

Export úloh 03. 04. 2022

Obsah

TO DO.....	9
[AMS-205] Štatistika domén, ktoré vracia Google News - koľko ich je, početnosť vrátenia článkov Created: 31/Mar/22 Updated: 31/Mar/22	9
[AMS-203] Náhľad článku Created: 24/Mar/22 Updated: 24/Mar/22.....	10
[AMS-202] Zobrazenie archivovaného článku Created: 24/Mar/22 Updated: 28/Mar/22 ...	11
[AMS-201] Spustenie scraperu na veľá dátach Created: 24/Mar/22 Updated: 24/Mar/22 ...	12
[AMS-198] Frontend keywords Created: 24/Mar/22 Updated: 24/Mar/22	13
[AMS-197] Otestovať filtrovanie Created: 24/Mar/22 Updated: 24/Mar/22.....	14
[AMS-192] Profil používateľa Created: 24/Mar/22 Updated: 24/Mar/22	15
[AMS-191] Update user account data Created: 24/Mar/22 Updated: 24/Mar/22	16
[AMS-190] Zobrazit' iba doménu stránky vo výsledkoch vyhľadávania Created: 20/Mar/22 Updated: 24/Mar/22	17
[AMS-183] Menu pre každý zobrazený článok Created: 10/Mar/22 Updated: 24/Mar/22 ..	18
[AMS-177] Boolean query implementácia Created: 03/Mar/22 Updated: 03/Mar/22	19
[AMS-170] Google analytics Created: 24/Feb/22 Updated: 24/Feb/22.....	20
[AMS-168] Obmedzenie API pre tretiu stranu Created: 24/Feb/22 Updated: 24/Feb/22.....	21
[AMS-166] Prevencia zahltienia API Created: 24/Feb/22 Updated: 24/Feb/22	22
[AMS-162] PDF report zdieľateľný linkom Created: 24/Feb/22 Updated: 24/Feb/22	23
[AMS-159] Dokumentácia - metodika definition of done Created: 19/Feb/22 Updated: 20/Feb/22.....	24
[AMS-158] Dokumentácia - metodika definition of ready Created: 19/Feb/22 Updated: 20/Feb/22.....	25
[AMS-151] Frontend registrácia Created: 19/Feb/22 Updated: 24/Mar/22.....	26
[AMS-112] Fuzzy search Created: 23/Nov/21 Updated: 20/Feb/22.....	27
[AMS-86] PDF report z nájdených článkov Created: 02/Nov/21 Updated: 30/Mar/22	28
[AMS-85] Prihlásenie sa na odber nejakého vyhľadaného človeka Created: 02/Nov/21 Updated: 20/Feb/22	29
[AMS-83] Hodnotenie risku pri nájdených článkoch Created: 02/Nov/21 Updated: 02/Nov/21	30
[AMS-82] Zvýraznenie hľadaných slovíčok Created: 02/Nov/21 Updated: 02/Nov/21	31
↳[AMS-80] geografická lokácia Created: 02/Nov/21 Updated: 02/Nov/21	32
↳[AMS-79] vyhľadanie na základe zločinov Created: 02/Nov/21 Updated: 17/Mar/22	33

↳[AMS-78] obmedzenie vyhľadavania na základe dátumu vydania článkov Created: 02/Nov/21 Updated: 02/Nov/21	34
[AMS-77] Filter na web stránke Created: 02/Nov/21 Updated: 24/Mar/22	35
↳[AMS-76] Iné možnosti prekladu článkov Created: 02/Nov/21 Updated: 02/Nov/21	36
[AMS-68] Analýza médií, ktoré dostaneme vo výstupoch a aké search queries treba poslať, aby sme mali výsledky ako z Google News Created: 02/Nov/21 Updated: 31/Mar/22	37
[AMS-67] Pozrieť sa na zahodené články, kt. neobsahovali paragrafy Created: 02/Nov/21 Updated: 20/Feb/22	38
[AMS-42] Porovnanie výsledkov nášho vyhľadavania s výsledkami Google News Created: 13/Oct/21 Updated: 20/Feb/22	39
[AMS-41] Vyriešiť, čo s článkami, ktoré budú v pôvodnom zdroji nekompletné Created: 13/Oct/21 Updated: 20/Feb/22	40
↳[AMS-38] Zistiť, ako funguje Google Translate API, ako ho vieme použiť Created: 13/Oct/21 Updated: 23/Nov/21	41
[AMS-37] Zistiť, ako fungujú iné zdroje Google News, prípadne aký mechanizmus používa Google News. Created: 13/Oct/21 Updated: 23/Nov/21	42
[AMS-33] Zistiť, čo spája kriminálne články Created: 13/Oct/21 Updated: 20/Feb/22	43
[AMS-19] Analyza - multilingvistika Created: 05/Oct/21 Updated: 27/Oct/21	44
[AMS-12] Nasu API dat na portal verejnych API v ramci propagacie Created: 05/Oct/21 Updated: 27/Oct/21	45
IN PROGRESS	46
[AMS-204] Refactor Flask server Created: 25/Mar/22 Updated: 02/Apr/22	46
[AMS-199] Sidebar Created: 24/Mar/22 Updated: 30/Mar/22	47
[AMS-186] Implementácia integračného testu Created: 10/Mar/22 Updated: 24/Mar/22 ...	48
[AMS-185] Komunikácia medzi frontendom a web serverom - poslať autentifikačné a registračné údaje Created: 10/Mar/22 Updated: 24/Mar/22	49
[AMS-182] Dokumentácia web frontendu Created: 10/Mar/22 Updated: 24/Mar/22	50
[AMS-178] Request na MongoDB, ktorý vráti články na základe keywords Created: 03/Mar/22 Updated: 02/Apr/22	51
[AMS-138] Dokumentácia User guide Created: 14/Feb/22 Updated: 31/Mar/22	53
[AMS-137] Dokumentácia elasticsearch Created: 14/Feb/22 Updated: 31/Mar/22	54
[AMS-136] Dokumentácia fungovania NGINX Created: 14/Feb/22 Updated: 31/Mar/22 ..	55
[AMS-135] Dokumentácia fungovania dockeru Created: 14/Feb/22 Updated: 31/Mar/22 ..	56
[AMS-134] Dokumentácia node serveru Created: 14/Feb/22 Updated: 31/Mar/22	57
[AMS-36] User databáza + zbieranie aktivity používateľa Created: 13/Oct/21 Updated: 24/Mar/22	58
DONE	61

[AMS-189] Prerobiť mongodb Created: 17/Mar/22 Updated: 19/Mar/22 Resolved: 19/Mar/22	61
[AMS-181] Integrácia nového stroja do tímového prostredia Created: 10/Mar/22 Updated: 24/Mar/22 Resolved: 24/Mar/22	62
[AMS-180] Github actions na repo team web page Created: 10/Mar/22 Updated: 19/Mar/22 Resolved: 19/Mar/22	63
[AMS-179] Github repo pre mensie sucasti projektu Created: 06/Mar/22 Updated: 20/Mar/22 Resolved: 20/Mar/22	64
[AMS-175] Vyhľadať v scrapnutom článku všetky keywords Created: 03/Mar/22 Updated: 24/Mar/22 Resolved: 24/Mar/22	65
↳ [AMS-173] ťahanie dát Created: 24/Feb/22 Updated: 03/Mar/22 Resolved: 03/Mar/22	71
[AMS-172] Implementácia developerského prostredia Created: 24/Feb/22 Updated: 03/Mar/22 Resolved: 03/Mar/22	72
[AMS-171] Implementácia routera na frontende Created: 24/Feb/22 Updated: 11/Mar/22 Resolved: 11/Mar/22	74
[AMS-157] Pridať na stránku tímu priestor na dokumenty z LS Created: 19/Feb/22 Updated: 03/Mar/22 Resolved: 03/Mar/22	77
[AMS-156] Preklad keywordov do viacerých jazykov Created: 19/Feb/22 Updated: 24/Mar/22 Resolved: 24/Mar/22	79
[AMS-153] Named entity recognition Created: 19/Feb/22 Updated: 11/Mar/22 Resolved: 11/Mar/22	80
[AMS-149] Frontend obrazovka prihlásenia Created: 19/Feb/22 Updated: 11/Mar/22 Resolved: 11/Mar/22	82
[AMS-148] Update indexovacieho mechanizmu Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22	84
[AMS-147] Refaktoring MongoDB Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22	86
[AMS-146] Stránkovanie článkov na frontende Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22	88
[AMS-145] Stránkovanie API Created: 19/Feb/22 Updated: 10/Mar/22 Resolved: 10/Mar/22	90
[AMS-144] Autentifikácia usera Created: 19/Feb/22 Updated: 24/Mar/22 Resolved: 24/Mar/22	92
[AMS-143] Komunikácia medzi aplikačným serverom a user databázou Created: 19/Feb/22 Updated: 21/Mar/22 Resolved: 21/Mar/22	93
[AMS-142] Základná user tabuľka v user databáze Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22	94
[AMS-141] Password management Created: 14/Feb/22 Updated: 19/Feb/22 Resolved: 19/Feb/22	96

[AMS-140] Zabezpečenie endpointov cez HTTP basic auth Created: 14/Feb/22 Updated: 09/Mar/22 Resolved: 07/Mar/22	97
[AMS-139] Verziovanie docker compose Created: 14/Feb/22 Updated: 24/Feb/22 Resolved: 24/Feb/22	98
[AMS-133] Mapa domény na stránku tímu Created: 14/Feb/22 Updated: 24/Feb/22 Resolved: 24/Feb/22	99
[AMS-132] MongoDB unit test Created: 07/Dec/21 Updated: 14/Dec/21 Resolved: 14/Dec/21	100
[AMS-131] Testovanie dostupnosti kontajnerov Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21.....	101
[AMS-130] Konfigurácia NGINX Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21	102
[AMS-129] [RP] Exporty úloh Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21	103
[AMS-128] [RP] Globálna retrospektíva Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21.....	104
[AMS-127] [RP] Šprint č. 5 Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21	105
[AMS-126] [RP] Metodika docker Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21	106
[AMS-125] [RP] Metodika github actions Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21.....	107
[AMS-124] [ID] Sekv. diagram celkového fungovania Created: 07/Dec/21 Updated: 08/Dec/21 Resolved: 08/Dec/21.....	108
[AMS-123] [ID] Modul aplikačný server + testovanie Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21.....	109
[AMS-122] [ID] Modul frontend + testovanie Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21.....	110
[AMS-121] [ID] Modul na indexovanie + testovanie Created: 07/Dec/21 Updated: 10/Dec/21 Resolved: 10/Dec/21.....	111
[AMS-120] [ID] Modul API server + testovanie Created: 07/Dec/21 Updated: 14/Dec/21 Resolved: 14/Dec/21.....	112
[AMS-119] [ID] Modul scraper + testovanie Created: 07/Dec/21 Updated: 13/Dec/21 Resolved: 13/Dec/21.....	113
[AMS-118] Rozbehnúť PostgreSQL Created: 23/Nov/21 Updated: 28/Nov/21 Resolved: 28/Nov/21	114
[AMS-117] Komunikácia medzi aplikačným serverom a API serverom Created: 23/Nov/21 Updated: 07/Dec/21 Resolved: 07/Dec/21	115
[AMS-116] Komunikácia medzi API serverom a Elastic Search Created: 23/Nov/21 Updated: 06/Dec/21 Resolved: 06/Dec/21	117

