

Slovenská technická univerzita v Bratislave

Fakulta informatiky a informačných technológií



reCommers

Dokumentácia k inžinierskemu dielu

Tím: 03 - reCommers

Vedúci tímu: Ing. Ivan Srba, PhD.

Vyhotovil: Matúš Cuper, Ondrej Kaščák, Lukáš Manduch, Ivana Mujgošová, Jozef Pazúrik, Ondrej Pitoňák, Ondrej Selecký, Martin Schnürer

Dátum poslednej zmeny: 11.05.2018

Obsah

1. Úvod	3
2. Globálne ciele pre ZS	3
3. Globálne ciele pre LS	3
4. Celkový pohľad na systém	3
5. Hlavné moduly systému	6
5.1. Spracovanie udalostí	6
5.2. Odporúčanie	6
5.3. Webová aplikácia v Django	6
5.4. Zaznamenávanie a monitorovanie chýb	6
5.4. Asynchrónne spracovanie	6

1. Úvod

Tento dokument je dokumentáciou pre predmet Tímový projekt I., v ktorom študenti pracujú na rozsiahlych projektoch. Náš tím s názvom reCommers sa venuje odporúčaniam pre e-biznis.

V tejto dokumentácii k inžinierskemu dielu sa budeme venovať globálnym cieľom na zimný semester. Celkový pohľad na systém sa bude nachádzať v druhej kapitole. Tretia kapitola je zameraná na moduly systému, každý modul je opísaný v podkapitolách analýza, návrh, implementácia a testovanie.

2. Globálne ciele pre ZS

Globálne ciele pre zimný semester sú menšieho rozsahu ako sú plánované na letný semester. Je to z toho dôvodu, že len teraz spolu začíname pracovať ako tím a musíme sa naučiť tak aj efektívne vykonávať robotu. Zo začiatku je potrebné pridelenie úloh členom tímu a inicializácia podporných nástrojov. Tieto aktivity vyžadujú svoj čas, preto sa sme sa im venovali prvé týždne.

Ako hlavný cieľ pre zimný semester je nahradiť existujúcu infraštruktúru. To sa bude vykonávať nahradením a doplnením automatizovaných nástrojov na tvorbu štatistiky. Bude sa taktiež meniť spracovanie prichádzajúcich dát, ktoré sa posielajú ďalej. Ďalšou dôležitou súčasťou bude získavať a následne spracovať informácie o odporúčaníach

3. Globálne ciele pre LS

V letnom semestri by sme chceli náš produkt postupne vylepšovať a dopĺňať o nové súčasti a funkcionality, ktorá by umožnila jeho využitie tretími stranami. Aby bolo možné systém v budúcnosti rozvíjať a integrovať s ďalšími odporúčanými systémami, je nutné riešenia starostlivo zdokumentovať a otestovať. Výsledný produkt tak musí spĺňať všetky požiadavky na softvérový produkt. Výslednú kvalitu sa budeme snažiť dosiahnuť tvorivým prístupom, zodpovednou prácou na projekte a reálnou spoluprácou s elektronickými obchodmi a odborníkmi v oblasti odporúčania pôsobiacimi na fakulte.

4. Celkový pohľad na systém

Vytváraný systém možno rozdeliť na dve základné časti (obr. 1):

- **Front-End** - jedná sa o django aplikáciu umožňujúcu zobrazovať základné aj pokročilé štatistiky odporúčania ako napríklad CTR a počet impresií. Okrem zobrazovania štatistík poskytuje rozhranie možnosť konfigurácie odporúčačov a ich verzií. Z oblasti spracovania udalostí umožňuje zobrazovať a schvaľovať štruktúru a počty prichádzajúcich udalostí. Okrem toho obsahuje možnosť vyhľadávania v histórii používateľov a poskytuje široký druh historických údajov.
- **Back-End** - obsahuje všetku logiku potrebnú pre zber, spracovanie a uloženie udalostí spolu s poskytovaním odporúčaní.

V súčasnosti máme k dispozícii dve možnosti získavania udalostí (angl. events) od nášho zákazníka. Buď nám zákazník posiela udalosti na náš vystavený end-point, alebo udalosti získavame z Kafky zákazníka.

Po získaní udalostí nasleduje ich spracovanie, ktoré spočíva najmä v úprave ich štruktúry, keďže nie vždy prichádzajú udalosti s požadovanou štruktúrou. Modifikované udalosti potom možno uložiť do databázy odporúčacieho systému (v našom prípade platforma Prediction.io) a taktiež uložiť mapovanie ID používateľa a cookies do našej PostgreSQL databázy.

