

Export úloh - Hruška

[MON-124] Newspaper crawler a parser Created: 19/Nov/18 Updated: 01/Apr/19 Resolved: 01/Apr/19

Status:	Done
Project:	MonAnt
Component/s:	Data providers
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Ivan Srba	Assignee:	Veronika Žatková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified	Remaining Estimate:	Not Specified
Time Spent:	Not Specified	Time Spent:	Not Specified
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	<table border="1"><thead><tr><th>Key</th><th>Summary</th><th>Type</th><th>Status</th><th>Assignee</th></tr></thead><tbody><tr><td>MON-264</td><td>Newspaper crawler</td><td>Sub-task</td><td>Done</td><td>Veronika Žatková</td></tr><tr><td>MON-266</td><td>Newspaper parser</td><td>Sub-task</td><td>Done</td><td>Veronika Žatková</td></tr><tr><td>MON-267</td><td>Vytvori migráciu</td><td>Sub-task</td><td>Done</td><td>Veronika Žatková</td></tr><tr><td>MON-269</td><td>Overí parsovanie externého html</td><td>Sub-task</td><td>Done</td><td>Veronika Žatková</td></tr><tr><td>MON-297</td><td>Fix newspaper parser</td><td>Sub-task</td><td>Done</td><td>Veronika Žatková</td></tr></tbody></table>	Key	Summary	Type	Status	Assignee	MON-264	Newspaper crawler	Sub-task	Done	Veronika Žatková	MON-266	Newspaper parser	Sub-task	Done	Veronika Žatková	MON-267	Vytvori migráciu	Sub-task	Done	Veronika Žatková	MON-269	Overí parsovanie externého html	Sub-task	Done	Veronika Žatková	MON-297	Fix newspaper parser	Sub-task	Done	Veronika Žatková
	Key	Summary	Type	Status	Assignee																										
	MON-264	Newspaper crawler	Sub-task	Done	Veronika Žatková																										
	MON-266	Newspaper parser	Sub-task	Done	Veronika Žatková																										
	MON-267	Vytvori migráciu	Sub-task	Done	Veronika Žatková																										
	MON-269	Overí parsovanie externého html	Sub-task	Done	Veronika Žatková																										
	MON-297	Fix newspaper parser	Sub-task	Done	Veronika Žatková																										
Epic Link:	Novinové články a blogy																														
Sprint:	Hruška, Ipa																														

Description

Pridat crawler a parser zalozeny na kniznici Newspaper.

Konkrétne parametre a eventy (konfiguracny JSON) a dalsia specifikacia bude doplnena v ramci ulohy [MON-123](#).

Vytvorit dva data providery:

- crawler - na vstupe zoznam domen + newspaper parametrov
- parser - na vstupe zoznam URL + newspaper parametrov

[MON-127] Reazenie a kombinovanie data providerov Created: 19/Nov/18 Updated: 26 /Mar/19 Resolved: 26/Mar/19