[AMS-115] Zobrazenie výsledkov vyhľadávania Created: 23/Nov/21 Updated: 05/Dec/21 Resolved: 05/Dec/21.....	118
[AMS-114] Komunikácia medzi MongoDB a API serverom Created: 23/Nov/21 Updated: 07/Dec/21 Resolved: 07/Dec/21.....	119
[AMS-113] Indexovanie stiahnutých dát Created: 23/Nov/21 Updated: 29/Nov/21 Resolved: 29/Nov/21.....	120
[AMS-111] Dokumentácia k riadeniu projektu Created: 23/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21.....	122
[AMS-110] Dokumentácia k inžinierskemu dielu Created: 23/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21.....	123
↳ [AMS-107] Vytvorenie articles db na virtualnej masine Created: 21/Nov/21 Updated: 21/Nov/21 Resolved: 21/Nov/21.....	124
↳ [AMS-106] Analyza shard, nodes Created: 16/Nov/21 Updated: 21/Nov/21 Resolved: 21/Nov/21.....	125
↳ [AMS-105] Analyza River plugin Created: 16/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21.....	126
[AMS-103] Rozšírenie docker compose o MongoDB a Elasticsearch Created: 15/Nov/21 Updated: 21/Nov/21 Resolved: 21/Nov/21.....	127
[AMS-102] Deployment scrapera Created: 15/Nov/21 Updated: 29/Nov/21 Resolved: 28/Nov/21.....	129
↳ [AMS-101] vytvorit komprimovanu collection z main.py a zmerat cas indexovania Created: 12/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21.....	131
↳ [AMS-100] Vytvorit config pre articles_index Created: 12/Nov/21 Updated: 12/Nov/21 Resolved: 12/Nov/21.....	132
↳ [AMS-99] Skonzultovat komentý v storke Created: 12/Nov/21 Updated: 15/Nov/21 Resolved: 15/Nov/21.....	133
↳ [AMS-98] Napísať dokumentáciu do README.md Created: 12/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21.....	134
↳ [AMS-97] Vytvorenie docker-compose file s MongoDB, Elasticsearch a Kibana Created: 12/Nov/21 Updated: 12/Nov/21 Resolved: 12/Nov/21.....	135
[AMS-95] Implementácia dummy Flask API Created: 09/Nov/21 Updated: 15/Nov/21 Resolved: 15/Nov/21.....	136
[AMS-94] Návrh Flask API Created: 09/Nov/21 Updated: 26/Feb/22 Resolved: 15/Nov/21.....	137
[AMS-93] Analýza možností implementácie testovania Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21.....	145
[AMS-92] Zabezpečenie komunikácie medzi klientom a serverom Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21.....	148

[AMS-91] Úprava dokumentácie zo stretnutí a šprintov tak, aby mohli ísť na stránku tímu Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21	149
[AMS-90] Integrácia MongoDB a Elasticsearch Created: 09/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21.....	150
[AMS-89] Integrácia scraperu a vylepšeného parseru Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 22/Nov/21.....	152
[AMS-88] Rozšíriť úložisko pomocou nepriradeného disku Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 14/Nov/21.....	154
[AMS-87] Zabrániť zacykleným buildom Created: 07/Nov/21 Updated: 22/Nov/21 Resolved: 15/Nov/21.....	155
↳[AMS-74] Sekvenčné scrapovanie Created: 02/Nov/21 Updated: 08/Nov/21 Resolved: 06/Nov/21	156
[AMS-73] Vyhľadanie RSS dvojčat'a k HTML článku - všeobecný postup Created: 02/Nov/21 Updated: 23/Nov/21 Resolved: 09/Nov/21	158
[AMS-72] Technický návrh developerského prostredia Created: 02/Nov/21 Updated: 25/Feb/22 Resolved: 24/Feb/22	159
[AMS-71] Napĺňanie MongoDB priamo pri scrapovaní Created: 02/Nov/21 Updated: 23/Nov/21 Resolved: 14/Nov/21.....	160
[AMS-70] Rozšírenie google news scrapperu na lokáciu UK Created: 02/Nov/21 Updated: 02/Nov/21 Resolved: 02/Nov/21.....	161
↳[AMS-66] Ako mongoDB pracuje s vopred komprimovaným textom Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21	162
↳[AMS-65] vyskusat nad datach Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21	163
↳[AMS-64] realne kniznice Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21	165
↳[AMS-63] elastic search a komprimacia Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21.....	166
↳[AMS-62] analyza vyhladavania nad komprimovaným textom Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21	167
↳[AMS-61] bezstratova kompresia text Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21.....	168
[AMS-60] Pamäťová zložitosť elastic search na vzorke dát Created: 01/Nov/21 Updated: 12/Nov/21 Resolved: 12/Nov/21.....	169
[AMS-59] Parsovanie tagov pomocou lxml, readability, trafiletura Created: 30/Oct/21 Updated: 09/Nov/21 Resolved: 09/Nov/21	171
[AMS-57] Zavesenie klienta Created: 27/Oct/21 Updated: 07/Nov/21 Resolved: 07/Nov/21	174

↳[AMS-56] NGINX konfigurácia Created: 27/Oct/21 Updated: 31/Oct/21 Resolved: 31/Oct/21	175
↳[AMS-55] Kontajner Express server Created: 27/Oct/21 Updated: 31/Oct/21 Resolved: 31/Oct/21	176
↳[AMS-54] Kontajner Flask API Created: 27/Oct/21 Updated: 30/Oct/21 Resolved: 30/Oct/21	177
[AMS-53] Rozšíriť docker compose Created: 27/Oct/21 Updated: 31/Oct/21 Resolved: 31/Oct/21	178
[AMS-51] Pridať info o projekte a opísať členov tímu Created: 27/Oct/21 Updated: 02/Nov/21 Resolved: 01/Nov/21	180
[AMS-50] Request URL s odpoveďou inou ako 200 uložiť do súboru Created: 27/Oct/21 Updated: 23/Nov/21 Resolved: 01/Nov/21	181
[AMS-49] Lokálna MongoDB so získanými dátami Created: 27/Oct/21 Updated: 23/Nov/21 Resolved: 01/Nov/21	182
[AMS-48] Analýza bezstratovej kompresie textu Created: 27/Oct/21 Updated: 01/Nov/21 Resolved: 01/Nov/21	185
[AMS-47] Parsovanie tagov pomocou regex Created: 27/Oct/21 Updated: 02/Nov/21 Resolved: 02/Nov/21	187
[AMS-46] Parsovanie tagov pomocou CSS Selector Created: 27/Oct/21 Updated: 08/Nov/21 Resolved: 08/Nov/21	188
[AMS-44] Stránka tímu Created: 18/Oct/21 Updated: 22/Nov/21 Resolved: 01/Nov/21	190
[AMS-43] Otvorenie portov na serveri Created: 18/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21	191
[AMS-39] Vyriešiť, aby fiitkar nemohol urobiť "sudo su" na virtuálnom stroji (aby nemal šancu sa zmeniť na roota) Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 19/Oct/21	192
[AMS-35] Prieskum databáz Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21	193
[AMS-34] Vytvoriť github projekt Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21	194
[AMS-32] Prihláška na TP CUP Created: 13/Oct/21 Updated: 02/Nov/21 Resolved: 02/Nov/21	195
[AMS-31] Set up elastic search v docker container Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 16/Oct/21	196
↳[AMS-30] Redukcia a kategorizácia zoznamu zločínov Created: 13/Oct/21 Updated: 23/Nov/21 Resolved: 23/Nov/21	198
[AMS-29] Získanie dát pre prototyp Created: 13/Oct/21 Updated: 23/Nov/21 Resolved: 26/Oct/21	199
[AMS-21] Vytvoriť Slack server Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 11/Oct/21	201

[AMS-20] Analyza parametrov Google News Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 13/Oct/21	202
↳ [AMS-17] API / scrapovanie Created: 05/Oct/21 Updated: 10/Oct/21 Resolved: 10/Oct/21	203
↳ [AMS-16] Ziskanie vzorky RSS Created: 05/Oct/21 Updated: 08/Oct/21 Resolved: 08/Oct/21	204
[AMS-15] Realny scrapping - vytvorenie vzorky Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 10/Oct/21	206
[AMS-13] Zistit, ktore miestnosti su na karticky Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21	207
[AMS-11] Lofi pouzivatel'skeho rozhrania - main page Created: 05/Oct/21 Updated: 20/Feb/22 Resolved: 20/Feb/22	208
[AMS-10] Revizia poziadaviek Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 13/Oct/21	209
[AMS-9] Spísať denník z prvého stretnutia s vedúcim Created: 03/Oct/21 Updated: 22/Nov/21 Resolved: 05/Oct/21	210
[AMS-8] Pripady pouzitia Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 05/Oct/21	211
[AMS-7] Prieskum ziskavania dat, praca s kniznicou Created: 01/Oct/21 Updated: 23/Nov/21 Resolved: 05/Oct/21	212
[AMS-6] Specifikacia poziadaviek Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 03/Oct/21	213
[AMS-5] Vysoka architektura Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 03/Oct/21	214
[AMS-4] Zaobstarat stroj v škole Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 13/Oct/21	215
[AMS-3] Zoznam trestnych cinov, ktore budu pouzite ako queries na google news Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 04/Oct/21	216
[AMS-200] Prerobiť developerské prostredie Created: 24/Mar/22 Updated: 29/Mar/22...	217
[AMS-193] Prepojenie servisov Created: 24/Mar/22 Updated: 26/Mar/22	218
[AMS-154] Článok IIT.SRC Created: 19/Feb/22 Updated: 25/Mar/22	219
[AMS-184] Implementovať spôsob pamätania si usera Created: 10/Mar/22 Updated: 24/Mar/22	220

TO DO

[AMS-205] [Štatistika domén, ktoré vracia Google News - koľko ich je, početnosť vrátenia článkov](#) Created: 31/Mar/22 Updated: 31/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0016e:

[AMS-203] [Náhl'ad článku](#) Created: 24/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zw

[AMS-202] [Zobrazenie archivovaného článku](#) Created: 24/Mar/22 Updated: 28/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0 i00166:

[AMS-201] [Spustenie scraperu na veľ'a dátach](#) Created: 24/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0 i0015y:

[AMS-198] [Frontend keywords](#) Created: 24/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0ji0014u:

Description

skonzultovat s Davidom

[AMS-197] [Otestovať filtrovanie](#) Created: 24/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Issue links:	Blocks		
	is blocked by	AMS-178	Request na MongoDB, ktorý vráti článk... IN PROGRESS
Story point estimate:	8		
Sprint:	AMS Sprint 8		
Rank:	0 i0014m:		

Description

Podľa keywords, krajiny, dátumu

Obsah response koresponduje s requestom?

