceho stretnutia

Zápis retrospektívy

Tento zápis vytvára manažér dokumentácie v priebehu stretnutia, počas ktorého sa uzatvára aktuálny šprint. V tomto dokumente sú uvedené informácie a zhodnotenie priebehu uzatváraného šprintu. Dokument je viac zameraný na manažérsky pohľad na priebeh šprintu. Tím v tomto dokumente zhodnotí, čo počas skončeného šprintu robil dobre, s čím nebol dostatočne spokojný alebo čo nefungovalo a taktiež čo plánuje zlepšiť alebo začať robiť v nasledujúcom šprinte.

Po skončení stretnutia sa tieto zápisky upravujú do vopred dohodnutej formálnej podoby, v akej bude zverejnená na stránkach projektu a finálnej dokumentácii. Každý člen tímu má právo zobrazit' a pripomienkovať / upraviť vytvorený dokument, aby bola dosiahnutá najvyššia možná kvalita.

Formálny dokument s retrospektívou sa skladá zo 5 častí:

1. Nadpis vo forme "Retrospektíva šprintu "XXX"", kde reťazec XXX je nahradený názvom šprintu, ku ktorému sa vzťahuje.
2. Hlavička dokumentu - obsahuje základné informácie o stretnutí
 - a. Dátum konania stretnutia
 - b. Miesto stretnutia
 - c. Vedúci stretnutia - člen tímu zodpovedný za priebeh stretnutia
 - d. Zapisovateľ - člen tímu poverený zápisom priebehu stretnutia
 - e. Prítomní - Menný zoznam členov tímu, ktorý boli prítomný na stretnutí
3. Čo robíme dobre? - niekoľko úrovňový zoznam odrážok ktorý opisuje, ktoré činnosti sme ako tím počas šprintu vykonávali dobre a mali by sme v tom pokračovať.
4. Čo robíme zle? - niekoľko úrovňový zoznam odrážok ktorý opisuje, ktoré činnosti sme ako tím počas šprintu robili zle, a je potrebné ich riešiť.
5. Čo zlepšiť? - niekoľko úrovňový zoznam odrážok, v ktorých sú zhrnuté kroky, ktoré plánujeme ako tím spraviť pre zlepšenie fungovania tímu.

Zápis sprint review

Tento zápis vytvára manažér dokumentácie v priebehu stretnutia, počas ktorého sa uzatvára aktuálny šprint. V tomto dokumente sú uvedené informácie a zhodnotenie priebehu uzatváraného šprintu. Dokument je viac zameraný na technické problémy, ktoré sa objavili počas šprintu. Sú tu zhrnuté problémy a ich prípadné riešenia

Po skončení stretnutia sa tieto zápisky upravujú do vopred dohodnutej formálnej podoby, v akej bude zverejnená na stránkach projektu a finálnej dokumentácii. Každý člen tímu má právo zobrazovať a pripomienkovať / upraviť vytvorený dokument, aby bola dosiahnutá najvyššia možná kvalita.

Formálna zápisnica zo stretnutia sa skladá zo 3 častí:

6. Nadpis vo forme "Review šprintu "XXX"", kde reťazec XXX je nahradený názvom šprintu, ku ktorému sa vzťahuje.
7. Hlavička dokumentu - obsahuje základné informácie o stretnutí
 - a. Dátum konania stretnutia
 - b. Miesto stretnutia
 - c. Vedúci stretnutia - člen tímu zodpovedný za priebeh stretnutia
 - d. Zapisovateľ - člen tímu poverený zápisom priebehu stretnutia
 - e. Prítomní - Menný zoznam členov tímu, ktorý boli prítomný na stretnutí
8. Šprint review - niekoľko-úrovňový zoznam odrážok, v ktorom sú zhrnuté technické problémy a možné riešenia, ktoré sa objavili počas šprintu.

Web stránka projektu

Aktuálny stav projektu je priebežne aktualizovaný aj na webstránke projektu <http://timak.moderateit.tech>. Po každom tímovom stretnutí je manažér dokumentácie zodpovedný na aktualizovanie informácií na stránke projektu a doplnenie dokumentov, ktoré na stretnutí vznikli.

Prístup repozitáru s webstránkou má každý člen tímu, aby v prípade výpadku manažéra dokumentácie nedošlo k neaktuálnosti stránky.

Dokumentácia k moderateIT API

Keďže hlavnou časťou nášho projektu je REST API, ktoré obsluhuje viacero rôznych volaní, vedíme aj osobitnú dokumentáciu určenú na tieto účely.

Táto dokumentácia je dostupná na adrese <http://docs.moderateitapi.apiary.io> . Ku každému requestu je napísaný stručný popis s informáciou načo dané volanie slúži. Je uvedená adresa na ktorú sa daný request robí, čo musí obsahovať hlavička requestu, atribúty obsiahnuté v tele requestu... Ku každému volaniu sú uvedené aj príkladové requesty.

Každý člen tímu, ktorý pracuje na úlohe, pri ktorej je potrebné meniť niektoré z volaní, je zodpovedný za aktualizovanie tejto dokumentácie a oboznámenie zákazníka ostatných členov tímu o tejto zmene.

Jira exporthy

Exporthy z nástroja na manažment úloh sú rovnako súčasťou dokumentácie a dokumentujú progres tímu na týždňovej báze. Sú zverejnené na webstránke projektu v časti "Jira exporthy" a dostupné širokej verejnosti.

Ich obsahom je tabuľka s aktuálnym stavom backlogu tabuľka s aktuálnym stavom úloh z aktuálne bežiaceho šprintu.

Globálna retrospektíva ZS/LS

Čo robíme dobre?

Komunikácia v tíme. Na komunikáciu v tíme používame Slack, do ktorého máme zavedený aj reporting udalostí. Do spoločných kanálov píšeme vždy keď máme nejaký problém a spravidla sa vieme rýchlo dostať spoločnými silami k riešeniu.

Spoločné stretnutia. Okrem stretnutí, ktoré sú z hľadiska scrumu potrebné, sa stretávame len preto, aby sme spolu boli a riešili úlohy. Takýmto spôsobom si vieme oveľa rýchlejšie poradiť a dopracovať sa k riešeniu. Tieto programovacie stretnutia mávame raz do týždňa.

Opis úloh. V systéme JIRA sú takmer všetky úlohy detailne opísané a keď sa niekto do nejakej úlohy pustí, zvyčajne mu je z opisu úlohy jasné, čo treba spraviť, na čo si dať pozor, popr. na čo nezabudnúť pri testovaní. Písanie opisov nám síce berie viac času, ako by sa mohlo stať užitočným, ale myslíme si, že v skutočnosti to tak nie je.

Čo robíme zle?

Chaotické vyvíjanie. Aj keď už manažujeme úlohy v systéme JIRA, stane sa nám, že niekedy nepracujeme na úlohách v JIRE, ale vyvíjame ich na vlastných lokálnych vetvách, o ktorých nikto nič nevie. Potom sa zrazu zjaví pull request, v ktorom ani nevieme, čo sa spravilo, lebo nie je priradený ku žiadnej úlohe a teda nemá napísané, čo sa v ňom má spraviť.