Status:	Done
Project:	MonAnt
Component/s:	Data providers
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Ivan Srba	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes	Remaining Estimate:	0 minutes
Time Spent:	1 week, 4 days, 6 hours	Time Spent:	4 hours
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Attachments:	navrh gui z tabule.jpg																																							
Sub-Tasks:	<table border="1"> <thead> <tr> <th>Key</th> <th>Summary</th> <th>Type</th> <th>Status</th> <th>Assignee</th> </tr> </thead> <tbody> <tr> <td>MON-206</td> <td>Refaktoring nazvov v DB tabulkach</td> <td>Sub-task</td> <td>Done</td> <td>Matúš Barabás</td> </tr> <tr> <td>MON-207</td> <td>Rozsirenie grafickeho rozhrania</td> <td>Sub-task</td> <td>Done</td> <td>David Tran Duc</td> </tr> <tr> <td>MON-208</td> <td>Vytvorenie podpory pre trigger eventu...</td> <td>Sub-task</td> <td>Done</td> <td>Dung Lam Tuan</td> </tr> <tr> <td>MON-209</td> <td>Pridanie podpory triggerovania evento...</td> <td>Sub-task</td> <td>Done</td> <td>Patrik Židuliak</td> </tr> <tr> <td>MON-211</td> <td>Krajšie admin rozhranie</td> <td>Sub-task</td> <td>Done</td> <td>Milan Cák</td> </tr> <tr> <td>MON-238</td> <td>Pridanie celery tasku k data provider...</td> <td>Sub-task</td> <td>Done</td> <td>Patrik Židuliak</td> </tr> </tbody> </table>					Key	Summary	Type	Status	Assignee	MON-206	Refaktoring nazvov v DB tabulkach	Sub-task	Done	Matúš Barabás	MON-207	Rozsirenie grafickeho rozhrania	Sub-task	Done	David Tran Duc	MON-208	Vytvorenie podpory pre trigger eventu...	Sub-task	Done	Dung Lam Tuan	MON-209	Pridanie podpory triggerovania evento...	Sub-task	Done	Patrik Židuliak	MON-211	Krajšie admin rozhranie	Sub-task	Done	Milan Cák	MON-238	Pridanie celery tasku k data provider...	Sub-task	Done	Patrik Židuliak
Key	Summary	Type	Status	Assignee																																				
MON-206	Refaktoring nazvov v DB tabulkach	Sub-task	Done	Matúš Barabás																																				
MON-207	Rozsirenie grafickeho rozhrania	Sub-task	Done	David Tran Duc																																				
MON-208	Vytvorenie podpory pre trigger eventu...	Sub-task	Done	Dung Lam Tuan																																				
MON-209	Pridanie podpory triggerovania evento...	Sub-task	Done	Patrik Židuliak																																				
MON-211	Krajšie admin rozhranie	Sub-task	Done	Milan Cák																																				
MON-238	Pridanie celery tasku k data provider...	Sub-task	Done	Patrik Židuliak																																				
Epic Link:	Manažment web monitoringu																																							
Sprint:	Fernet, Gin, Hruška, Ipa																																							

Description	
--------------------	--

Pre retazenie a kombinovanie data providerov bude potrebné:

- len minimalne zrefaktorovat nazvy v DB tabulkach
 - DB schema v [EAP projekte](#)
- rozšírit graficke rozhranie o zadavanie eventov a na ne nadviazanych data providerov
 - schema EA obsahuje aj konkretne priklady JSONov s rozšírenou konfiguraciou data providerov a priklad JSONu s konfiguraciou monitora
 - dodam aj obrazok pouzivatel'skeho rozhrania, resp. prekreslim ho na tabulu
- prichystat podporu pre data providery, aby mohli vykonat trigger eventu aj s prislusnymi atributmi
- nachystat jeden univerzalny listener, ktory bude pocuvat na vyvolany event trigger a zabezpeci naschedulovanie dalsieho data providera
- upravit existujuce data providery, aby triggerovali eventy
 - rss parser trigeroval event "new-article" s atributom "url"
 - [badatel.net](#) crawler trigeroval event "new-article" s atributom "url" a "html_id", cez tento event sa bude aj volat [badatel.net](#) parser a povodne priamo volanie parsera sa zrusi

[MON-130] Endpoint pre automaticke spustanie monitorov Created: 19/Nov /18 Updated: 25/Mar/19 Resolved: 24/Mar/19