[AMS-192] [Profil používateľa](#) Created: 24/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Issue links:	Blocks
	blocks AMS-191 Update user account data To Do
Sprint:	
Rank:	0ji0013a:

Description

Predpripraviť na zmenu použ. údajov a pdf reporty história, prípadne ich generovanie

[AMS-191] [Update user account data](#) Created: 24/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Issue links:	Blocks
	is blocked by AMS-192 Profil používateľa To Do
Sprint:	
Rank:	0ji00132:

Description

full stack

frontend napojiť na výsledok storky <https://tim8-2021.atlassian.net/browse/AMS-192>

[AMS-190] [Zobrazit iba doménu stránky vo výsledkoch vyhľadávania](#) Created:
20/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 8
Rank:	0ji0012u:

[AMS-183] [Menu pre každý zobrazený článok](#) Created: 10/Mar/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0j00112:

[AMS-177] [Boolean query implementácia](#) Created: 03/Mar/22 Updated: 03/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji000zi:

[AMS-170] [Google analytics](#) Created: 24/Feb/22 Updated: 24/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000xq:

[AMS-168] [Obmedzenie API pre tretiu stranu](#) Created: 24/Feb/22 Updated: 24/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji000xa:

Description

Pre tretiu stranu vráti API max 10 výsledkov.

Po zaplattení vráti neobmedzené množstvo.

[AMS-166] [Prevencia zahltenia API](#) Created: 24/Feb/22 Updated: 24/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000wu:

[AMS-162] [PDF report zdieľateľný linkom](#) Created: 24/Feb/22 Updated: 24/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000vy:

Description

- Vygenerovanie url smerujúcej na pdf report
- Jednorázové heslo k url

[AMS-159] [Dokumentácia - metodika definition of done](#) Created: 19/Feb/22 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000va:



[AMS-158] [Dokumentácia - metodika definition of ready](#) Created: 19/Feb/22 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000v2:

[AMS-151] [Frontend registrácia](#) Created: 19/Feb/22 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	3
Sprint:	AMS Sprint 8
Rank:	0ji0012y:

Description

zatial staci username a password

[AMS-112] [Fuzzy search](#) Created: 23/Nov/21 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji000m7:

[AMS-86] [PDF report z nájdených článkov](#) Created: 02/Nov/21 Updated: 30/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	2 hours		
Original estimate:	Not Specified		

Story point estimate:	13
Sprint:	AMS Sprint 8
Rank:	0ji0015i:

Description

používateľ vyberie, z ktorých článkov chce vytvoriť report

bude na tom robiť aj Jakub

[AMS-85] [Prihlásenie sa na odber nejakého vyhľadaneého človeka](#) Created:
02/Nov/21 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000hj:

Description

vždy keď sa pridá nový článok o hľadanom človeku

Continuous monitoring

[AMS-83] [Hodnotenie risku pri nájdenných článkoch](#) Created: 02/Nov/21 Updated: 02/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji000h3:

Description

Používateľ si ohodnotí, aký risk má daný článok + nejaká poznámka k tomu

[AMS-82] [Zvýraznenie hľadaných slovíčok](#) Created: 02/Nov/21 Updated: 02/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000gv:

0

 **[AMS-80] [geografická lokácia](#)** Created: 02/Nov/21 Updated: 02/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 8
Rank:	0 i000gf:

0

 **[AMS-79] [vyhl'adanie na základe zločinov](#)** Created: 02/Nov/21 Updated: 17/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 8
Rank:	0 i000g7:

Description

pridať do requestu nový parameter keywords=[a, b, ...]

jazyk keywordov sa zmení pri zmene regiónu

keywordy môžu zostať aj v anglickom jazyku (murder / vražda)

0

 **[AMS-78] obmedzenie vyhľadavania na základe dátumu vydania článkov** Created:

02/Nov/21 Updated: 02/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 8
Rank:	0ji000fz:

[AMS-77] [Filter na web stránke](#) Created: 02/Nov/21 Updated: 24/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
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

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-78	obmedzenie vyhľadávania na základe dá...	Subtask	To Do	
	AMS-79	vyhľadanie na základe zločinov	Subtask	To Do	
	AMS-80	geografická lokácia	Subtask	To Do	
Story point estimate:	5				
Sprint:	AMS Sprint 8				
Rank:	0 i00130:				

0

 **[AMS-76] Iné možnosti prekladu článkov** Created: 02/Nov/21 Updated: 02/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000fj:

Description

Z článku v ľubovoľnom jazyku dostať článok v ľubovoľnom jazyku.

[AMS-68] [Analýza médií, ktoré dostaneme vo výstupoch a aké search queries treba poslať, aby sme mali výsledky ako z Google News](#) Created: 02/Nov/21 Updated: 31/Mar/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000dr:



[AMS-67] [Pozriet' sa na zahodené články, kt. neobsahovali paragrafy](#) Created:
02/Nov/21 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000dj:

[AMS-42] [Porovnanie výsledkov nášho vyhľadávania s výsledkami Google News](#)

Created: 13/Oct/21 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0008n:

[AMS-41] Vyriešiť, čo s článkami, ktoré budú v pôvodnom zdroji nekompletné

Created: 13/Oct/21 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0008f:

Description

Články môžu byť vymazané, prípadne iba ich obrázky. Návrhy: vytvoriť pdf, screenshot, ...

0

[AMS-38] Zistiť, ako funuguje Google Translate API, ako ho vieme použiť

Created: 13/Oct/21 Updated: 23/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i00063:

Comments

Comment by [Adam Šípka](#) [23/Nov/21]

Google Translate API je spoplatnené od istého množstva preložených znakov, avšak existujú aj varianty, ktoré sú zadarmo a majú dostatočnú podporu jazykov (t. j. všetky, alebo väčšina jazykov štátov Európskej únie)

<https://english.api.rakuten.net/hongson890/api/translate-api/details>

<https://english.api.rakuten.net/zakutynsky/api/YandexTranslate>

[AMS-37] [Zistiť, ako fungujú iné zdroje Google News, prípadne aký mechanizmus používa Google News.](#) Created: 13/Oct/21 Updated: 23/Nov/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zx

Comments

Comment by [Adam Šípka](#) [23/Nov/21]

Väčšina iných zdrojov je buď obmedzená na určitý počet requestov za deň, alebo je spoplatnená. Prípadne sa jedná o príliš špecifické zdroje, ako napríklad New York Times, čiže by sme mali jednostranne zamerané dáta.

[AMS-33] [Zistiť, čo spája kriminálne články](#) Created: 13/Oct/21 Updated: 20/Feb/22

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0006f:

Description

Napríklad aké frázy a slová sa nachádzajú v dokumentoch o podozrivých osobách a v iných článkoch nie (slovo "commit").

[AMS-19] [Analyza - multilingvistika](#) Created: 05/Oct/21 Updated: 27/Oct/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	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

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-38	Zistiť, ako funguje Google Translate...	Subtask	To Do	Adam Šípka
	AMS-76	Iné možnosti prekladu článkov	Subtask	To Do	
Sprint:					
Rank:	0 i0001y:zz				

Description

API na Google translate, preklad klucovych slov

[AMS-12] [Nasu API dat na portal verejnych API v ramci propagacie](#) Created:
05/Oct/21 Updated: 27/Oct/21

Status:	To Do
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0001z:

IN PROGRESS

[AMS-204] Refactor Flask server Created: 25/Mar/22 Updated: 02/Apr/22	
Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Dominik Horvath	Assignee:	Adam Šípka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 8
Rank:	0 i000pj:i

Description

Connection na mongoDB je vytvoreny vo vlastnom subore prostrednictvom vlastneho objektu, co je OK. Connection na elastic a vsetka logika s elastikom sa ale deje v search_routes.py, co je subor urceny na definovanie API routov.. Bolo by potrebne elastic logiku oddelit od routovacieho kodu.

[AMS-199] [Sidebar](#) Created: 24/Mar/22 Updated: 30/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	3
Sprint:	AMS Sprint 8
Rank:	0ji000qa:

Description

navigacia pre prihlaseneho usera

[AMS-186] [Implementácia integračného testu](#) Created: 10/Mar/22 Updated: 24/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Jakub Müller
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	8
Sprint:	AMS Sprint 8
Rank:	0 i000pj:

[AMS-185] [Komunikácia medzi frontendom a web serverom - poslať autentifikačné a registračné údaje](#) Created: 10/Mar/22 Updated: 24/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	3 hours		
Original estimate:	Not Specified		

Story point estimate:	3
Sprint:	AMS Sprint 8
Rank:	0 j0012s:r

Comments

Comment by [Táňa Poláková](#) [16/Mar/22]

Formik vracia dáta v json formáte, takže sú rovno poslané na server bez ďalších úprav. Omylom som si loggla o 2h naviac, takže skutočný čas na storke je hodina 😊

[AMS-182] [Dokumentácia web frontendu](#) Created: 10/Mar/22 Updated: 24/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	2 hours		
Original estimate:	Not Specified		

Story point estimate:	3
Sprint:	AMS Sprint 8
Rank:	0 i0012t:

Description

+ struktura kodu

Comments

Comment by [Táňa Poláková](#) [16/Mar/22]

Na stránke dokumentácie som pridala dokumentáciu k MUI a user managementu.