Rovnako máme niekedy problém s tým, že robíme na úlohách mimo šprintu. Tieto by sme mali normálne pridať do šprintu a nie potom naceňovať

Duplicitné úlohy. Duplicitné úlohy sú bežným scenárom a nie je na tom nič zlé. Je normálne, že sa stane, že dvaja ľudia nahlásia to isté. Ide však skôr o spôsob akým s týmito úlohami jednáme. Stalo sa nám, že sa nám do šprintu dostali úlohy, pričom jedna bola obsiahnutá v druhej. A teda keď prišlo na tú vnorenú úlohu, už bola spravená.

Nejednotné pomenovania. Pri vytváraní pull requestov máme nejednotné názvy. Je to spojené aj s tým, že náš git server, Bitbucket, vytvára pull requestom rôzne názvy podľa toho, koľko commitov je v pull requeste. Rovnako máme problém s tým, že niekedy pomenovávame nejednotne premenné v kóde. Povedali sme si síce, ako budeme premenné pomenovávať a je to v súlade s odporúčaniami, ale nie vždy sú tieto konvencie dodržané.

Návrhy na zlepšenie

Spísať metodiku pre pomenovávanie. Pre odstránenie problémov s nejednotnými pomenovaniami by bolo dobré vytvoriť metodiky pre pomenovávanie či už pull requestov, alebo aj premenných, názvov metód a tried v kóde. Niektoré z týchto vecí sme si síce povedali ústnou formou, ale zjavne to nie je postačujúce.

Komentáre k úlohám. Snažiť sa písať komentáre k úlohám. Tieto komentáre pomáhajú ostatným členom tímu vidieť, na čom sa v rámci úlohy pracovalo. Takisto poukazujú na aktuálny stav úlohy, pokiaľ je In progress. Komentáre pomáhajú vývojárovi zastaviť sa, urobiť krok dozadu a pozrieť sa na to, čo robí, či ide správnym smerom.

Vyriešiť duplicity. Pokiaľ v systéme vzniknú duplicity, je normálne označiť úlohu ako duplicitu s inou úlohou a takto sa nám do šprintu nedostanú úlohy, ktoré sú rovnaké.

PREBERACÍ PROTOKOL

Odozdávajúci subjekt: Tím moderatIT

Preberajúci subjekt : Ing. Jakub Šimko, PhD.

Tento protokol slúži ako potvrdenie o odovzdaní dokumentácie k inžinierskemu dielu a dokumentácie k riadeniu tímového projektu moderatIT.

Poznámka :

Preberajúca strana potvrdzuje svojím podpisom odovzdanie dvoch dokumentov vo vytlačenej forme:

- Dokumentácia k inžinierskemu dielu : listov
- Dokumentácia k riadeniu: listov

.....
podpis zástupcu odovzdávajúcej strany

.....
podpis zástupcu preberajúcej strany

V , dňa

Príloha A: Motivačný dokument

V tejto časti je motivačný dokument v takej podobe, ako bol odovzdaný na začiatku semestra.

Tím

Náš tím tvoria 4 študenti, ktorí práve ukončili bakalárske štúdium na FIIT. Počas predchádzajúceho roku sme pod vedením Ing. Jakuba Šimka, PhD. spoločne pracovali na projekte moderateIT a radi by sme v našom úsilí pokračovali aj na predmete Tímový projekt.

Medzi skúsenosti členov tímu, ktoré by mali byť prínosom pri práci na tímovom projekte, patria najmä skúsenosti so systémovou administráciou serverov a rôznych typov databáz, ktoré nadobudol hlavne Filip počas svojej práce, ako systémový administrátor a backend developer. Filip má taktiež skúsenosti so správou cloudovej infraštruktúry v prostredí Microsoft Azure, ktoré nadobudol počas doterajšej práce na projekte moderateIT.

V osobe Andreja má tím skúsenosti s tvorbou mobilných aplikácií pre OS android, keďže pracuje ako Android developer. Tiež má prehľad a skúsenosti s problematikou strojového učenia, ktorá je najdôležitejšou časťou nášho projektu. V spojitosti so strojovým učením sme tiež získali skúsenosti so základnými prístupmi k počítačovému spracovaniu textu.

Všetci členovia tímu majú tiež skúsenosti s tvorbou webových aplikácií využitím moderných technológií. Majú tiež skúsenosti s prácou na väčších projektoch, ktoré nadobudli v zamestnaní a tiež počas predchádzajúcej práce na projekte. Jakub a Monika majú tiež skúsenosti s testovaním webových aplikácií, ktoré nadobudli hlavne počas predchádzajúcich zamestnaní.

Taktiež keďže spolu na projekte pracujeme už viac ako rok, získali sme základné zručnosti s manažmentom práce v tíme a tiež skúsenosti s prezentovaním projektu potencionálnym zákazníkom z domáceho aj zahraničného trhu na akciách rôzneho druhu (Imagine Cup, Tech Inno Day, Student Startup CUP, MSFEST).

K predmetom Ing. štúdia, ktoré majú členovia tímu aktuálne navolené a mali by byť prínosom k práci na projekte, patria hlavne Architektúra informačných systémov, Pokročilé databázové technológie a Bezpečnosť systémov informačných a komunikačných technológií.

Motivácia

Naša téma sa týka problematiky nevhodných komentárov na webových portáloch. Čoraz viac sa touto oblasťou zaoberajú nielen naše slovenské médiá, ale aj zahraničné. Je to spôsobené tým, že tento problém so sebou prináša nepríjemné prostredie v diskusiách a diskusné portály sú právne zodpovedné za obsah diskusií. Trpia tým portály a zároveň aj samotní navštevníci, ktorí sa chcú dozvedieť zaujímavosti z domova a zo sveta, ale nie je im to v dostatočnej miere umožnené, a preto je potrebný softvér, ktorý by dokázal uľahčiť prácu moderátorom.

Tento nápad vznikol v rámci medzinárodnej súťaže Imagine Cup organizovanej Microsoftom. Následne, potom ako sme zostavili tím, sme boli vedení k vytvoreniu myšlienky, ktorá by mala v sebe biznis potenciál a dokázala by byť do veľkej miery inovatívna. Zisťovali sme ako vieme pomôcť ľuďom zefektívniť ich život. Po dlhom čase

skúmania dier na trhu a identifikovania najrôznejších problémov v spoločnosti sme sa rozhodli zamerať sa na moderovanie diskusií. Bolo potrebné overiť si ako veľmi je táto téma aktuálna a či je to skutočným problémom, preto sme kontaktovali čo najviac ľudí z danej oblasti či už moderátorov, editorov, ba dokonca aj sociológa, ktorý vedie dlhoročný výskum v oblasti diskusií na webe a ich moderovania. Naše domnieky sa potvrdili. Moderátori boli nadchnutí a zároveň nám pomohli hlbšie sa zorientovať v problematike a uviesť nám svoje požiadavky. Dali sme si na tom záležať, aby sme nevyvíjali niečo, o čo následne nikto nebude mať záujem.

Medzitým sme však pracovali aj na vstupoch potrebných do medzinárodnej súťaže Imagine Cup. Podarilo sa nám získať prvé miesto na celom svete za nápad a biznis model, čo bolo pre nás veľmi pozitívne znamenie, že postupujeme správnym smerom.