Status:	Accepted
Project:	MonAnt
Component/s:	Central storage
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Ivan Srba	Assignee:	Matúš Barabás
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Time Spent:	7 hours, 20 minutes	Time Spent:	Not Specified
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	<table><thead><tr><th>Key</th><th>Summary</th><th>Type</th><th>Status</th><th>Assignee</th></tr></thead><tbody><tr><td>MON-255</td><td>Uprava endpointu /monitors GET - prid...</td><td>Sub-task</td><td>Done</td><td>Matúš Barabás</td></tr></tbody></table>	Key	Summary	Type	Status	Assignee	MON-255	Uprava endpointu /monitors GET - prid...	Sub-task	Done	Matúš Barabás
	Key	Summary	Type	Status	Assignee						
MON-255	Uprava endpointu /monitors GET - prid...	Sub-task	Done	Matúš Barabás							
Epic Link:	Manažment web monitoringu										
Sprint:	Hruška										

Description

Endpoint /monitors, GET sa rozšíri o novy parameter, ktoreho nastavenim sa vrati zoznam len takych monitorov a ich konfiguracii, ktore maju byt spustene.

Monitor ma byt spusteny, ak jeho posledna extrakcia ma cas ukoncenia (tabulka extractions, stlpec end_time) pred dlhsim casom, ako je zadenovana frekvencia spustania pre monitor (tabulka monitor, stlpec frequency, hodnota je v minutach).

[MON-131] Endpointy pre CRUD extrakcii Created: 19/Nov/18 Updated: 25/Mar /19 Resolved: 24/Mar/19

Status:	Accepted
Project:	MonAnt
Component/s:	Central storage
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Ivan Srba	Assignee:	Matúš Barabás
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Time Spent:	1 day, 4 hours, 35 minutes	Time Spent:	Not Specified
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	MON-246	Vytvori tabuky pre extrakcie a extr...	Sub-task	Done	Matúš Barabás
	MON-245	Vytvorit endpointy pre extrakcie	Sub-task	Done	Matúš Barabás
	MON-247	Vytvori endpointy pre ítanie dát	Sub-task	Done	Matúš Barabás
	MON-259	Pridanie stpca task_stats do extrakcii	Sub-task	Done	Matúš Barabás
	Epic Link:	Manažment web monitoringu			
Sprint:	Hruška				

Description	
--------------------	--

Vytvorit tabulky (presna struktura je v EAPe):

- extractions (tu treba pridať este stĺpce pre stav uloh z Celery)
- extraction_logs

Vytvorit endpointy pre zapis dat (detailne informacie su v [Google docu](#)):

- /extractions, POST
- /extractions/{extraction_id}, PUT
- /extractions/{extraction_id}/logs, POST

Vytvorit endpointy pre citanie dat:

- /monitors/{monitor_id}/extractions, GET
- /extractions/{extraction_id}, GET
- /extractions/{extraction_id}/logs, GET

[MON-139] Automaticke spustanie monitorov Created: 19/Nov/18 Updated: 30/Mar/19

Status:	In review
Project:	MonAnt
Component/s:	Web monitoring management
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Ivan Srba	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Time Spent:	3 hours	Time Spent:	Not Specified
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	MON-253	Spustanie python skriptu (data_provid...	Sub-task	Rejected	Dung Lam Tuan
	MON-254	Ziskanie zoznamu monitorov na spustan...	Sub-task	In review	Dung Lam Tuan
Epic Link:	Manažment web monitoringu				
Sprint:	Hruška, Ipa				

Description	
--------------------	--

Automaticke spustanie bude skript, ktorý bude súčasťou web monitoringu. Tento skript sa bude spúšťať v pravidelných intervaloch (každú minútu) z CRONu alebo Celery (čo bude jednoduchšie na setup) (frekvencia spustania je z configu alebo rozumne nastaviteľná).

Skript si cez endpoint centrálného úložiska (~~MON-139~~) vypýta zoznam monitorov, ktoré majú byť spustené. Prejde týmto zoznamom a jednotlivé extrakcie nascheduluje aj so zadanou konfiguráciou.

