

Splnené stories

[CAR-102] Implementovať URL Updater Created: 21/Nov/19 Updated: 05/Dec/19 Resolved: 05/Dec/19	
Status:	Done
Project:	carAD
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Branislav Baláž	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	Not Specified	Remaining Estimate:	Not Specified
Σ Time Spent:	Not Specified	Time Spent:	Not Specified
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-116	Prechod DB a vkladanie do queue	Sub-task	Done	Peter Bokor
	CAR-125	vytvorit zapisovanie casov created/mo...	Sub-task	Done	Peter Bokor
Sprint:	CAR Sprint 4				

Description

Potrebuje mechanizmus, ktorý bude prechádzať databázu a vkladat odkazy do queue, aby sme vedeli, či je ešte inzerát platný.

[CAR-101] Vytvorit' API, ktorá bude z DB vracat' inzeráty Created: 21/Nov/19 Updated: 04/Dec/19 Resolved: 04/Dec/19

Status:	Done
Project:	carAD
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Veronika Vejcikova	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Σ Time Spent:	6 hours	Time Spent:	Not Specified
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-112	Vytvorit route	Sub-task	Done	Ondrej Harnúšek
	CAR-113	Vratit inzerat podľa URL	Sub-task	Done	Ondrej Harnúšek
Sprint:	CAR Sprint 4				

Description


Nech vracia len aktuálne inzeráty

Dopyt na unikátny identifikátor - hash URL - obojsmerny hash nieco na styl <https://www.stubbornjava.com/posts/obfuscating-and-shortening-sequential-ids-with-hashids>

[CAR-100] [Vytvorenie crawlera a generatora](#) Created: 21/Nov/19 Updated: 02/Dec/19 Resolved: 02/Dec/19

Status:	Done
Project:	carAD
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Max Karel	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	0 minutes
Σ Time Spent:	7 hours	Time Spent:	4 hours
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified

Attachments:	 20191125_145032-1-1.jpg				
Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-103	1. Dummy Queue - Pozriet Multiprocessing	Sub-task	Done	Peter Bokor
	CAR-104	2. Callbacky - Printovanie	Sub-task	Done	Ondrej Harnúšek
	CAR-105	3. generovanie URL	Sub-task	Done	Max Karel
	CAR-106	4. Multiprocessing	Sub-task	Done	Peter Bokor
	CAR-107	5. Asynchrone	Sub-task	Done	Peter Bokor
	CAR-114	6. Zdefinovat a pouzivat interface pr...	Sub-task	Done	Max Karel
Sprint:	CAR Sprint 4				

Description

Ako product owner chcem mať crawler ktorý bude z queue získavať URL, ktoré má stiahnuť a asynchrónne vytvárať callbacky na spracovanie stiahnutého HTML.

Generátor bude vytvárať postupne vyberať záznamy z queue.

DoD: Crawler a generátor vedia vyberať z štandardnej Python queue nespracované URL a spracovať ich. Vie aj naplniť tu queue

Musí aj správne zachytávať a spracovať error message z http get

Vieme proces spustiť paralelne pre rozne portali

[CAR-66] [Skompletizovať veci na odovzdanie po 9. týždni](#) Created: 31/Oct/19 Updated: 25/Nov/19 Resolved: 22/Nov/19

Status:	Done
Project:	carAD
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Peter Bokor	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	Not Specified	Remaining Estimate:	Not Specified
Σ Time Spent:	Not Specified	Time Spent:	Not Specified
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-74	Zistiť čo sa odovzdáva	Sub-task	Done	Max Karel
	CAR-85	Vytvoriť dokumentáciu inžinierskeho d...	Sub-task	Done	Branislav Baláž
	CAR-86	Vytvoriť dokumentáciu riadenia projek...	Sub-task	Done	Max Karel
	CAR-87	Odovzdať všetko	Sub-task	Done	Branislav Baláž
Sprint:	CAR Sprint 3, CAR Sprint 4				

Description


DoD:

Pripravené veci na odovzdanie podľa pokynov na stránke

Nesplnené stories

[CAR-98] Implementovat queue pre URL inzeratov Created: 14/Nov/19 Updated: 05/Dec/19	
Status:	In Progress
Project:	carAD
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Peter Bokor	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Σ Time Spent:	30 minutes	Time Spent:	Not Specified
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified

Attachments:	 20191125_145933-1.jpg				
Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-108	1. Mongo queue	Sub-task	Done	Peter Bokor
	CAR-109	2. Callbacky cez Mongo - mapovanie na...	Sub-task	Done	Ondrej Harnúšek
	CAR-110	3. generovanie URL - vybrat starting ...	Sub-task	Ready for Review	Branislav Baláž
	CAR-111	4. prioritizovanie user requestov	Sub-task	Done	Tomáš Vrtal
	CAR-122	Pospajat všetky komponenty	Sub-task	In Progress	Max Karel
Sprint:	CAR Sprint 4, CAR Sprint 5				

Description

Ako product owner chcem mať queue, ktorá bude

- potrebujeme collection v MongoDB pre vkladanie a citanie zaznamov

- ~~potrebujeme procesor zaznamov ktory bude citat a spracovavat zaznamy~~
- po vybaveni zaznamu musi procesor poslat data asynchrone spat
- treba vyriesit' duplicitne zaznamy v queue - checkne sa DB a prida/neprida sa

Dôvod. previazanosť s ostatnými úlohami spomalovala proces

[CAR-89] [Parsovanie inzerátov z auto-moto inzeretných portálov](#) Created: 14/Nov/19 Updated: 05/Dec/19

Status:	In Progress				
Project:	carAD				
Component/s:	None				
Affects Version/s:	None				
Fix Version/s:	None				

Type:	Story	Priority:	Medium		
Reporter:	Miroslav Rác	Assignee:	Unassigned		
Resolution:	Unresolved	Votes:	0		
Labels:	None				
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified		
Σ Time Spent:	1 day, 7 hours, 2 minutes	Time Spent:	Not Specified		
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified		

Issue Links:	Blocks				
	blocks	CAR-61	Vykonať exploratívnu analýzu	To Do	
	blocks	CAR-70	Filtrovanie relevantných inzerátov	To Do	
	is blocked by	CAR-71	Analyzovať portály inzerátov	Done	
Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-91	Parsovanie pre bazos.sk	Sub-task	Done	Branislav Baláž
	CAR-92	Parsovanie pre autovia.sk	Sub-task	Done	Tomáš Vrtal
	CAR-93	Parsovanie pre auto.bazar.sk	Sub-task	Done	Richard Letanec
	CAR-94	Parsovanie pre autobazar.eu	Sub-task	Done	Veronika Vejcikova

	CAR-95	Parsovanie pre autobazar.sk	Sub-task	In Progress	Veronika Vejcikova
	CAR-115	Spravit jednotne interface pre bazary	Sub-task	Done	Richard Letanec
	CAR-117	Handlovanie errorov pre vsetky parser...	Sub-task	In Progress	
	CAR-124	Error handling pre bazos.sk	Sub-task	Done	Peter Bokor
Sprint:	CAR Sprint 4				

Description

Ako používateľ
chcem mať dostupné dáta z čo najviac inzerátov na slovenskom internete,
aby boli štatistiky a predikcie relevantné.

DOD:

- crawler pre kazdy portal
- dáta sú ukladané do DB v predpísanej štruktúre

Pre kazdy bazar implementovat 2 funkcie:

- `def parse_index(self, raw_text: str) -> list:` - vrati list URL koncovych inzeratov
- `def parse_detail(self, raw_text: str, url: str) -> str:` - vrati json objekt dat o aute

treba riesit aj spracovanie ked sa vrati stranka nenajdena

Dôvod: casova narocnost, vzajomna previazanost s inymi ulohami a zmeny, neskoro definovanie interface, slaba komunikacia o pomoc s taskom

[CAR-75] [Tool pre rychle anotovanie a anotovanie dat](#) Created: 07/Nov/19 Updated: 05/Dec/19

Status:	In Progress
Project:	carAD
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Branislav Baláž	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Σ Remaining Estimate:	Not Specified	Remaining Estimate:	Not Specified
Σ Time Spent:	Not Specified	Time Spent:	Not Specified
Σ Original Estimate:	Not Specified	Original Estimate:	Not Specified

Issue Links:	Blocks blocks CAR-59 Extrakcia údajov z textu inzerátu - w... Done				
Sub-Tasks:	Key	Summary	Type	Status	Assignee
	CAR-80	Anotovať dostupné dáta	Sub-task	In Progress	Tomáš Vrtal
	CAR-84	Analyzovať nástroje pre anotovanie dát	Sub-task	Done	Tomáš Vrtal
Sprint:	CAR Sprint 3, CAR Sprint 4				
Reviewer:	Branislav Baláž				

Description

DoD:1000 anotovaných inzerátov.

Comments

Comment by [Tomáš Vrtal](#) [05/Dec/19]

anotovane data

- <https://drive.google.com/drive/u/1/folders/1FT2ccNNbW48EELNmbOF2ffcdxyq-UW75>

Dôvod: nedodržanie akceptacnych kriterii a ignorovanie review procesu.