

Review šprintu “Ixobrychus + Jynx”

Dátum: 26.4.2017
Miesto: Projektové laboratórium
Vedúci stretnutia: Filip Vozár
Zapisovateľ: Jakub Adam
Prítomní: Jakub Adam
Monika Filipčíková (on-line)
Jakub Šimko
Andrej Švec
Filip Vozár

Zhrnutie

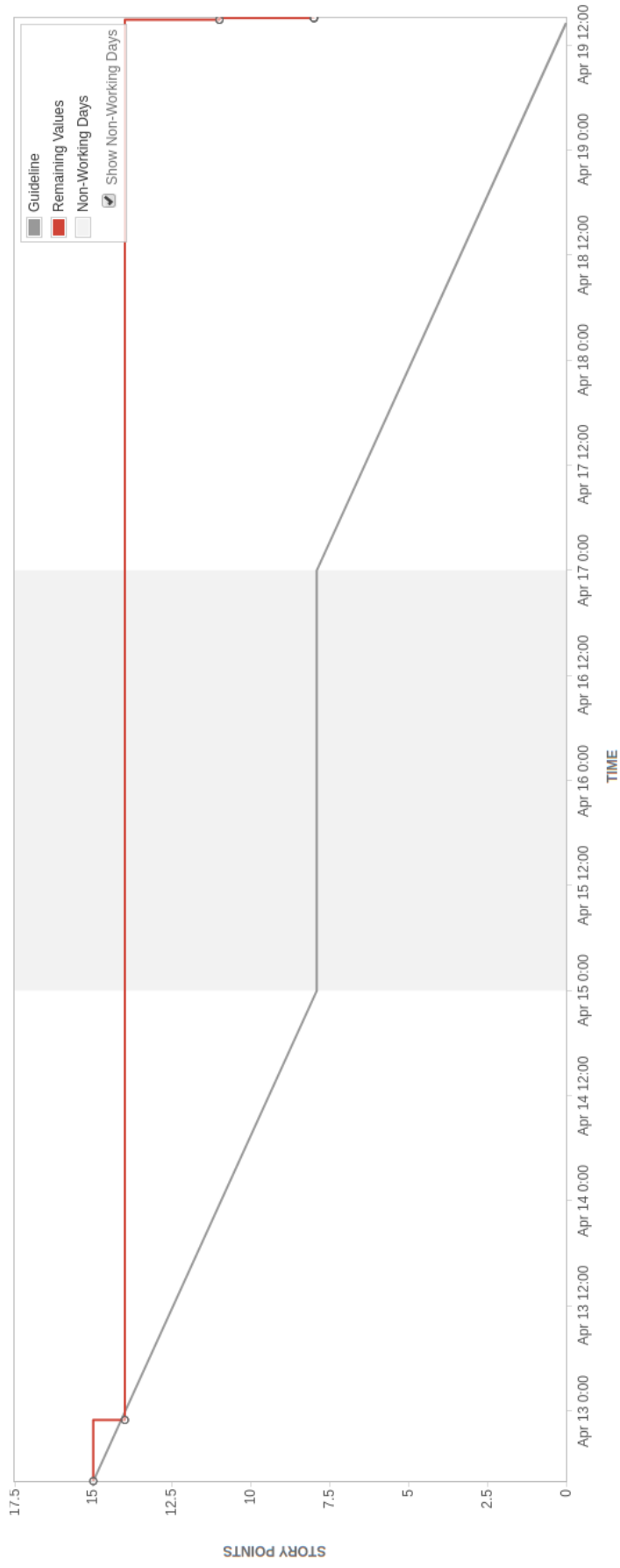
Šprint Ixobrychus bol prvým pokusom o zavedenie týždňových šprintov. Splnila sa väčšina úloh zaradených do šprintu. Nepodarilo sa dokončiť len úlohu MD-1, ktorá pozostávala z dvoch podúloh. Jednou bolo vytvorenie endpointu do API a druhou bola úprava portálu, aby volala tento endpoint a zobrazovala korektne výsledky. Obe tieto podúlohy boli značne v rozpracovanom stave, no nestihli sa dokončiť kvôli zlému definovaniu požadovaných dát. Nedostatočne sme si definovali obsah návratového JSON-u, takže sa robili časte zmeny implementácie a možno aj kvôli veľkonočným sviatkom sa to nestihlo dokončiť.

V šprinte Jynx sa nám podarilo splniť všetky zaradené úlohy, medzi ktorými boli aj tie nestihnuté z predošlého šprintu.

Šprint review

- Používať N1QL na dopytovanie v couchbase
- Bude sa treba zamyslieť nad zavedeným indexom v DB
- Zohľadniť viac črt pri fuzzy klasifikátore
- Definovať rozhrania pred implementáciou
- Nemáme gausovské rozdelenie pri výsledkoch. Objavujú sa tam 2 peaky, takže je potrebné zistiť príčinu pozretím sa na konkrétne dáta.
- Nad zatiaľ ohodnotenými komentármi z topiek boli spustené detectory, ale bez offtopicu, keďže nemáme k dispozícii text článkov.
- Boli vizualizované získané dáta, ale grafy je potrebné prerobiť znova, keďže došlo k nesprávnemu behu skriptu na nekompatibilnej verzii pythonu.
- Pri vyhľadavaní nevhodných komentárov sa prejavil veľký vplyv BagOfWords detectoru. Výsledky sa v druhej iterácii značne zlepšili a cca polovica vyzerá byť nevhodná.