Ďalším dôležitým bodom bol pre nás postup do semifinále Imagine Cupu. Mediá zareagovali okamžite a podarilo sa nám vzbudiť záujem u verejnosti a taktiež poukázať na významnosť našej fakulty, ktorá sa do veľkej miery podieľa na tom, aby z jej študentov vyrastali cieľavedomí ľudia.

Plán na nasledujúce obdobie

Náš projekt chceme uviesť na trh a získať prvých zákazníkov a čo je pre tento predmet dôležitejšie, získať prvých používateľov. Ide najmä o novinový a diskusný portál SME.sk, s ktorými sa dohadujeme na testovacej prevádzke. Vďaka tomuto dostaneme podnety, priamo od našich zákazníkov a používateľov. Toto veľké množstvo spätnej väzby budeme chcieť samozrejme využívať a zapracovať do našej služby. Z tohto dôvodu plánujeme tiež osloviť ďalšie portály a inštitúcie, či už na Slovensku alebo aj v zahraničí.

Naša práca sa teda bude do istej časti odvíjať od ich ďalších požiadaviek. Momentálne máme za sebou stretnutie s moderátormi v SME.sk a človekom zodpovedným za diskusie, Matejom Čorejom. Ukázali sme im, čo sa nám podarilo a ako naša služba vyzerá teraz. Tí boli nadšení z toho, že im vieme ponúknuť niečo, čo im v moderovaní pomôže a začali sme sa rozprávať o nasadení a konkrétnych technických detailoch.

Zatiaľ sme sa dohodli s ich CTO na tom, že by radšej ako naše používali svoje moderačné rozhranie. Preto sme sa rozhodli, že naše služby im poskytneme prostredníctvom iframe, ktorý si budú integrovať do svojho rozhrania. Ukázali nám aj presné miesto, kde by chceli iframe mať. Ide o pomerne málo vysoký priestor, ktorý sa nachádza v okienku pod príspevkom, ktorého sa týka. Našou úlohou bude teda navrhnuť a vytvoriť iframe, ktorý sa zmestí do tohto malého priestoru. V interných kruhoch sme sa zatiaľ rozprávali o tom, že do tohto priestoru by sme im umiestnili graficky pekne a prehľadne spracovanú informáciu o celkovej kvalite daného príspevku našou službou.

Po kliknutí na mini ukazovateľ celkovej kvality príspevku by sme radi otvorili vyskakovacie okno, v ktorom by sa nachádzali podrobné informácie, ktoré sme zistili analýzou príspevku. Ide najmä o počet linkov, počet nadávok, informácia o bohatosti slovnej zásoby diskutéra, náladu príspevku, gramatickú správnosť, ale aj o prekryv témy komentára s témou článku na ktorý reaguje a príspevkom, na ktorý priamo reaguje, ak takýto príspevok existuje. Toto vyskakovacie okno chceme poskytnúť pre našich zákazníkov tiež formou iframe, avšak jeho

technická realizácia bude omnoho náročnejšia ako realizácia prvého iframe. Naši zákazníci si ho totiž budú musieť vložiť do vyskakovacieho okna sami a rovnako toto okno obsahuje značné množstvo javascriptom generovaných častí. Ide najmä o grafy, ktoré sprehľadňujú celú aplikáciu.

V moderačnom portáli SME.sk sme si všimli aj viacero možností filtrovania a zoradovania komentárov, napríklad filtrovanie podľa článku, pod ktorým sú napísané. Naším cieľom je ukázať, ako dobre naša služba funguje a aký benefit má táto služba pre moderátorov. Chceme teda poskytnúť našim používateľom možnosť pozrieť sa na 100 podľa našej služby najhorších komentárov za daný deň. Na tomto mieste by moderátori mali objaviť komentáre, ktoré im “ušli pomedzi prsty” a našli ich len vďaka našej službe. Týmto spôsobom si uvedomia, aká je naša služba užitočná. Táto funkcionálnosť je pre nás kľúčová avšak zatiaľ ju nemáme vytvorenú. Bude preto nutné dohodnúť sa so zákazníkmi akým spôsobom im doručíme 100 najhorších príspevkov. Jeden z nápadov, ktorý padol, bolo volanie zo serveru diskusného portálu na náš server. Týmto volaním by sa vypýtal zoznam identifikátorov, ktoré zodpovedajú najhorším príspevkom. Následne by sa tieto príspevky vyhľadali na serveri diskusného portálu a boli by zobrazené v zozname. Avšak aj túto možnosť musíme ešte prekonzultovať s predstaviteľmi diskusného portálu.

Popri práci na integrácii našich výsledkov do už existujúcich moderačných rozhraní chceme pokračovať aj vo vývoji nášho používateľského rozhrania (portálu), ktoré chceme optimalizovať pre efektivitu moderovania. Naším cieľom je teda spraviť rozhranie v ktorom možno čo najefektívnejšie označovať príspevky za vhodné, alebo nevhodné do diskusie. Všetky rozhodnutia chceme zjednodušovať pomocou analýz z našej služby a takto moderátorom poskytovať nástroj pre podporu rozhodovania. Vyššie sme písali, že chceme získať spätnú väzbu od moderátorov. Avšak túto spätnú väzbu vieme dostať aj inak ako tak, že nám moderátori budú spísovať pripomienky a návrhy. Požiadavky vieme získavať aj automatizovane, s použitím trackovania. V našom portáli chceme teda zapracovať trackovanie mnohých udalostí. Neskôr, keď už naše rozhranie bude používať značné množstvo používateľov a bude vykonaných veľa udalostí, sa pozrieme na dáta a budeme vedieť povedať, ktorú akciu vykonávajú moderátori najčastejšie a práve túto akciu im budeme vedieť sprístupniť tak, aby sa k nej rýchlejšie dostali. Buď použijeme klávesové skratky, alebo len vytiahneme akciu von z kontextového menu, aby bolo treba na jej vykonanie menej klikov.

Naše moderačné rozhranie (portál) je taktiež momentálne len prototypom a teda nie je pripravený na úplne reálnu prevádzku. Sú v ňom síce zobrazené komentáre, ale portál funguje svižne iba pre málo komentárov. Toto však nie je prípad diskusných portálov. Na niektorých slovenských denníkoch pribudnú denne tisíce komentárov. Niektoré články obsahujú stovky komentárov a našou úlohou bude zachovať aj pri takýchto množstvách komentárov prehľadnosť a svižnosť portálu. Potrebné bude najmä zavedenie stránkovania, alebo tzv. nekonečného skrollovania s postupným načítavaním komentárov do zoznamu. Nad takýmto množstvom príspevkov budeme musieť umožniť aj efektívne filtrovať a zoradovať. Niektoré akcie a možnosti filtrovania a zoradovania, ktoré by radi uvítali

moderátori sme konzultovali aj s moderátorom z Denníku N. Išlo o filtrovanie podľa článku, autora článku, podľa viacerých článkov a zoraďovanie podľa času alebo kvality komentára odhadnutej našou službou.

