

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 zužitkovať 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 navštevovníkov stránok z dlhodobého hľadiska, dosiahneme tým ďalšie vstupy, ktoré skvalitnia našu analýzu a zároveň dokážeme portálom ukázať bližšie infomácie 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 funkcionálnu ktorá vyplynie z požiadaviek zákazníkov pri testovaní.