

20. Stretnutie 14.4.2020

Agenda

1. Ukoncenie sprintu
2. Retrospektiva
3. Planovanie sprintu

Priebeh stretnutia (hangouts call)

- Ukoncenie sprintu
 - [CAR-203](#) - Pridat detaily auta pod vyhľadany inzerat
 - pocitanie percentilu chceme inak - nezaratat tam to konkretne auto lebo nedosiahneme nikdy maximum progress baru
 - [CAR-206](#) - Automatické rozpoznávanie entít
 - Anotovali sme vsetci
 - Word2vec na vsetkych freetextoch
 - Klasifikacia aj s okolim
 - Implementovane v api - vracia zoznam atribs z freetextu
 - Uz teda spracovavame bazos ale je tam nejake oneskorenie kym sa zapne text procesor od cca 30teho inzeratu (pri iba bazosi) to az zacne spracovavat
 - [CAR-201](#) - Obnova stavu crawlovania po restarte
 - [CAR-202](#) - Upravenie typu vystupnych hodnot parsera
 - storka je ok ale vynara sa nam tu problem Max ked updatoval celu DB - trvalo to velmi dlho - a toto bude iba narastat s poctom inzeratov - je optimalne vyuzivane mongo? Ak ano je tu priestor zvazit pouzitie SQL databazy
 - [CAR-176](#) - Hladanie podobnosti cez vektory
 - Divny jav ze najpodobnejsie auta nemaju podobny brand
- Retro
 - Pri opise storiek pri akceptacii - nerobit take vhupnutie do situacie ktore povazuje za implicitne ze vsetci, obzvlast MIro vedia o co ide, lepsie perzentovat
 - NLP model - sucast nejakeho repa by mal byt nie len v colabe
 - FE tasky - dokladnejsie review? Vychadzame z nepochopenia percentil baru, Zaver: ked sa nasadi FE vsetci to pojdu pozriet
- Planovanie sprintu
- Chceme si dat spolocny call k dokumentacii -> v nedelu