[AMS-178] [Request na MongoDB, ktorý vráti články na základe keywords](#) Created: 03/Mar/22 Updated: 02/Apr/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day		
Original estimate:	Not Specified		

Issue links:	Blocks
	blocks AMS-197 Otestovať filtrovanie To Do
Story point estimate:	8
Sprint:	AMS Sprint 8
Rank:	0ji000pk:

Description

Z aplikačného serveru príde request s novým parametrom keywords=[a, b, ...] a treba vytvoriť nový request na MongoDB, ktorý vráti iba články, ktoré tieto slová obsahujú.

flask dostane anglicke keywords a flask vráti clanky s ang keyword, jedine, co sa meni, je region.

Cize napr. clanok zo spanielska s anglickymi keywords

Comments

Comment by [Adam Šípka](#) [22/Mar/22]

nemôžem zatiaľ dokončiť, ale zmenil som OR operátor na AND 😊

Comment by [Adam Šípka](#) [02/Apr/22]

Keywords musia vyzerat' presne tak, ako v Elasticu (*keywords=[Arson, Domestic Violence, ...]*), čiže prvé písmeno musí byť veľké, ináč filtrovanie nefunguje.

Taktiež som pridal aj filter na regióny, tiež by to mal byť zoznam: *regions=[gb, sk, hu, ...]*, možno ešte úplne nefunguje, keďže všetky články sú z Británie.

V response už nie je html článkov, na to bude potom iná funkcia.

[AMS-138] [Dokumentácia User guide](#) Created: 14/Feb/22 Updated: 31/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zv1

Description

user dokumentácia

[AMS-137] [Dokumentácia elasticsearch](#) Created: 14/Feb/22 Updated: 31/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zv2

[AMS-136] [Dokumentácia fungovania NGINX](#) Created: 14/Feb/22 Updated: 31/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zv4

[AMS-135] [Dokumentácia fungovania dockeru](#) Created: 14/Feb/22 Updated: 31/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zv9

[AMS-134] [Dokumentácia node serveru](#) Created: 14/Feb/22 Updated: 31/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None



Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zvi

[AMS-36] [User databáza + zbieranie aktivity používateľa](#) Created: 13/Oct/21 Updated: 24/Mar/22

Status:	IN PROGRESS
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 Conceptual Model.bmp  Tables.bmp
Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0 i0015q:

Description

pridala by som aj tabulku o pdf reportoch

Comments

Comment by [Táňa Poláková](#) [17/Oct/21]



Model, ktorý veľmi na hrubo zobrazuje ako by mohla byť pointa user databázy.

Comment by [Táňa Poláková](#) [17/Oct/21]

DONE

[AMS-189] Prerobit' mongodb Created: 17/Mar/22 Updated: 19/Mar/22 Resolved: 19/Mar/22	
Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	30 minutes		
Original estimate:	Not Specified		

Story point estimate:	0.5
Sprint:	AMS Sprint 7
Rank:	0ji0012e:

Description
inicializačný skript

[AMS-181] [Integrácia nového stroja do tímového prostredia](#) Created: 10/Mar/22 Updated: 24/Mar/22 Resolved: 24/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 7 hours		
Original estimate:	Not Specified		

Story point estimate:	13
Sprint:	AMS Sprint 7
Rank:	0 i0011e:



[AMS-180] [Github actions na repo team web page](#) Created: 10/Mar/22 Updated: 19/Mar/22 Resolved: 19/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	45 minutes		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 7
Rank:	0 i0012m:



[AMS-179] [Github repo pre mensie sucasti projektu](#) Created: 06/Mar/22 Updated: 20/Mar/22 Resolved: 20/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Dominik Horvath	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	3
Sprint:	AMS Sprint 7
Rank:	0 i0012q:

Description

Nas projekt obsahuje niekoľko sucasti ktore su privelmi male na to, aby mali samostatny repozitar. Sem spadaju Dockerfiles pre mongo, postgres, elastic, ich init scripty.

Podobne sucasti pravdepodobne casom pribudnu a je potrebne aj tieto male sucasti zalohovat do spolocneho repozitara.

[AMS-175] [Vyhl'adat' v scrapnutom článku všetky keywords](#) Created: 03/Mar/22 Updated: 24/Mar/22 Resolved: 24/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 30 minutes		
Original estimate:	Not Specified		

Attachments:	image-20220313-173704.png image-20220313-173733.png image-20220315-203205.png image-20220318-152626.png image-20220318-152659.png
Issue links:	Blocks is blocked by AMS-156 Preklad keywordov do viacerých jazykov Done
Story point estimate:	8
Sprint:	AMS Sprint 7
Rank:	0 j001 i:

Description

Ako súčasť pipeline. Hľadať z toho setu keywords, z ktorého jazyka je článok.

Comments

Comment by [Jakub Hlavačka](#) [13/Mar/22]

Otázky:

- Co s viac jazyčným scrapovaním?
- Ukladať keywords iba v ENG?
- Budem musieť všetko na lower case, aby to matchlo...

Comment by [Jakub Hlavačka](#) [13/Mar/22]

For more images please see the album: "The 100 Most Beautiful Views of Prague" that has been published by the city of Prague.

[Flashtext](#) je knižnica, ktorá rýchlo vyhľadáva alebo replacuje list reťazcov v texte. Ale na prvomobrazku je vidieť termov na vyhľadavanie/replacovanie okolo 500 (My máme 100 termov).

No a na záver [tu](#) vidím, že jednoduche “in” v pythone je rýchlejšie ako regex.

Comment by [Jakub Hlavačka](#) [15/Mar/22]

Otázky na stretnutie vol.2:

- Ukladať keywords v ENG?
- Ako chceme robiť continuous monitoring (to sa do scraperu asi nezmestí asi na tu bude potrebné novú službu aj na indexovanie...)?

Navrh na storku:

- Spraviť poriadne logovanie pre celý proces od scrapnutia po zaindexovanie (keď sa niečo pokáži, tak logy rozdeliť pre pokážené indexovanie a pokážené uloženie do db)
- Vyriešiť, prečo píše logy z elasticu namiesto klasických logov zo scrapy

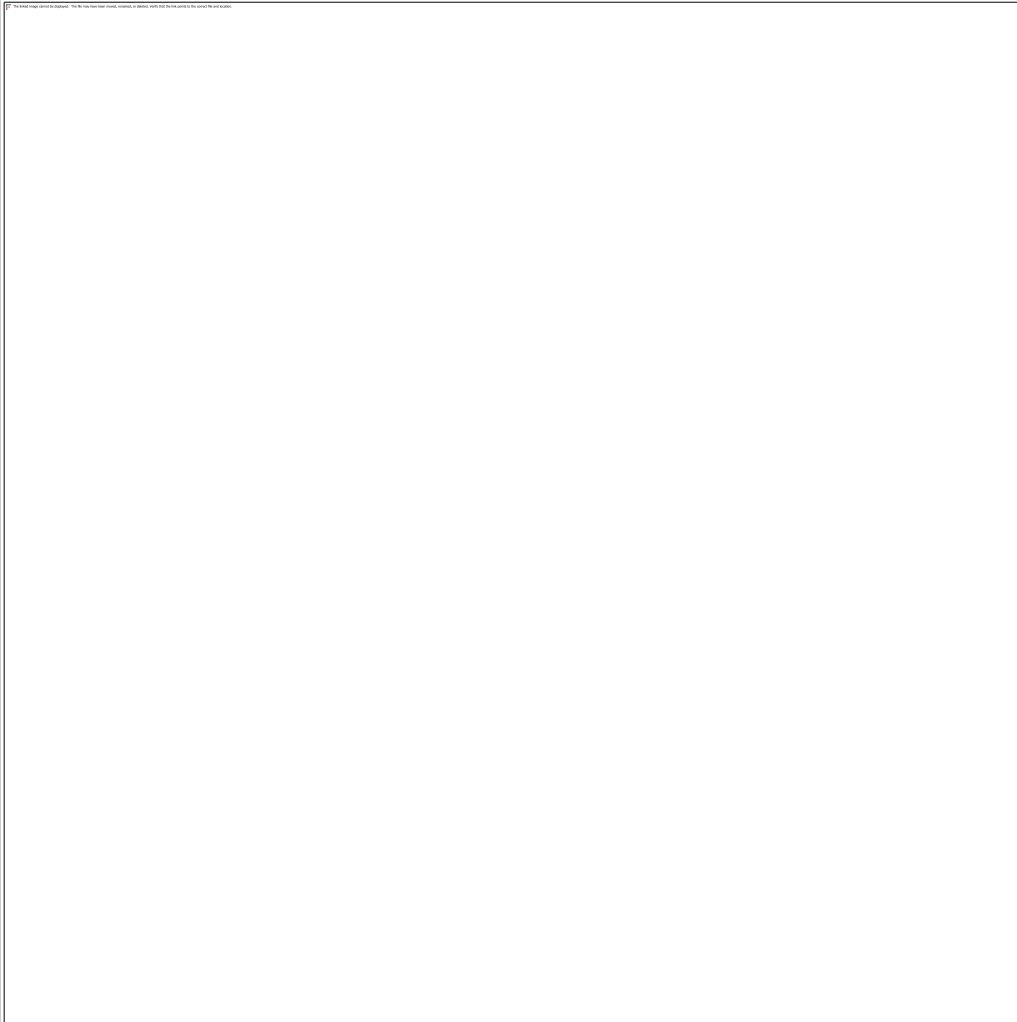
Comment by [Jakub Hlavačka](#) [15/Mar/22]

Rucne som prezrel asi 5 clankov, ci naozaj obsahuju keywords, ktore nasiel scraper a sedelo to.