[MON-140] Zbieranie a odoslanie informácií o extrakcii Created: 20/Nov/18 Updated: 25/Mar/19

Status:	To Do
Project:	MonAnt
Component/s:	Data providers
Affects Version/s:	None
Fix Version/s:	None

Type:	Story	Priority:	Medium
Reporter:	Ivan Srba	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Time Spent:	1 day, 3 hours	Time Spent:	Not Specified
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	MON-258	Získa stavby úloh z Celery a zapísa ...	Sub-task	Done	Dung Lam Tuan
	MON-260	Zapísanie zaiatku extrakcie do CÚ	Sub-task	In review	Dung Lam Tuan
	MON-261	Zapísanie ukonenia extrakcie do CÚ	Sub-task	Blocked	Patrik Židuliak
	MON-262	Zapísanie potu eventov do CÚ	Sub-task	To Do	
	MON-263	Pridanie extraction_id do articles	Sub-task	Done	Peter Lipták
	MON-310	Zdokerizova flower a exposnú ho	Sub-task	Done	Peter Lipták
	Epic Link:	Manažment web monitoringu			
Sprint:	Hruška, Ipa				

Description	
--------------------	--

Informacie o extrakcii budu odosielane na centralne ulozisko s vyuzitim endpointov (vytvorenych v ramci [MON-134](#)).

Informacie o extrakcii sa budu posielat zo schedulera v nasledujucich situaciach:

- pri zaciatku extrakcie - este pred vlozenim tasku do Celery (/extractions, POST) - nastavi sa zaciatok extrakcie a stav extrakcie sa nastavi na "running"
- pri kazdom spusteni schedulera (/extractions/{extraction_id}, PUT) - aktualizuju sa statistiky o pocte subtaskov v jednotlivych stavoch (done, pending, running, failed)
- pri konci extrakcie, t.j. ked uz nebudu existovat ziadne tasky, ktore maju zadany extraction_id (/extractions/{extraction_id}, PUT) - nastavi sa cas konca extrakcie, finalny stav extrakcie (uspesna, neuspesna)

Hlavny data provider ako aj vsetky vytvorene subtasky budu zdielat to iste extraction_id.

Treba premysliet, ako sa budu kontrolovat beziace extrakcie (zavisi to od moznosti API poskytovanej Celery). Ako najvhodnejšia moznost sa javi vypytat si z centralneho uloziska zoznam ID extrakcii, ktore aktualne bezia, cez tento zoznam preiterovat a vypytat si zoznam Celery taskov s danym extraction ID ako argumentom.

Do kazdeho data providera treba pridať event "start" a "end". Pri odchytení eventu end, hlavný data provider pošle počet výskytov jednotlivých eventov (je to zoznam eventov len z neho, nie z jeho detí). Tento počet sa bude zbierať cez špeciálny listener, ktorý bude robiť len inkrementáciu.

Navše treba rozšíriť tabuľku article o FK "extraction_id", aby sme mali uložené aj informáciu o tom, v ktorej extrakcii sa vyextrahovali konkrétne články.

[MON-239] Logovanie z data providerov Created: 11/Mar/19 Updated: 25/Mar/19 Resolved: 24/Mar/19

Status:	Accepted
Project:	MonAnt
Component/s:	None
Affects Version/s:	None
Fix Version/s:	None

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

Sub-Tasks:	Key	Summary	Type	Status	Assignee
	MON-248	Vytvorenie schémy správy	Sub-task	Done	Veronika Žatková
	MON-249	Vytvorenie pomocného handlera	Sub-task	Done	Veronika Žatková
	MON-250	Posielanie správ v existujúcich data ...	Sub-task	Done	Veronika Žatková
	MON-257	Zmeni databázové logovanie	Sub-task	Done	Veronika Žatková
Epic Link:	Manažment web monitoringu				
Sprint:	Hruška				

Description	
--------------------	--

Pocas extrakcie bude mozne zalogovat spravu. Priklad logovanej spravy je odpoved 404 z URL, ktora sa mala nacitat.