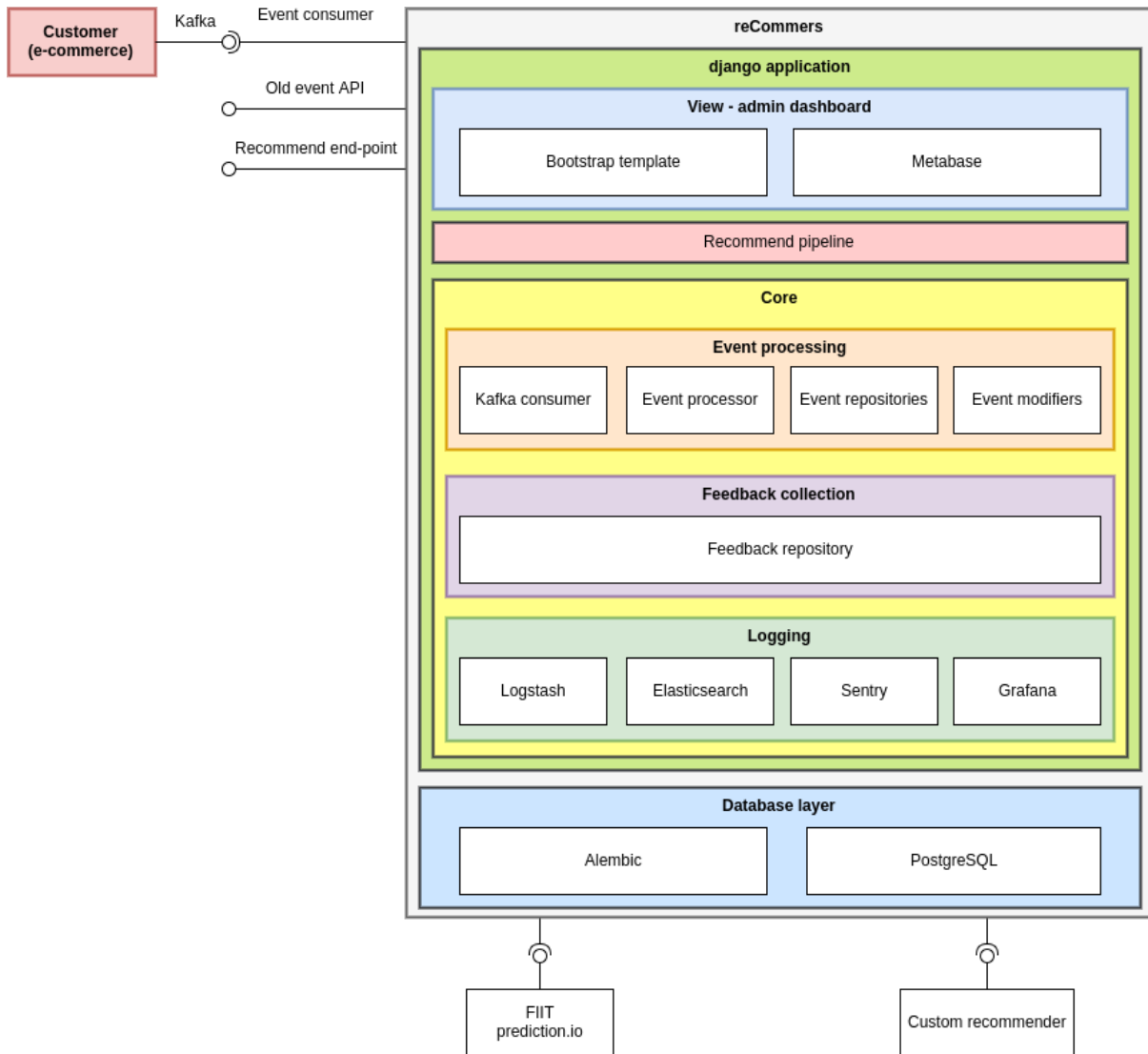
Súčasťou aplikácie je aj kontrola štruktúry a počtu prichádzajúcich udalostí, ktoré je možné skontrolovať vo webovom rozhraní prostredníctvom grafov a tabuliek.

Ďalšou veľmi dôležitou časťou je získavanie a ukladanie spätnej väzby (angl. feedback) kvôli možnosti vyhodnotenia úspešnosti odporúčania a spomínaných štatistík vo webovej aplikácii.

Chod celého systému je zaznamenávaný pomocou logov, ktoré putujú cez Logstash do databázy Elasticsearch. Na notifikovanie v prípade aplikačných chýb a varovaní využívame službu Sentry. Službu Grafana využívame na sledovanie štatistík týkajúcich sa rýchlosti odozvy našej aplikácie.

Pre poskytnutie odporúčania máme vystavený end-point, na ktorý ak príde požiadavka pre poskytnutie odporúčania, tak zákazníkovi obratom zašleme vygenerované odporúčanie so zoznamom položiek.

V prípade odporúčania je vytvorený systém na prídávanie typov odporúčačov a ich verzií. Tu je možné pre zvolený odporúčač a verziu vybrať, aké percento odporúčaní vygenerujú.



Obr. 1: Architektúra systému

5. Hlavné moduly systému

Tieto moduly sú opísané v dokumente o modeloch systému:

5.1. Spracovanie udalostí

Opisuje spracovanie udalostí a s ním súvisiace podmoduly:

- Spracovanie udalostí
- Kafka konzument

5.2. Odporúčanie

Opis odporúčačov a ich konfigurácie.

- Odporúčanie
- A/B filtrovanie
- Šablóna odporúčacieho systému
- Schema builder

5.3. Webová aplikácia v Django

Hlavné časti webovej aplikácie.

- Webová aplikácia v Django
- Správa používateľov a autentifikácia
- Zoznam nástrojov v menu
- Ladenie histórie používateľa
- CRUD používateľov
- Definícia odporúčačov a ich verzíí

5.4. Zaznamenávanie a monitorovanie chýb

Informácie o logovacích a monitorovacích nástrojoch.

5.4. Asynchrónne spracovanie