Comment by [Jakub Hlavačka](#) [18/Mar/22]

Elastic:

MongoDB:



Ulozime keywords aj v jazyku, podla ktoreho sa clanky scrapovali (mozne testovanie, lepsie overenie spravnosti)

Warning: region moze byt prazdny, pretoze pre niektore krajiny existuju **locale** dopyty iba podla jazyka bez reg sa bude **vyhladavat nad elasticom**, tak je potrebne to brat do uvahy a namiesto **regionu** vtedy pouzit **language**)

Comment by [Jakub Hlavačka](#) [18/Mar/22]

TODO:

- este zmenit type of keywords in elasticsearch configuration
- update readme scraper

0

 **[AMS-173] ťahanie dát** Created: 24/Feb/22 Updated: 03/Mar/22 Resolved: 03/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 2 hours, 30 minutes		
Original estimate:	Not Specified		

Reviewer:	Dominik Horvath
Story point estimate:	3
Sprint:	AMS Sprint 6
Rank:	0 i000ye:

Comments

Comment by [Jakub Hlavačka](#) [03/Mar/22]

V buducnosti by sa mohlo zrusit reverse proxy na MongoDB a Postgres na produkcii. Pretoze sa do databazy pripaja cez ssh tunneling, tym padom by to malo fungovat bez toho

Comment by [Jakub Hlavačka](#) [03/Mar/22]

Neskor zabezpecit, aby sa data tahali iba na prvý docker-compose up.

[AMS-172] [Implementácia developerského prostredia](#) Created: 24/Feb/22 Updated:
03/Mar/22 Resolved: 03/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	0 minutes
Σ Time Spent:	2 days, 7 hours, 30 minutes	Time Spent:	1 day, 5 hours
Σ Original Estimate:	Not Specified	Original estimate:	Not Specified

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-173	ťahanie dát	Subtask	Done	Jakub Hlavačka
Reviewer:	Jakub Hlavačka				
Story point estimate:	13				
Sprint:	AMS Sprint 6				
Rank:	0 i000sy:				

Description

- wrapper nad docker compose
- prehodit' adresy
- NGINX

Comments

Comment by [Dominik Horvath](#) [26/Feb/22]

premenne prostredia boli presunute do externych konfiguracnych suborov. Bol vytvoreny specialny subor s premennymi suvisiacimi s pripojenim akehokolvek servicu na mongoDB. Tento file by mal pouzít akykolvek service pracujuci s mongoDB, aby sme mohli upraviť udaje suvisiace s databazou na jednom mieste naraz pre všetky služby.

Podobny opstup by mal byt pouzity aj pri elasticu aj pri postgresQL databaze. Subory obsahujuce parametre databazy a podobnych sluzieb by mali byt pomenovane ako <MENO>_connection.env

Za tymto ucelom bolo doposial vytvorene **elastic_connection.env** a **mongo_connection.env**

Comment by [Dominik Horvath](#) [27/Feb/22]

Prva verzia dev prostredia je hotova, dokumentacia je dostupna na https://github.com/FIIT-TEAM8/dev_environment

[AMS-171] [Implementácia routera na frontende](#) Created: 24/Feb/22 Updated: 11/Mar/22 Resolved: 11/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 hour		
Original estimate:	Not Specified		

Attachments:	image-20220304-204107.png image-20220306-111531.png
Reviewer:	Táňa Poláková
Story point estimate:	1
Sprint:	AMS Sprint 6
Rank:	0 i000te:

Comments

Comment by [David Silady](#) [04/Mar/22]

Found in branch: *basic_router*

Comment by [David Silady](#) [04/Mar/22]

Je tam aj zopár fixov na backende.

Comment by [David Silady](#) [04/Mar/22]

Hlavná časť je v:

App.jsx (router - je nested)

index.js (router wrapper - tam sa určuje podľa env premennej, že aka je home)

TitleSearch.jsx (pridaný Outlet pre router [tam, kde je outlet sa vyrenderuje child routa], tiež upravená logika pre

SearchResults.jsx (presunutý apiCall, niekoľko častí zakomentovaných, pretože hlavné paging nefunguje a chybné vyhľadávanie query parametrov)

Comment by [David Silady](#) [04/Mar/22]

Comment by [David Silady](#) [04/Mar/22]

Zakladna route teraz presmeruje na /search, kde sa nachadza komponent <TitleSearch/>

Po vyhľadani sa presmeruje na /search/results, cim sa pod <TitleSearch/> na mieste <Outlet/> vyrenderuje <SearchResults/> potom na zaklade parametrov vykona vyhľadanie.

Navigácia v react-router@6 sa používa primárne pomocou:

`<Navigate/>`
`useNavigate()`

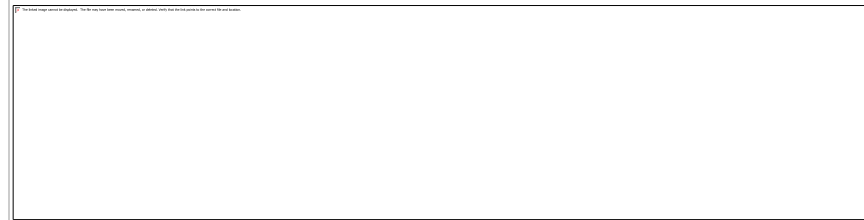
`<Route/>`

Cesty sa používajú aj relatívne.

Takže ak v TitleSearch (route: `/search`) použijem `navigate('results')`, prejdem do `/search/results`

Strom aktuálnych routov:

- `/`
- `/search`
 - `/results`
- `/example`



Comment by [Táňa Poláková](#) [11/Mar/22]

Router je implementovaný veľmi elegantne, prakticky a je jednoducho rozšíriteľný.

[AMS-157] [Pridať na stránku tímu priestor na dokumenty z LS](#) Created: 19/Feb/22 Updated: 03/Mar/22 Resolved: 03/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	2 hours		
Original estimate:	Not Specified		

Attachments:	 image-20220228-205356.png
Reviewer:	Jakub Müller
Story point estimate:	3
Sprint:	AMS Sprint 6
Rank:	0 j000uu:

Description

Dokumenty: zápis zo stretnutia a export úloh.

Návrh:

- dropdown (vybrať semester a týždeň)
- tabuľka ostane a iba sa bude voliť semester pomocou nejakého buttonu, čiže sa zmení len obsah tabuľky

Comments

Comment by [Táňa Poláková](#) [28/Feb/22]

Na stránku tímu <https://team08-21.studenti.fiit.stuba.sk/#sprintovanie> som pridala sekciu pre letný semester. Me implementované v JavaScripte a spôsobí zmenu tabuľky pre príslušný semester.

© 2022 Microsoft Corporation. All rights reserved. Microsoft, the Microsoft Dynamics logo, and the Microsoft Dynamics logo are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Comment by [Jakub Müller](#) [03/Mar/22]

Všetko ako má byť.

[AMS-156] [Preklad keywordov do viacerých jazykov](#) Created: 19/Feb/22 Updated:
24/Mar/22 Resolved: 24/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	2 hours, 10 minutes		
Original estimate:	Not Specified		

Issue links:	Blocks
	blocks AMS-175 Vyhl'adat' v scrapnutom článku všetky k... Done
Story point estimate:	5
Sprint:	AMS Sprint 7
Rank:	0 i000yx:

Description



všetky oficialne eu jazyky:

Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish and Swedish

[AMS-153] [Named entity recognition](#) Created: 19/Feb/22 Updated: 11/Mar/22 Resolved: 11/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day		
Original estimate:	Not Specified		

Attachments:	 ner_eval.xlsx  ner_testing.py
Reviewer:	Táňa Poláková
Story point estimate:	8
Sprint:	AMS Sprint 6
Rank:	0 i000t1:r

Description

nlTK python knižnica

chceme vidieť pekný výsledok

na 100 článkov

skúsiť viac jazykov

niektoré jazyky aj ručne

Comments

Comment by [Jakub Müller](#) [09/Mar/22]

Named entity recognition bolo testované na 50-tich článkoch, ktoré boli získané z našej databázy. Na identifikovanie požadovaných entít bola použitá knižnica NLTK, ktorá identifikuje

entity ako: osoba, lokácia, organizácia a pod. Bol vytvorený jednoduchý skript, ktorý identifikuje všetky typy menných entít v testovacích článkoch. Výsledok rozpoznávania entít bol vyhodnocovaný ručne a výsledky sú uvedené v prílohe **ner_eval.xlsx**. Daných 50 článkov dokázal skript spracovať za cca 17 sekúnd a je uvedený v prílohe **ner_testing.py**.


Comment by [Táňa Poláková](#) [11/Mar/22]

Storka bola pekne spracovaná, dozvedeli sme sa, akú úspešnosť má knižnica nltk.

[AMS-149] [Frontend obrazovka prihlásenia](#) Created: 19/Feb/22 Updated: 11/Mar/22 Resolved: 11/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	Not Specified		

Attachments:	 image-20220304-170103.png
Reviewer:	David Silady
Story point estimate:	5
Sprint:	AMS Sprint 6
Rank:	0 i000us:

Description

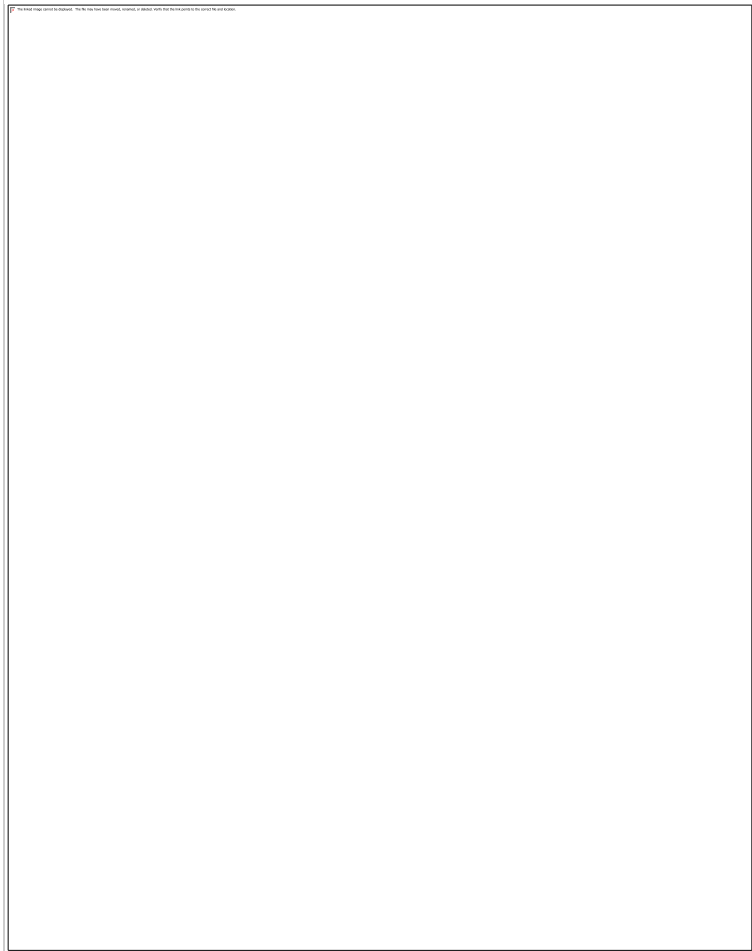
<https://www.npmjs.com/package/react-jsonschema-form>

Comments

Comment by [Táňa Poláková](#) [04/Mar/22]

Namiesto pôvodnej odporúčanej knižnice som použila <https://formik.org/docs/overview> , keďže zaručila jednoduchšiu úpravu a jednoduchšie validation handleovanie jednotlivých komponentov v rámci tagu form.