Sledovanie správania sa používateľa na diskusiách je významnou vecou, ktorá nás môže posunúť ďalej. Portály nemajú do dostatočnej miery prepracované štatistiky o používateľoch, ktorí navštevujú chodia na diskusie. Veľmi zaujímavou úlohou bude naprogramovať skript, ktorý dokáže zaznamenávať akcie používateľa. Napríklad aké články si otvára a ku ktorým z nich pridáva komentáre. Na základe toho, bude možné pridať do nášho klasifikátora ďalšie črty. Ak budeme sledovať kroky jednotlivých navštevujúcich stránok z dlhodobého hľadiska, dosiahneme tým ďalšie vstupy, ktoré skvalitnia našu analýzu a zároveň dokážeme portálom ukázať bližšie informácie o ich čitateľoch aj v našom grafickom rozhraní.

Naša služba je založená na klasifikátore, ktorému sú predložené črty a ktorý sa na základe už ohodnotených dát učí. Do toho by sme však radi zapracovali fuzzy vyhodnocovanie, kde by sme si dokázali určiť, pri akej hodnote, ktorej črty môžeme považovať komentár za zlý alebo naopak dobrý. Mohli by sme sa riadiť kódexmi jednotlivých denníkov, ale vyžaduje si to aj ďalšie konzultácie s denníkmi, ktoré majú o týchto pravidlách lepší prehľad už priamo z praxe.

Veľkú prioritu má aj implementácia vrátenia výsledkov analýz pomocou webhookov. Toto bude uprednostňovaný spôsob vrátenia výsledkov kvôli jednoduchšiemu škálovaniu a nižšej záťaži na naše servery v porovnaní s dopytovaním sa na výsledky. Pri konzultácii so sli.do ohľadom prípadnej integrácie sme sa rozprávali o možnostiach použitia technológie websocketov na medziserverovú komunikáciu. Ak by aj spolupráca so sli.do nevyšla, tak websockety budeme naďalej vyhodnocovať ako jednu z reálnych možností na prepojenie serverov diskusných portálov s našimi API servermi.

Okrem zlepšovania výsledkov analýz sa potrebujeme sústrediť aj zlepšovanie a udržiavanie infraštruktúry. Cez letné prázdniny sa nám podarilo implementovať základy mechanizmu autentifikácie a autorizácie prístupu k API s možnosťou obmedzenia na čítanie výsledkov, vkladanie nových záznamov a administráciu, založeného na Json Web Tokenoch. Tento mechanizmus bude potrebné dôkladne otestovať a v prípade potreby implementovať dodatočnú funkcionálnosť ktorá vyplynie z požiadaviek zákazníkov pri testovaní.

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

Jakub Adam, Monika Filipčíková, Andrej Švec, Filip Vozár

moderateIT

Dokumentácia k inžinierskemu dielu z predmetu Tímový projekt

Študijný program: Informačné systémy / Softvérové inžinierstvo

Vedúci tímu: Ing. Jakub Šimko, PhD.

Zimný semester šk. r. 2016 / 2017

Dokumentácia k inžinierskemu dielu

Big picture

Úvod

S pribúdajúcim počtom používateľov webu, rastie každodenne aj počet komentárov, v rôznych diskusiách. Novinové portály majú veľký záujem o to, aby sa na ich stránkach vo veľkom diskutovalo. Avšak čím ďalej, tým viac je web zaplavovaný negativitou, nenávisťnými komentármi a tzv. "trollmi", ktorí sa snažia diskusie znehodnocovať a napádať ostatných návštevníkov.

Portály sú za diskusie zodpovedné z právneho hľadiska a taktiež je v ich záujme, aby sa návštevníci cítili na ich stránkach príjemne. Často však nie je v silách moderátorov novinových portálov odstrániť všetky komentáre nepatriace do diskusie. Práve preto sme sa rozhodli, že vytvoríme nástroj, ktorý bude slúžiť, ako poradný hlas pre moderátorov. Jeho účelom je ušetriť im čas a zároveň urýchliť ich prácu.

Služba, ktorú sme vytvorili je založená na strojovom učení. Z komentára sú abstrahované črty opisujúce komentár, autora a článok. Tieto črty sú následne spracovávané pomocou rozhodovacích stromov. Hodnota, ktorú získame zodpovedá kvalite komentára a slúži nám na zoraďovanie komentárov podľa kvality. Moderátor sa tak môže zamerať primárne na toxické komentáre a nemusí čítať úplne všetko.

Globálne ciele pre ZS/LS

Nasadenie do SME.sk

Naším prvým a najdôležitejším cieľom, ktorý je zároveň aj trošku biznis typu, je nasadenie našej služby pre nášho partnera SME.sk. S týmto je spojené množstvo práce inžinierskeho typu. V prvom kroku sme so SME.sk dohodli, ako budeme na ich systém napojení.

API. Dohodli sme, že nám budú posielat' do nášho API vystaveného na URL <https://api.moderateit.tech> jednotlivé články, autorov článkov a komentáre. My budeme komentáre analyzovať a výsledky si ukladať do databázy. Nebudeme ich SME.sk nijako vracat', budú si ich musieť sami vypýtať.

iframe. Našou úlohou je navrhnúť a vytvoriť iframe, ktorý si SME.sk vloží na svoju stránku a vďaka tomu ich moderátori uvidia vo svojom moderačnom rozhraní výsledky našich analýz. Do iframe umiestnime graficky pekne a prehľadne spracovanú informáciu o celkovej kvalite daného príspevku ohodnotenej našou službou a o reputácii autora, tiež ohodnotenej našou službou.

Podrobné štatistiky. Po kliknutí do iframe sa v novom okne otvoria podrobné štatistiky komentára, alebo autora. Tieto štatistiky už máme pripravené, akurát SME.sk ich potrebuje v novom okne a nie vy vyskakovacím okne, čiže ich bude treba trošku upraviť

Model. Pre hodnotenie príspevkov SME.sk je potrebné na ich dátach natrénovať aj nový model. Náš model beží v prostredí Microsoft Azure Machine Learning Studio. Trénovanie v praxi predstavuje zobrať nový dataset a použiť hyperparametre modelu, ktoré už máme prednastavené a pohrať sa s nimi, aby sme dosiahli čo najlepšie výsledky.

Ďalšie dáta. Od SME.sk sme dostali aj ďalšie dáta, konkrétne všetky dáta za rok 2016. Tieto dáta by sme chceli analyzovať a potom trénovať model znova. Získaním väčšieho množstva tréningových dát by sme sa mohli pokúsiť natrénovať modely, ktoré vyžadujú viac dát na tréning, napr. nejaký hĺbkovo sa učiaci model.

Optimalizácie databázy. V projekte nepoužívame relačné databázy. Všetky perzistentné dáta sú uložené v databáze Couchbase. Avšak spôsob akým využívame Couchbase z neho môže časom spraviť úzke hrdlo systému. Navyše tento spôsob nie je v súlade s odporúčaným používaním Couchbase. Ide konkrétne o to, že často využívame views aj na vyhľadanie nejakých konkrétnych záznamov, čo by sme nemali. Views by sme mali používať len pri hľadaní množstva príspevkov, ktoré spĺňajú nejaké požiadavky.