Kazda sprava bude mat:

- cas, kedy bola vygenerovana
- svoj text
- level (info, warning, error)

Sprava sa bude posielat:

- hned pri jej vyskyte (/extractions/{extraction_id}/logs, POST)
- priamo z data providera (treba vytvorit nejaky helper, ktory poskytne metodu, ktorej sa ako argumenty posle text spravy a jej level).

V ramci existujucich data providerov doplnte logovanie tam, kde sa Vam bude zdať uzitocne.

[MON-240] Management - dummy story Created: 11/Mar/19 Updated: 30/Mar/19 Resolved: 12/Mar/19

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

Type:	Story	Priority:	Medium
Reporter:	Veronika Žatková	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes	Remaining Estimate:	Not Specified
Time Spent:	3 days, 25 minutes	Time Spent:	Not Specified
Original Estimate:	Not Specified	Original Estimate:	Not Specified

Sub-Tasks:

Key	Summary	Type	Status	Assignee
MON-241	Pridanie logu hodin do retrospektivy ...	Sub-task	Done	Milan Cák
MON-242	Pridanie logu hodin do retrospektivy ...	Sub-task	Done	Milan Cák
MON-243	Export uloh z Jira sprintu Gin (vytvo...	Sub-task	Done	Milan Cák
MON-244	Doplňit tabuľky o taskoch v retrospek...	Sub-task	Done	Milan Cák
MON-251	Nahodi dokumenty na web	Sub-task	Done	Milan Cák
MON-252	Pozrie podmienky k dokumentom odovzd...	Sub-task	Done	Milan Cák
MON-202	Ošetrovanie výnimiek parsovania	Sub-task	In Progress	Veronika Žatková
MON-256	Zmenit home page v UI	Sub-task	Done	David Tran Duc
MON-265	Prezentácia MVP	Sub-task	To Do	
MON-268	Seeder na data providers	Sub-task	To Do	Veronika Žatková
MON-270	Prepoji docker kontajnery	Sub-task	Done	Peter Lipták
MON-271	Vyplni dotaznik na TP Cup	Sub-task	Done	Milan Cák
MON-272	Posla dotaznik a prezentáciu na tpcu...	Sub-task	Done	Milan Cák
MON-273	Prípravi video	Sub-task	Done	Milan Cák
MON-274	Prípravi prezentáciu (robime.it blog)	Sub-task	Done	Milan Cák
MON-275	Zmeni nastavenie nginxu	Sub-task	Done	Peter Lipták
MON-276	Prípravi tímovú fotku	Sub-task	Done	Milan Cák
MON-279	Prida Travisa do monant-web-monitori...	Sub-task	Done	Peter Lipták
MON-280	Nahra tímové foto	Sub-task	Done	Milan Cák
MON-281	Ošetri zlyhanie migrácií na produk...	Sub-task	To Do	Peter Lipták
MON-282	Refaktor article scheme and event tri...	Sub-task	Done	Veronika Žatková
MON-283	Informova Ivana o dokonení user sto...	Sub-task	Done	Patrik Židulak
MON-284	Rozbeha pgsqľ aby fungoval na serveri	Sub-task	Done	Peter Lipták
MON-291	Sfinalizova príspevok na IIT.SRC, do...	Sub-task	Done	Milan Cák
MON-307	Prerobi a odovzda lánok na IIT.SRC	Sub-task	Done	Milan Cák
MON-308	Prípravi A1 plagát	Sub-task	In review	Milan Cák
MON-311	Pridanie logu hodin do retrospekt...	Sub-task	In Progress	Milan Cák
MON-312	Export uloh z Jira sprintu Hruška (vy...	Sub-task	In Progress	Milan Cák
MON-313	Nahodi dokumenty na web	Sub-task	In Progress	Milan Cák
MON-315	Spojazzdnenie Jiry	Sub-task	Done	Patrik Židulak

Sprint:

Hruška, Ipa