Momentálne je hotová obrazovka prihlásenia a je potrebné dorobiť komunikáciu s aplikačným serverom, ako aj vymyslieť preklik na túto obrazovku z aktuálne dostupných obrazoviek.



Comment by [David Silady](#) [04/Mar/22]

Knižnica môže byť, ale myslím, že úprava cez json by bola stále jednoduchšia, pokiaľ by sa spravil wrapper na RJSF (ja som to v minulosti robil takto nejak <https://github.com/DavidSilady/sitecreator/blob/master/src/Components/JsonSchemaWrapper.js> a šlo to docela fajn). Ale Formik vyzerá tiež dosť powerful a šikovne.

Výzorovo super.

Kód ale asi nie je na githube - neviem ho nájsť. (btw vytvoril som dnes nanovo branch pre sprint6, tak potom by som to mergoval tam a na konci sprintu sa to mergne to mainu)

[AMS-148] [Update indexovacieho mechanizmu](#) Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	7 hours, 45 minutes		
Original estimate:	Not Specified		

Reviewer:	Dominik Horvath
Story point estimate:	8
Sprint:	AMS Sprint 6
Rank:	0 i0001y:zk

Comments

Comment by [Jakub Hlavačka](#) [07/Mar/22]

Keď modul scraper uloží článok do MongoDB, tak následne pošle request Elasticu, aby ho zaindexoval.

Funguje to cez database.py nie cez Scrapy pipelines (keď tak prerobit)

TODO pred review:

- napísať dokumentáciu
- pridať env v produkčnej masíne
- zistiť, prečo Scraper vypisuje logy z elasticsearch a nie tie naše

Comment by [Jakub Hlavačka](#) [08/Mar/22]

samotný text field v elastic nie je uložený, ale zistil som, že sa v source automaticky jeho hodnota uklada.

Na linku su popisane nevyhody, ked text field v source prestaneme ukladat.

<https://www.elastic.co/guide/en/elasticsearch/reference/current/mapping-source-field.html>


Comment by [Jakub Hlavačka](#) [08/Mar/22]

Ukladaju sa uz clanky do ams_db v MongoDB alebo stale do admin ?

[AMS-147] [Refaktoring MongoDB](#) Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 hour		
Original estimate:	Not Specified		

Attachments:	 Snímka obrazovky 2022-02-26 175533.png
Reviewer:	Adam Šípka
Story point estimate:	2
Sprint:	AMS Sprint 6
Rank:	0 i0001y:zki

Description

Prichystať novovytvorenú DB - *ams*.

Prepis calls na API.

Príprava inicializačného scriptu.

Comments

Comment by [Dominik Horvath](#) [26/Feb/22]

MongoDB vytvori databazu az ked su v nej nejake data, nie je potrebne pripravit infrastrukturu dopredu, bude stacit vlozit zdrojove kody v komponentoch, ktore s databazou pracuju.

Comment by [Dominik Horvath](#) [06/Mar/22]

Custom image pre mongoDB bol vytvoreny + pushnuty na docker-hub <https://hub.docker.com/r/fiitteam8/mongo>

[AMS-146] [Stránkovanie článkov na frontende](#) Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	5 hours		

Attachments:	Screenshot 2022-03-11 161606.png image-20220311-151743.png
Reviewer:	Adam Šípka
Story point estimate:	5
Sprint:	AMS Sprint 6
Rank:	0 j0001y:zm

Description

Na jednu stranu 10 článkov.

Frontend bude posielat' request s číslom požadovanej stránky.

Comments

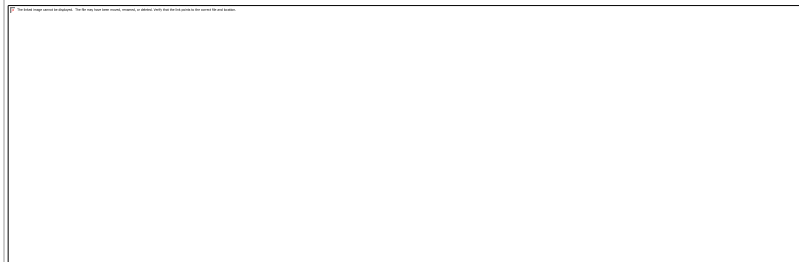
Comment by [Táňa Poláková](#) [11/Mar/22]

Pri prvom requeste pre hľadaný výraz je automaticky do parametrov pridaný parameter "page" s hodnotou 1. Následne po kliknutí na ďalšie strany je parameter upravený podľa aktuálneho čísla strany a prebehne nový request na server.

Request na server (vždy po zmenení parametrov v route):



Update aktualnej stránky:



[AMS-145] [Stránkovanie API](#) Created: 19/Feb/22 Updated: 10/Mar/22 Resolved: 10/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 2 hours		
Original estimate:	Not Specified		

Reviewer:	Dominik Horvath
Story point estimate:	13
Sprint:	AMS Sprint 6
Rank:	0 i000st:

Description

Jeden response vráti 20 článkov = jedna stránka.

Výskum, nastavenie Elasticsearch.

Request:

- číslo stránky

Do response sa pridajú fields:

- počet všetkých výsledkov
- číslo stránky
- articles per page
- počet stránok

Dokumentácia.



[AMS-144] [Autentifikácia usera](#) Created: 19/Feb/22 Updated: 24/Mar/22 Resolved: 24/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Reviewer:	Jakub Hlavačka
Story point estimate:	8
Sprint:	AMS Sprint 7
Rank:	0ji0012o:

Description

Nadväzuje na <https://tim8-2021.atlassian.net/browse/AMS-143>.

- endpoint na registráciu, login, tokeny (vytvorenie, refresh, vymazanie)
- komunikácia s user DB
- middleware

[AMS-143] [Komunikácia medzi aplikačným serverom a user databázou](#) Created: 19/Feb/22 Updated: 21/Mar/22 Resolved: 21/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Reviewer:	Jakub Hlavačka
Story point estimate:	2
Sprint:	AMS Sprint 7
Rank:	0 i0012s:

[AMS-142] [Základná user tabuľka v user databáze](#) Created: 19/Feb/22 Updated: 17/Mar/22 Resolved: 17/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	3 hours		
Original estimate:	Not Specified		

Acceptance criteria:	Splnené description požiadavky
Reviewer:	Jakub Hlavačka
Story point estimate:	5
Sprint:	AMS Sprint 6
Rank:	0 i0001y:zl

Description

- Vytvorenie PostgreSQL usera.
- Vytvorenie databázy "ams".
- Vytvorenie jednoduchej tabuľky "user": id, username, password
- Naming convention:
 - všade snake_case
 - tabuľky - lowercase, singular
 - fields - lowercase, singular
- Sprostredkovanie cez NGINX.

Comments

Comment by [David Silady](#) [28/Feb/22]

+ nahodiť heslo do KeyPassu



[AMS-141] [Password management](#) Created: 14/Feb/22 Updated: 19/Feb/22 Resolved: 19/Feb/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000ra:



[AMS-140] [Zabezpečenie endpointov cez HTTP basic auth](#) Created: 14/Feb/22 Updated: 09/Mar/22 Resolved: 07/Mar/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	30 minutes		
Original estimate:	Not Specified		

Acceptance criteria:	<ul style="list-style-type: none">• pri načítaní endpointu budú požadované credentials• implementované na všetky endpointy
Reviewer:	Dominik Horvath
Story point estimate:	3
Sprint:	AMS Sprint 6
Rank:	0 i000so:

Comments

Comment by [Dominik Horvath](#) [02/Mar/22]

vsetko OK

[AMS-139] [Verziovane docker compose](#) Created: 14/Feb/22 Updated: 24/Feb/22 Resolved: 24/Feb/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji000qu:

Description

+ github actions a deploy

service config

[AMS-133] [Mapa domény na stránku tímu](#) Created: 14/Feb/22 Updated: 24/Feb/22 Resolved: 24/Feb/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i000pi:

[AMS-132] [MongoDB unit test](#) Created: 07/Dec/21 Updated: 14/Dec/21 Resolved: 14/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	3 hours		
Original estimate:	3 hours		

Story point estimate:	3
Sprint:	AMS Sprint 5
Rank:	0 i000pa:

Comments

Comment by [Adam Šípka](#) [14/Dec/21]

Implementované dva unit testy, jeden testuje, či sa z kolekcie articles vráti správny dokument, druhý testuje, či sa z kolekcie crimemaps vracajú správne kľúčové slová.

[AMS-131] [Testovanie dostupnosti kontajnerov](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	2
Sprint:	AMS Sprint 5
Rank:	0 i000nw:

Description

Po skončení deploy sa pingnú všetky services, či nespadli.

[AMS-130] [Konfigurácia NGINX](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	30 minutes		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 5
Rank:	0 i000ob:

[AMS-129] [\[RP\] Exporty úloh](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	0.5
Sprint:	AMS Sprint 5
Rank:	0 i000p7:9

[AMS-128] [\[RP\] Globálna retrospektíva](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 5
Rank:	0 i000p8:

Comments

Comment by [Táňa Poláková](#) [15/Dec/21]

Na spoločnom stretnutí mi členovia tímu dali nápady, čo by som do tejto sekcie mohla napísať. Retrospektíva je v dokumentácii k riadeniu projektu.

[AMS-127] [\[RPI\] Šprint č. 5](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	0.5
Sprint:	AMS Sprint 5
Rank:	0 i000p7:i

Comments

Comment by [Táňa Poláková](#) [15/Dec/21]

Šprint č. 5 je pridaný do dokumentácie k riadeniu projektu.