Preto budeme potrebovať zmeniť spôsob, akým Couchbase využívame. Pôjde najmä o systém generovania takých identifikátorov pre jednotlivé uložené entity, že namiesto hľadania entít vo views ich budeme vedieť následne nájsť iba pomocou kľúča.

Autentifikácia. Pre nasadenie do produkcie je veľmi dôležitá aj bezpečnosť. Doteraz sme vyvíjali len u seba na lokálnom počítači a teda sme neriešili produkciu. Avšak momentálne s príchodom nášho prvého používateľa je nutné vyvinúť systém autentifikácie. Autentifikáciu a autorizáciu prístupu k API implementujeme s možnosťou obmedzenia na čítanie výsledkov, vkladanie nových záznamov a administráciu.

Autentifikácia je založená na Json Web Tokenoch, a preto nám odpadá nutnosť mať akúkoľvek databázu pre autentifikáciu. Tento mechanizmus bude potrebné dôkladne otestovať a v prípade potreby implementovať dodatočnú funkcionálnu funkciu ktorá vyplynie z požiadaviek zákazníkov pri testovaní.

100 najhorších. Výhodou používania nášho systému je hlavne rýchle nájdenie zlých komentárov. Naším cieľom je teda ponúknuť SME.sk 100 najhorších príspevkov za posledných 24 až 72 hodín, podľa ich nastavenia. Týmto spôsobom nájdú zlé komentáre oveľa rýchlejšie. Pre túto funkcionálnu funkciu bude potrebné vystaviť v api prístupový bod, ktorý bude pristupovať cez nejaké view do couchbase a hľadať najhoršie príspevky za zadaný čas. Výsledkom bude zoznam identifikátorov 100 najhorších príspevkov, ktorý si oni vyrenderujú v ich rozhraní.

Refaktor javascriptu. SME.sk chce výsledky analýz zobrazovať priamo v našom portáli a aj kvôli nim, aj do budúcnosti bude pre nás dôležité urobiť refaktor javascriptu, ktorý zobrazuje grafy, nakoľko sa pri ňom využíva veľa princípov metaprogramovania a nie je v tom prehľad. Upravovať tento javascript je veľmi zdĺhavé a často to môže viesť k nečakaným dôsledkom. Preto je veľmi dôležité urýchlene tento javascript upraviť, aby bolo možné a jednoduchšie vyvíjať nad ním niečo.

Ďalší vývoj

Filtrovanie a zoradovanie. Naše moderačné rozhranie (portál) je taktiež momentálne len prototypom a teda nie je pripravený na úplne reálnu prevádzku. Sú v ňom síce zobrazené

komentáre, ale portál funguje svižne iba pre málo komentárov. Toto však nie je prípad diskusných portálov. Na niektorých slovenských denníkoch pribudnú denne tisícky komentárov. Niektoré články obsahujú stovky komentárov a našou úlohou bude zachovať aj pri takýchto množstvách komentárov prehľadnosť a svižnosť portálu. Potrebné bude najmä zavedenie stránkovania, alebo tzv. nekonečného skrollovania s postupným načítavaním komentárov do zoznamu. Nad takýmto množstvom príspevkov budeme musieť umožniť aj efektívne filtrovať a zoradovať. Niektoré akcie a možnosti filtrovania a zoraďovania, ktoré by radi uvítali moderátori sme konzultovali aj s moderátorom z Denníku N. Išlo o filtrovanie podľa článku, autora článku, podľa viacerých článkov a zoraďovanie podľa času alebo kvality komentára odhadnutej našou službou.

Portál. Popri práci na integrácii našich výsledkov do už existujúcich moderačných rozhraní chceme pokračovať aj vo vývoji nášho používateľského rozhrania (portálu), ktoré chceme optimalizovať pre efektivitu moderovania. Naším cieľom je teda spraviť rozhranie v ktorom možno čo najefektívnejšie označovať príspevky za vhodné, alebo nevhodné do diskusie. Všetky rozhodnutia chceme zjednodušovať pomocou analýz z našej služby a takto moderátorom poskytovať nástroj pre podporu rozhodovania.

Webhooky. Veľkú prioritu má aj implementácia vrátenia výsledkov analýz pomocou webhookov. Toto bude uprednostňovaný spôsob vrátenia výsledkov kvôli jednoduchšiemu škálovaniu a nižšej záťaži na naše servery v porovnaní s dopytovaním sa na výsledky. Pri konzultácii so sli.do ohľadom prípadnej integrácie sme sa rozprávali o možnostiach použitia technológie websocketov na medziserverovú komunikáciu. Ak by aj spolupráca so sli.do nevyšla, tak websockety budeme naďalej vyhodnocovať ako jednu z reálnych možností na prepojenie serverov diskusných portálov s našimi API servermi.

Pravidlový klasifikátor. Naša služba je založená na klasifikátore, ktorému sú predložené črty a ktorý sa na základe už ohodnotených dát učí. Do toho by sme však radi zapracovali fuzzy vyhodnocovanie, kde by sme si dokázali určiť, pri akej hodnote, ktorej črty môžeme považovať komentár za zlý alebo naopak dobrý. Mohli by sme sa riadiť kódexmi jednotlivých denníkoch, ale vyžaduje si to aj ďalšie konzultácie s denníkmi, ktoré majú o týchto pravidlách portálov s našimi API servermi.