- Spraviť lepší výber “dobrých” príspevkov do trénovacej množiny. Zistilo sa, že takmer všetky “dobré” komentáre sa vyjadrujú k Ficovi. Pravdepodobne preto, že príspevky spomínajúce Fica majú veľa upvote-ov.
- Zamyslieť sa nad možnosťou pravidelného pretrenovávania klasifikátora, aby sme lepšie vedeli hodnotiť aktuálne témy diskusií.



Completed Issues

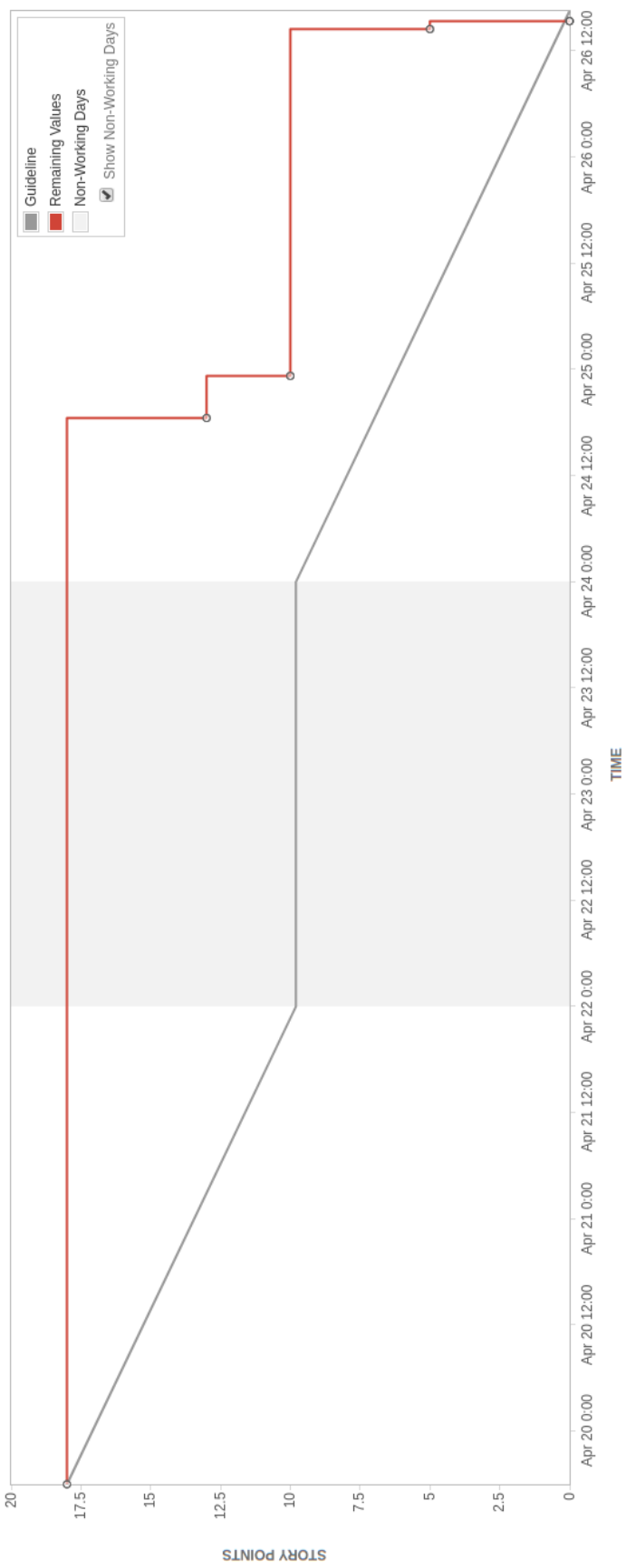
[View in Issue Navigator](#)

Key	Summary	Issue Type	Priority	Status	Story Points (7)
MD-152	Opravit Vocabulary Richness detector	Bug	↑ Medium	DONE	1
MD-156	Napisat pravidla pre fuzzy - 2. iteracia	Story	↑ Medium	DONE	3
MD-163	Pohladat zle komentare v SME datach VOL. 2	Story	↑ Medium	DONE	3

Issues Not Completed

[View in Issue Navigator](#)

Key	Summary	Issue Type	Priority	Status	Story Points (8)
MD-1	Zobrazit 100 najhoršich príspevkov za posledných 24 hodin	Story	↑ Medium	TO DO	8



[View in Issue Navigator](#)

Completed Issues

Key	Summary	Issue Type	Priority	Status	Story Points (18)
MD-1	Zobrazit 100 najhorších príspevkov za posledných 24 hodín	Story	↑ Medium	DONE	5
MD-166	Pohrat sa s datami z Topiek	Story	↑ Medium	DONE	5
MD-167	Vyhodnocovanie uspesnosti fuzzy - T-test	Story	↑ Medium	DONE	3
MD-169	Precistenie datasetu zo SME	Story	↑ Medium	DONE	5