[AMS-126] [\[RP\] Metodika docker](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 hour		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 5
Rank:	0 i000p9:i

Comments

Comment by [Adam Šípka](#) [15/Dec/21]

Zjednotené s GitHub Actions metodikou do metodiky nasadzovania.

[AMS-125] [\[RP\] Metodika github actions](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 hour		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 5
Rank:	0 i000o1:

Comments

Comment by [Adam Šípka](#) [15/Dec/21]

Zjednotené s docker metodikou do metodiky nasadzovania.

[AMS-124] [\[ID\] Sekv. diagram celkového fungovania](#) Created: 07/Dec/21 Updated: 08/Dec/21 Resolved: 08/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 hour		
Original estimate:	Not Specified		

Story point estimate:	1
Sprint:	AMS Sprint 5
Rank:	0 i000p3:

[AMS-123] [ID] Modul aplikačný server + testovanie Created: 07/Dec/21 Updated:
15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	30 minutes		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 5
Rank:	0 i000ov:

[AMS-122] [\[ID\] Modul frontend + testovanie](#) Created: 07/Dec/21 Updated: 15/Dec/21 Resolved: 15/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 5
Rank:	0 i000ox:

Comments

Comment by [Táňa Poláková](#) [15/Dec/21]

Do dokumentácie inžinierskeho diela som pridala informácie o module pre webovú aplikáciu, ktoré sa týkajú frontendu. Zhodli sme sa, že nepotrebujeme členenie na frontend a backend.

[AMS-121] [\[ID\] Modul na indexovanie + testovanie](#) Created: 07/Dec/21 Updated: 10/Dec/21 Resolved: 10/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	4 hours		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 5
Rank:	0 i000p7:

Comments

Comment by [Jakub Hlavačka](#) [10/Dec/21]

V rámci modulu indexer som vytvoril jednotkové testy na Elasticsearch a MongoDB.

Prerobil som modul Indexovanie v Inžinierskom diele a dopísal modul Elasticsearch.

Konzultácia stretnutie: Treba písať aj modul MongoDB?

[AMS-120] [\[ID\] Modul API server + testovanie](#) Created: 07/Dec/21 Updated: 14/Dec/21 Resolved: 14/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	5 hours		

Story point estimate:	5
Sprint:	AMS Sprint 5
Rank:	0 i000p9:



[AMS-119] [\[ID\] Modul scraper + testovanie](#) Created: 07/Dec/21 Updated: 13/Dec/21 Resolved: 13/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	4 hours		
Time Spent:	Not Specified		
Original estimate:	4 hours		

Story point estimate:	5
Sprint:	AMS Sprint 5
Rank:	0 i000oz:



[AMS-118] [Rozbehnúť PostgreSQL](#) Created: 23/Nov/21 Updated: 28/Nov/21 Resolved: 28/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	2 hours		
Original estimate:	Not Specified		

Story point estimate:	3
Sprint:	AMS Sprint 4
Rank:	0 i000ml:

Comments

Comment by [Jakub Hlavačka](#) [28/Nov/21]

POSTGRES_DB=users, prihlasovacie udaje su v /home/fiitkar/docker-file/postgres_db.env

[AMS-117] [Komunikácia medzi aplikačným serverom a API serverom](#) Created: 23/Nov/21 Updated: 07/Dec/21 Resolved: 07/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	4 hours, 30 minutes		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 4
Rank:	0 j000nb:

Description

Zabezpečiť komunikáciu medzi Aplikačným containerom a Flask API (data) containerom, aby aplikačný container dokázal spracovať, ukladať, hľadané requesty vzhľadom na užívateľa (napr. história vyhľadávania)

Comments

Comment by [David Silady](#) [07/Dec/21]

Zabezpečená komunikácia medzi aplikačným serverom (ExpressJS) a data API serverom (Flask)

V DEV prostredí sa zatiaľ odkazuje na produkčnú data API:

<https://team08-21.studenti.fiit.stuba.sk/api/v2>

V PROD prostredí (už je možná komunikácia v rámci docker containerov):

http://flask_server:5000/v2

Implementovaná základná funkcionálna pre endpoint:

`/ams/api/data/search/?q=""`

Aktuálne posunie request (všetky query parametre) na

```
{{  
  
{flask_api_host}  
  
/v2/search/  
  
{query_params}  
  
}}
```

Výsledok následne posunie ako response klientovi.

Poznámka:

Možno bude nutné premenovanie NGINX routes k aplikačnému webu a k data API. V aktuálnom stave to dokáže byť celkom mäťúce.

[AMS-116] [Komunikácia medzi API serverom a Elastic Search](#) Created: 23/Nov/21 Updated: 06/Dec/21 Resolved: 06/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	5 hours		
Time Spent:	3 days, 3 hours		
Original estimate:	Not Specified		

Story point estimate:	13
Sprint:	AMS Sprint 4
Rank:	0 i000mj:

Comments

Comment by [Jakub Müller](#) [06/Dec/21]

Bola zabezpečená komunikácia medzi API serverom a Elasticsearchom. Táto komunikácia prebieha tak, že po zadaní dopytu API server nadviaže komunikáciu s Elasticsearchom a pošle mu daný dopyt. Následne Elasticsearch vráti serveru "hity", kde sú uvedené ID-čka dokumentov obsahujúcich hľadaný výraz. Keďže response z Elasticsearchu je vo formáte JSON, tak musia byť ID-čka vyextrahované. Následne sú tieto ID-čka použité na získanie samotných článkov z MongoDB.

[AMS-115] [Zobrazenie výsledkov vyhľadávania](#) Created: 23/Nov/21 Updated: 05/Dec/21 Resolved: 05/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 4 hours		
Original estimate:	Not Specified		

Attachments:	 obrázok_2021-12-05_171113.png
Story point estimate:	13
Sprint:	AMS Sprint 4
Rank:	0 i000ma:

Comments

Comment by [Táňa Poláková](#) [05/Dec/21]

Na frontende sa zobrazujú výsledky vyhľadávania vo forme: dátum publikovania, link, názov článku a kľúčové slová. Prvé tri vety článku sa zatiaľ nezobrazujú, keďže k tomu je potrebná implementácia na backende.

Mohli by sme pouvažovať na upravení formy dátumu tak, aby som na frontende nemusela riešiť orezávanie a rovnako by sme mohli skúsiť extrahovať z linku iba odkaz na doménu, pričom preklik by ostal pôvodný.

[AMS-114] [Komunikácia medzi MongoDB a API serverom](#) Created: 23/Nov/21 Updated: 07/Dec/21 Resolved: 07/Dec/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	7 hours		
Time Spent:	6 hours		
Original estimate:	1 day, 5 hours		

Story point estimate:	13
Sprint:	AMS Sprint 4
Rank:	0 i000mc:

Description

+ stránkovanie

Comments

Comment by [Adam Šípka](#) [07/Dec/21]

Komunikácia medzi MongoDB a API serverom bola zabezpečená. ID hodnoty dokumentov získaných z Elasticu sú použité pri vyhľadávaní v databáze a pre každé z nich nám Mongo vráti celý dokument. Pre ten ešte nájdeme kľúčové slová v kolekcii crimemaps a tie priložíme k danému dokumentu. Výsledné články sú poslané ako response vo formáte JSON na NodeJS server.

[AMS-113] [Indexovanie stiahnutých dát](#) Created: 23/Nov/21 Updated: 29/Nov/21 Resolved: 29/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	7 hours, 30 minutes		
Original estimate:	Not Specified		

Attachments:	 image-20211129-155001.png
Story point estimate:	8
Sprint:	AMS Sprint 4
Rank:	0ji000me:

Description

Rozbehnutie docker kontajneru pre elastic search.

Zaindexovanie UK clankov za posledne dva tyzdne.

Comments

Comment by [Jakub Hlavačka](#) [29/Nov/21]

https://github.com/FIIT-TEAM8/elasticsearch_mongo

Github actions nepodporuje deploy ale iba build. Pretoze by bolo zbytocne zakazdym indexovat.

Zakomentoval som tento kontajner v produknej verzii, pretoze by sa pri kazdom docker-compose up indexoval

Comment by [Jakub Hlavačka](#) [29/Nov/21]

[AMS-111] [Dokumentácia k riadeniu projektu](#) Created: 23/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	6 hours		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i0008v:2


[AMS-110] [Dokumentácia k inžinierskemu dielu](#) Created: 23/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	4 hours		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i0008v:9

0

 **[AMS-107] [Vytvorenie articles db na virtualnej masine](#)** Created: 21/Nov/21 Updated: 21/Nov/21 Resolved: 21/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i000lj:

0

 **[AMS-106] [Analyza shard, nodes](#)** Created: 16/Nov/21 Updated: 21/Nov/21 Resolved: 21/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i000lb:

0

 [AMS-105] [Analyza River plugin](#) Created: 16/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0ji00013:

Description

<https://hevodata.com/learn/integrating-elasticsearch-and-mongodb/>

Comments

Comment by [Jakub Hlavačka](#) [16/Nov/21]

River plugin a Mongo connector su deprecated pre MongoDB v.4 a Elasticsearch v.7 . Myslim si, ze bude najvhodnejsie pouzit logstash, kedze to je sucast balika elastic.

<https://stackoverflow.com/questions/56628675/sync-mongodb-with-elasticsearch>

[AMS-103] [Rozšírenie docker compose o MongoDB a Elasticsearch](#) Created:
15/Nov/21 Updated: 21/Nov/21 Resolved: 21/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	3 hours	Remaining Estimate:	3 hours
Σ Time Spent:	2 hours	Time Spent:	2 hours
Σ Original Estimate:	5 hours	Original estimate:	5 hours

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-105	Analyza River plugin	Subtask	Done	Jakub Hlavačka
	AMS-106	Analyza shard, nodes	Subtask	Done	Jakub Hlavačka
	AMS-107	Vytvorenie articles db na virtualnej ...	Subtask	Done	Jakub Hlavačka
Story point estimate:	5				
Sprint:	AMS Sprint 3				
Rank:	0ji0008v:i				

Comments

Comment by [Jakub Hlavačka](#) [21/Nov/21]

Zistit ako funguje logstash.

Comment by [Jakub Hlavačka](#) [21/Nov/21]

<https://www.elastic.co/elasticon/conf/2016/sf/quantitative-cluster-sizing> - Tu pisu, ze je vhodne otestovat na zaindexovanych datach scenare zistit, ze kedy jednotlivé shardy pomáhajú a kedy nie... Shary môžu zahliť CPU, pretože vyhľadávanie v každom sharde funguje na samotnom CPU vlakne.