Celkový pohľad na systém

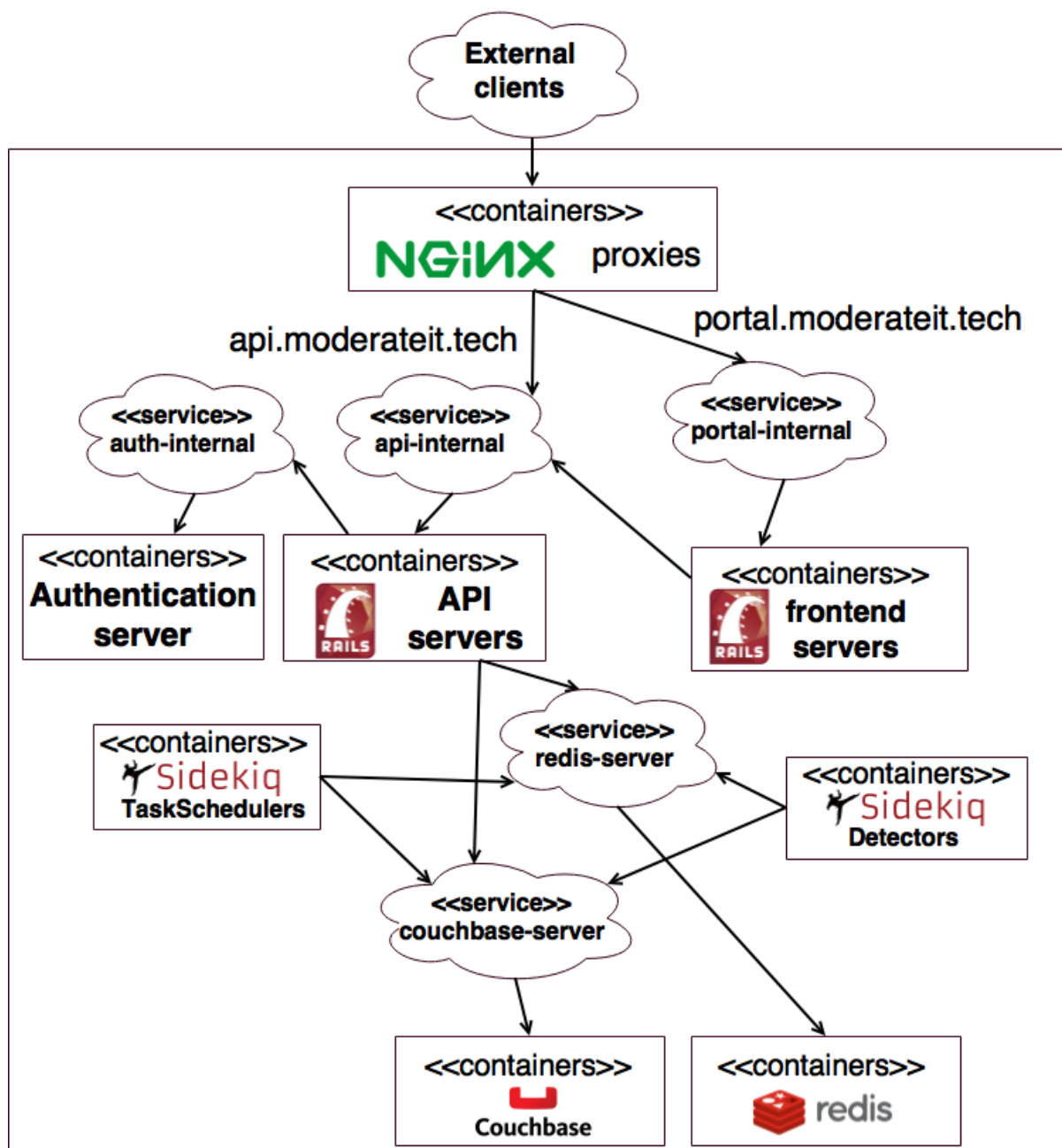
Úvod do architektúry

Jadrom služby je schopnosť prijať nový komentár, analyzovať ho a poskytnúť výsledky analýzy zákazníkovi. Keďže pri analýze využívame aj externé služby a celková doba analýzy môže narásť aj na niekoľko sekúnd, tak sme sa rozhodli, že analýzy sa budú vykonávať asynchrónne na pozadí. Požiadavka na analýzu nového komentára teda končí potvrdením o prijatí a nečaká sa na výsledky analýzy. Po prijatí komentára sa v odpovedi vráti aj URL odkazujúci na miesto kde budú dostupné výsledky po skončení analýzy. V ďalších iteráciách plánujeme implementovať vrátenie výsledkov pomocou Webhookov - teda preddefinovaných HTTP endpointov, na ktoré po skončení analýzy klientovi pošleme výsledok.

Rozmiestnenie a nasadenie komponentov

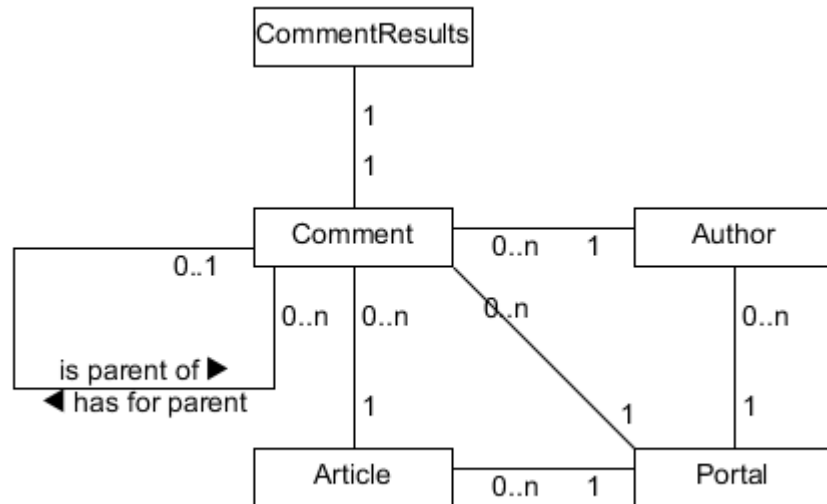
Služba moderateIT je nasadená v Microsoft Azure cloude, pretože na Azure máme k dispozícii kredit prostredníctvom programu Microsoft BizSpark. V Azure sme vytvorili cluster pozostávajúci z niekoľkých virtuálnych strojov. Služby nasádzame použitím Docker kontajnerov, o ich rozmiestnenie v clustri sa stará Openshift, ktorý nám výrazne zjednodušuje činnosti súvisiace s prevádzkou infraštruktúry (napríklad vnútorné DNS, rozmiestnenie kontajnerov, automatické zotavovanie po zlyhaní, load balancing, bezvýpadkové aktualizácie služieb).

Nasledujúci diagram zobrazuje vysokoúrovňové prepojenie medzi jednotlivými komponentami aplikácie.



Dátový model

Dátový model možno vidieť na nasledujúcom obrázku. Tento model sme vytvorili v databáze Couchbase. Couchbase je databáza, ktorá obsahuje dokumenty uložené v úložisku typu kľúč-hodnota. Nepodporuje žiadne cudzie kľúče, a preto si všetky validácie robíme ručne.



Moduly

- API
- Autentifikačný server
- Klasifikátor
- Diakritikovac
- Portal

Moduly systému

Autentifikačný server

Analýza

Pri každom requeste do API potrebujeme overiť a zistiť o ktorého zákazníka ide. Preto potrebujeme efektívny spôsob ako tieto informácie z requestu vytiahnuť. Zároveň by sme chceli zachovať jednoduché a bezpečné používanie API z pohľadu zákazníka. Zvolili sme autentifikáciu pomocou JSON Web Tokenov (JWT). JWT je JSON, ktorý v sebe nesie akúkoľvek informáciu, ktorú potrebujeme, v našom prípade nás zaujíma predovšetkým zákazníckove ID a čas expirácie tokenu. JWT je bezpečne podpísaný našim súkromným kľúčom, ktorý sa používa aj pri overení tokenu, takže je zabezpečené, že token ktorý prejde validáciou sme vydali my a nevyrobil ho niekto iný. V neskorších fázach do JWT môžeme uložiť aj informáciu o právach, ktoré je s ním možné vykonávať - zákazníci si tak budú môcť vygenerovať osobitné tokeny pre rôzne prípady použitia, napríklad moderátori by mohli mať právomoc čítať, schvaľovať a zamietat' komentáre, avšak nemohli by meniť konfiguráciu systému.

Identifikovali sme 4 hlavné požiadavky na túto službu:

1. dokázať vytvoriť nový JWT s informáciami o zákazníkovi
2. dokázať validovať JWT - teda povedať, či bol JWT vytvorený našim systémom (podľa podpisu) a či už nie je expirovaný
3. vytiahnuť z JWT údaje o zákazníkovi
4. rýchla odozva - keďže každé jedno volanie do API musí prejsť autentifikačným serverom, tak je potrebné, aby tento server nespôsoboval ďalšie spomalenie alebo úzke hrdlo (bottleneck) pri vybavovaní requestov

Návrh

Autentifikačný mechanizmus sme sa rozhodli implementovať ako službu, pretože predpokladáme, že v budúcnosti to bude potrebovať viac aplikácií a nechceme mať viac implementácií tej istej logiky v rôznych aplikáciách. Jednotlivé časti systému by teda mali vedieť fungovať bez akejkoľvek znalosti o tom čo je to JWT a ako sa spracováva, pretože túto logiku za nich rieši autentifikačný server. V prvej fáze od služby požadujeme aby pomocou HTTP volania vedela overiť platnosť JWT a vytiahnuť z neho ID zákazníka. Pri generovaní nového JWT pomocou HTTP volania by bolo nutné zabezpečiť, aby toto volanie okrem nás nemohol urobiť nik iný. Táto operácia nazačiatku nebude častá a nebude potrebné ju automatizovať, preto stačí, aby sa nové JWT dali generovať ručne z príkazového riadku.

Implementácia

Na implementáciu sme zvolili jazyk Go, ktorého vlastnosti ho robia takmer ideálnym na tento typ služby - malá, jednoduchá, výkonná služba dostupná pomocou HTTP API. Jazyk Go má v sebe všetko potrebné na implementáciu tejto služby, ako jedinú externú knižnicu používame knižnicu go-jwx, ktorá obsahuje všetko potrebné na prácu s JWT.

Veľkou výhodou použitia JWT je to, že potrebné informácie o zákazníkovi dokážeme uložiť priamo do tokenu. Toto umožňuje aby služba bola tzv. bezstavová (stateless), čiže nepotrebuje načítavať a ukladať údaje do databázy alebo na disk.

Služba má 2 návratové kódy:

- 403 Forbidden - poskytnutý JWT je z nejakého dôvodu neplatný (môže byť zle naformátovaný, expirovaný, neplatný podpis,...).
- 200 OK - poskytnutý JWT je platný. V tomto prípade je v HTTP odpovedi nastavená hlavička X-Moderateit-Portal, ktorej hodnota je zákazníkove ID v našom systéme (vytiahnuté z JWT)

Testovanie

Keďže vytvorená služba poskytuje zatiaľ len jednoduchú funkcionálnosť, tak sme ku nej nepísali automatické testy. Jej funkčnosť je možné overiť niekoľkými HTTP volaniami, ktoré sa dajú rozdeliť do 2 kategórií:

- poskytnutý platný JWT - návratový kód volania musí byť 200 a v HTTP odpovedi musí byť nastavená hlavička X-Moderateit-Portal s hodnotou zákazníkovo ID vytiahnutého z JWT
- čokoľvek iné musí vrátiť návratový kód 403 BEZ nastavenej HTTP hlavičky X-Moderateit-Portal. Čokoľvek iné zahŕňa expirované tokeny, tokeny vygenerované s použitím iného podpisu, náhodné a prázdne (resp. neposkytnuté) tokeny.

Jednou z požiadaviek na službu bola aj jej rýchla odozva, čo sme otestovali jednoduchým benchmarkom s použitím programu Apache Benchmark. Test pozostával zo 100 000 HTTP volaní na overenie poskytnutého tokenu, pričom bolo robených vždy 20 volaní súbežne.

Tu sú výsledky po niekoľkých testoch:

Requestov za sekundu: 5758.67

Priemerná doba odozvy: 3.473ms

Z testov tiež vyšlo, že 95% requestov malo dobu odozvy 10ms a menej, 99% malo dobu odozvy 19ms a menej, najhoršia doba odozvy bola 168ms. Tieto výsledky považujeme za dostatočne dobré pre naše použitie, obzvlášť keď vezmeme do úvahy, že testovaná záťaž bola približne 1000x vyššia, ako očakávame pri nasadení služby pre zákazníka veľkosti SME.sk.

Klasifikátor

Analýza

Pri analýze komentárov z príspevku extrahujeme množstvo znakov. Vstupom pre klasifikátor sú tieto jednotlivé znaky a výstupom musí byť jedno číslo, celková kvalita komentára. Klasifikátor bol vytvorený už minulý akademický rok v rámci bakalárskej práce. Trénovaný a testovaný je na dátach z portálu SME.sk, nakoľko je to náš prvý zákazník. Pre každého zákazníka budeme vytvárať vlastný klasifikátor.

Návrh

Klasifikátor sme tiež navrhovali ešte minulý rok v rámci bakalárskej práce. Je možnosť použiť viacero modelov. Ide najmä o rozhodovacie stromy, neurónové siete, alebo SVM. Tieto všetky sme vyskúšali a následne sme vybrali taký model, ktorý mal najväčšiu správnosť nad testovacím datasetom a zároveň mal čo najvyššie F1 skóre.

Implementácia

Náš klasifikátor beží v prostredí Microsoft Azure Machine Learning. V tomto prostredí sa dá jednoducho natrénovať model, ktorý je možné vystaviť vo forme HTTP REST API. Naše API sa následne pripája ku klasifikátoru prostredníctvom tohto API kedykoľvek je potrebné získať celkové hodnotenie komentára na základe jeho opisu pomocou znakov, ktoré z neho extrahujeme.

Azure Machine Learning Studio používame kvôli tomu, že minulý rok sme sa zúčastnili technologickej súťaže Imagine Cup, kde bolo použitie Azure jednou z postačujúcich podmienok účasti.

Testovanie

Úspešnosť klasifikátora testujeme použitím testovacej množiny dát priamo v Azure Machine Learning Studio. To, či funguje vystavené API testujeme ukázkovým volaním, ktoré je možno spraviť opäť priamo v Azure Machine Learning Studio.

Diakritikovač

Analýza

Keďže veľká časť diskusných príspevkov je písaná bez použitia diakritiky, rozhodli sme sa na zvýšenie presnosti analýzy túto diakritiku do príspevkov doplniť. Existuje viacero verejne dostupných služieb, ktoré sú určené práve na túto úlohu.

Jednou z nich je napríklad štatistický diakritikovač vytvorený na Fakulte informatiky a informačných technológií, dostupný na adrese <http://text.fiit.stuba.sk:8081/>, ktorý dosahuje úspešnosť až 98%. Použitie tejto služby je pomerne jednoduché no rýchlosť jeho odozvy je niekedy nedostatočná.

Ďalšou možnosťou je služba poskytovaná Jazykovedným ústavom Ľudovíta Štúra, dostupná na adrese <http://korpus.juls.savba.sk/diakritik.html>. Táto služba v niektorých prípadoch dosahuje lepšie výsledky a má aj kratšiu dobu odozvy, ako vyššie spomínaná služba, no jej nevýhodou je, že nie je poskytovaná ako API, takže práca s ňou je zložitejšia.

Návrh

Rozhodli sme sa používať vyššie spomínaný štatistický diakritikovač, ktorý sme so súhlasom autora modifikovali a kvôli zlepšeniu odozvy nasadili na vlastný server, aby sme neboli závislí na službe poskytovanej treťou stranou.

Oproti pôvodnej službe došlo k niekoľkým zmenám. Prvou zmenou je zmena formátu správ ktorými toto API komunikuje. Ďalšou je zmena databázy, ktorú používa.

Implementácia

Keďže konkrétna implementácia tejto služby nie je dielom žiadneho z členov tímu, nebude opisovaná do detailov a vyjadríme sa len k zmenám, ktoré boli nami vykonané a všeobecnému opisu tejto služby.

Tento diakritikovač pri rekonštrukcii vychádza z kontextu. Využíva sa jazykový model slovenčiny, ktorý bol vytvorený na JÚLŠ SAV v Bratislave. Úspešnosť rekonštrukcie sa pohybuje nad úrovňou 98%. Diakritika je doplnená len do gramaticky správne napísaných slov. Slová s preklepmi nie sú rekonštruované.

Je implementovaný v jazyku python a využíva hlavne opensource knižnice, ako sú kenlm (práca s jazykovým modelom) a web.py (spracovanie http requestov).

Ako bolo vyššie spomenuté, oproti pôvodnej implementácii sme urobili niekoľko zmien. Prvou zmenou je, že ako databázu sme namiesto MySQL použili redis, hlavne kvôli jeho rýchlosti. Druhou zmenou bola zmena API rozhrania. Namiesto pôvodne použitých GET requestov, kedy sa text určený na diakritikovanie posielal ako parameter v url, sme sa rozhodli použiť POST requesty obsahujúce jednoduchý JSON. Taktiež bola zmenená forma odpovedí z API. Namiesto pôvodných XML správ sa v súčasnosti taktiež posielajú jednoduché JSON-y.

Testovanie

Testovanie tejto služby je vykonané v rámci testovania triedy DiacriticDetector. Počas týchto testov sa porovnáva očakávaný textu s diakritikou a text, ktorý pre rovnaký text bez diakritiky vráti po zaslaní requestu diakritikovač. Taktiež je testovaný aj kód odpovede na http POST request.

Portál

Analýza

Primárnym účelom portálu bola priniesť verejnosti predstavu o práci moderátorov a taktiež priblížiť ako naša služba funguje. Či už grafickým znázornením jednotlivých znakov komentárov a autorov alebo možnosť napísať si vlastný komentár, ktorý bude do pár sekúnd vyhodnotený. Taktiež sa dali komentáre filtrovať podľa článkov, aby si moderátor mohol vybrať pod, ktorým článkom chce moderovať.

Portál bol predstavený denníku, ktorý nám zadefinoval svoje požiadavky, tak aby sa mohli dopracovať k výsledkom kvality komentára a a reputácie autora. Taktiež si denník vyžiadala prístup k bližším informáciám o komentári a autorovi.

Návrh

Pri vytváraní grafického rozhrania sme si dali záležať na tom, aby bolo čo najlepšie prispôbené pre používateľa. Snažili sme sa dodržiavať základné princípy UX dizajnu. Najskôr vznikol prvý prototyp na papieri, aby sme mali jasnú predstavu o tom kam umiestime rôzne prvky.

Na to, aby sme denníku mohli zobrazit' potrebné informácie sme museli náš portál upraviť, aby sa dal používať ako iframe. Moderátor bude mať k dispozícii výsledky komentára a autora v malom iframe. V závislosti od toho na aký graf klikne sa mu zobrazia podrobnejšie štatistiky autora alebo komentára. V neskorších fázach pribudne zoznam 100 najhorších príspevkov pre moderátora.

Implementácia

Pre používanie portálu denníkom je potrebné, aby si nastavili cookie `moderateit_api_token`, ktorá im zabezpečí, že sa nebudú musieť nikam prihlasovať na to, aby sa dostali k iframom a výsledkom, ktoré potrebujú. Do svojho moderačného rozhrania si implementujú iframy s odkazom na náš portál. Následne po kliknutí na výsledok komentára sa zobrazia v novom

okne grafy, ktoré znázorňujú črty komentára. Napríklad počet nadávok, gramatická správnosť komentára či bohatosť slovnej zásoby v komentári. To isté platí pre autora. Po kliknutí na malý graf v iframe znázorňujúci reputáciu používateľa sa zobrazia črty opisujúce autora. Napríklad odkedy je autor komentára zaregistrovaný alebo informácia o tom koľko príspevkov je zmazaných.

Zoznam článkov, zoznam komentárov a input pre zadanie nového komentára boli z portálu zmazané, pretože nie sú potrebné pre denník. Výsledky komentára a autora komentára a taktiež ostatné informácie sú dotiahnuté do portálu z api pomocou requestov.

Testovanie

Jedna z vecí, ktorú bolo pri portáli otestovať, bolo grafické rozhranie. Rozhranie sme dali vyskúšať známym, ale aj samotným predstaviteľom denníkov. Dostali sme spätnú väzbu o tom, že portál je veľmi ľahko použiteľný. Informácie sú zobrazené prehľadne.

Tiež sme testovali portál ručne či sa grafy správne zobrazujú aj v prípadoch keď výsledky nemáme k dispozícii. Na autentifikáciu cez cookie máme napísaný test.

API

Analýza

moderateIT API je hlavnou časťou nášho projektu. Do toto API sú zasielané dáta zákazníka, ktoré má služba ohodnotiť. Po prijatí potrebných dát a ich zvalidovaní je tento modul zodpovedný za spustenie potrebných analýz, uloženie ich výsledkov do databázy a ich následné zaslanie na vstup klasifikátora, ktorý bol opísaný vyššie.

S týmto API však nekomunikuje len strana zákazníka, ale rovnako aj nami vytvorený portál. Ten po prijatí požiadavky na zobrazenie výsledkov analýz urobí request do tohto api a zobrazí získané výsledky.

Návrh

Už od začiatku projektu sme sa rozhodli vytvoriť API, ktoré bude dodržiavať "best practices" pre tvorbu REST API. Ide hlavne o dodržiavanie správneho pomenovávania, využívania správnych návratových kódov pre jednotlivé volania a ďalších maličkostí, ktoré však v konečnom dôsledku majú veľký prínos pre jednoduchšiu prácu s týmto API. Aplikovanie týchto princípov je možné vidieť napríklad v dokumentácii k tomuto API (<http://docs.moderateitapi.apiary.io>).

Implementácia

API je rovnako ako väčšina projektu implementovaná v jazyky RUBY, s použitím frameworku Rails. Celé toto API beží na webserveri puma (<http://puma.io/>), ktorý poskytuje lepšie možnosti ako defaultný rails server webrick.

Ako databáza na ukladanie informácií o článkoch, komentároch, autoroch a výsledkov analýz využívame dokumentovú NoSQL databázu Couchbase (<http://www.couchbase.com/>). Na

ukladanie medzivýsledkov analýz a konštánt používaných pri analýze používame Redis (<http://redis.io/>).

Keďže jednotlivé detektory spracúvajú dáta na pozadí, na ich manažovanie využívame externú knižnicu Sidekiq (<http://sidekiq.org/>).

Testovanie

Keďže sa jedná o najdôležitejšiu časť projektu, aj jej testovaniu sa prikladá najväčší význam. Je našou snahou pokryť testami každú časť tohto API od reakcie na requesty, cez činnosť jednotlivých detektorov až po komunikáciu, či už s portálom alebo klasifikátorom.