<https://www.elastic.co/guide/en/elasticsearch/reference/current/size-your-shards.html#shard-size-recommendation> - Tu spominaju, ze by sme sa mali snazit drzat velkost jedneho shardu medzi 10GB az 50GB

<https://www.elastic.co/guide/en/elasticsearch/reference/current/size-your-shards.html#shard-count-recommendation> - Tu sa spomina, ze ak jeden node prekroci 20 shardov pre 1GB heap memory, tak sa ma pridat dalsi node

<https://discuss.elastic.co/t/how-many-nodes-should-an-elasticsearch-cluster-have/227674> - Tu opisuju, ze kolko nodov by mal mat cluster a vetu each node is one host (either physical or virtual) som pochopil tak, ze mi budeme mat jeden node.

Vysledok: Cluster ma jeden node, jeden shard a jednu repliku

Navrh na dalsi task <https://www.elastic.co/guide/en/elasticsearch/reference/current/size-your-shards.html#use-ds-ilm-for-time-series> - Zistit nieco o elastic ILM, pretoze pouzivame time series data a zjavne sa to hodi.

Dalsi task zistit, ci sa nam hodi logstash na updateovanie elasticsearch vzdy, ked pride novy zaznam do mognodb zo scrapera. Ak ano, tak ho pouzit.

[AMS-102] [Deployment scrapera](#) Created: 15/Nov/21 Updated: 29/Nov/21 Resolved: 28/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 5 hours		
Original estimate:	Not Specified		

Story point estimate:	13
Sprint:	AMS Sprint 4
Rank:	0 i000mh:

Description

+ systémové premenné na ovládanie kontajnera

Spustenie scrapera na posledné dva týždne (zatiaľ) z UK.

Comments

Comment by [Dominik Horvath](#) [28/Nov/21]

parser bol upraveny na konfiguraciu pomocou argumentov. Deploy bol vyrieseny pomocou scrapyd modulu. Tento modul vytvori jednoduchy HTTP API server nad nasim crawlerom. Cela aplikacia je dockerizovana a bezi na virtualnom stroji timu.

Ovladanie je teda prostrednictvom HTTP API, momentalne pouzitelne len priamo z virtualneho stroja - NGNINX konfiguracia z nejakého dôvodu zabranuje POST requestom z externých sietí. Toto je práca do ďalšej story.

Priklad ovladania:

Po prihlaseni na virtualny stroj je mozne spustit nas crawler nasledovnym sposobom:

```
curl localhost:6800/schedule.json -d spider=news_spider -d project=ams -d
search_from=2021-11-14 -d search_to=2021-11-28 -d locale=en-gb -d
crimes_file=list_of_crimes.txt
```

Toto spusti crawlovanie dat v danom casovom rozpati so specifikovanim suborom zlocinov.

Cely kod je k dispozicii na [githube](#).

Poznamka: v README na githube su argumenty posielane pomocou **-a** prepinača a pri vzorke kodu v tomto komentari je pouzity **-d** prepinač. Nie je to chyba, **-d** pri curl-e iba posielala POST data a samotny server ich automaticky potom prechodi na **-a**. Cize pri lokalnom developmente, ked scrapy nebezi na serveri, pouzivat **-a** argumenty a sposob specifikovany v README.

0

 **[AMS-101] vytvorit komprimovanu collection z main.py a zmerat cas**

indexovania Created: 12/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i000k3:



0

 **[AMS-100] [Vytvorit config pre articles_index](#)** Created: 12/Nov/21 Updated: 12/Nov/21 Resolved: 12/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i000jv:

0


 **[AMS-99] [Skonzultovat komentý v storke](#)** Created: 12/Nov/21 Updated: 15/Nov/21 Resolved: 15/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0ji000k7:

0

 **[AMS-98] [Napisat dokumentaciu do README.md](#)** Created: 12/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i000jz:

0

 **[AMS-97] Vytvorenie docker-compose file s MongoDB, Elasticsearch a Kibana**

Created: 12/Nov/21 Updated: 12/Nov/21 Resolved: 12/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 3
Rank:	0 i000jr:

[AMS-95] [Implementácia dummy Flask API](#) Created: 09/Nov/21 Updated: 15/Nov/21 Resolved: 15/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	5 hours		

Issue links:	Blocks
	is blocked by AMS-94 Návrh Flask API Done
Story point estimate:	5
Sprint:	AMS Sprint 3
Rank:	0 i000jj:

Comments

Comment by [Dominik Horvath](#) [15/Nov/21]

Dummy flask api hotova. Program momentalne vracia nahodny pocet clankov, vsetko jednoduche lorem-ipsum. Jednoducha dokumentacia je k dispozicii v README repozitaru [flask_serveru](#).

[AMS-94] [Návrh Flask API](#) Created: 09/Nov/21 Updated: 26/Feb/22 Resolved: 15/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	6 hours		
Original estimate:	5 hours		

Attachments:	image-20211114-140311.png image-20211114-142548.png image-20211114-142800.png
Issue links:	Blocks blocks AMS-95 Implementácia dummy Flask API Done
Story point estimate:	5
Sprint:	AMS Sprint 3
Rank:	0 i0008x:i

Description

Nástroj na dokumentáciu + dokumentácia.

Comments

Comment by [Jakub Hlavačka](#) [10/Nov/21]

[Táňa Poláková https://betterprogramming.pub/restful-api-design-step-by-step-guide-2f2c9f9fcd9f](https://betterprogramming.pub/restful-api-design-step-by-step-guide-2f2c9f9fcd9f)

Comment by [Táňa Poláková](#) [14/Nov/21]

Versioning: robi sa kvoli supportu starsich zakaznikov v pripade, ze sa nejake poziadavky zmenia a my spravime

Existuju 4 strategie versioningu:

1. URI path versioning - /api/v1/Customers/1 ... /api/v2/Customers/1

2. URL parameter versioning - /api/Categories?v=1.1
3. Content Negotiation - URLs sa nemenia, v accept header vidime, ktora verzia je ta, ktoru chceme:
 0. Accept: application/app.v1.categories
 1. Accept: application/app.v2.categories
4. Custom Header:
 0. x-App-version: 1.3
 1. x-App-version: 2017-08-12

Continuous versioning - neexistuju ziadne verzie, iba json responses vzdy nejak prisposobime, aby obsahovalo p...
sme pridali na zaklade zmenenych poziadaviek (viac: <https://www.youtube.com/watch?v=M2KCu0Oq3JE>).

Comment by [Táňa Poláková](#) [14/Nov/21]

Dokumentáciu je vhodné písať/generovať, až keď sú requesty implementované.

Postman

Vyhody

- intuitívne grafické rozhranie dostupné aj cez prehliadač
- podpora kolaborácie
- testovanie API

Nevyhody

- výsledná dokumentácia je dostupná cez postman URL a nie cez našu vlastnú - to sa dá len pri PRO verzii
- response json sa neda dokumentovať

Ukazka dokumentacie

Ako funguje

1. nainstalujeme si Postman (<https://www.postman.com/downloads/>)
2. do timoveho workspace sa vieme dostat cez invite link dostupny na nasom drive
3. vytvorime si collection, kam budu spadat nase implementovane requesty, mozeme vytvorit aj viac kolekci
4. nasa Collection je vlastne subor nasich requestov, ktoru vieme ako celok zdokumentovat - vieme jej dat nazov
Collections - nam asi bude stacit len Collection "Articles"
5. v ramci kolekcie si vytvorime vsetky implementovane requesty - rozdelené aj podla metod
6. pri kazdom requeste vieme pridat description, query parametre, ulozit, ako vyzera priklad json response

Flask-RESTPlus

Rozsirenie pre Flask, dostupne cez pip: pip install flask-restplus

<https://www.imaginarycloud.com/blog/flask-python/>

Vyhody

- vacsinu si naprogramujeme sami
- dokumentacia moze byt na nasej domene
- kolaboracia je mozna vlastne len tak, ze pristupujeme k spolocnemu kodu
- response json sa da dokumentovat

Nevyhody

- vacsinu si naprogramujeme sami
- nie je graficke rozhranie
- vysledny vzhľad dokumentacie sa mi nepaci (😞)
- testujeme iba cez konzolu

Ukazka kodu

Ukazka dokumentacie

https://flask-ic.herokuapp.com/ocumented_api/doc



[AMS-93] [Analýza možností implementácie testovania](#) Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	5 hours		

Story point estimate:	5
Sprint:	AMS Sprint 3
Rank:	0 i0008v:4

Comments

Comment by [Adam Šípka](#) [23/Nov/21]

Typy testovania:

1. Functional Testing
2. Usability testing - zatiaľ netreba
3. Interface testing
4. Compatibility testing - zatiaľ netreba
5. Performance testing - zatiaľ asi netreba
6. Security testing - zatiaľ asi netreba

Functional Testing

Toto testovanie kontroluje UI, API, databázu, bezpečnosť, klient/server komunikáciu a fungovanie aplikácie.

a) Unit testing

Účelom je otestovať každú funkciu poskytnutím vhodného vstupu a overením výstupu vzhľadom na funkčné požiadavky. Testujeme malé časti softvéru, vykonáva sa ako prvé Frameworky na testovanie Node.js aplikácií:

- Mocha
- Jest
- Chai
- Jasmine
- AVA
- Pre Python:
- unittest + Nose2 a Testify (optional)
- PyTest
- DocTest

Ako písať ľahko testovateľný kód:

- deterministické funkcie: výstupná hodnota závisí od vstupnej a nie od skrytých premenných v danej funkcii
- Inversion of Control technika: oddeliť decision making kód (kedy a za akých podmienok sa niečo vykoná) a action kód (čo sa za vykoná)
- Dependency Injection: Objekt (client) prijíma ďalšie objekty (services/dependencies) od ktorých je závislý. Kód kde je service vložený (injector) určí clientovi, akú službu bude používať, namiesto toho, aby si to client určil sám.

```
public class SmartHomeController
{
    private readonly IDateTimeProvider _dateTimeProvider; // Dependency

    public SmartHomeController(IDateTimeProvider dateTimeProvider)
    {
        // Inject required dependency in the constructor.
        _dateTimeProvider = dateTimeProvider;
    }

    public void ActuateLights(bool motionDetected)
    {
        // Delegating the responsibility
        DateTime time = _dateTimeProvider.GetDateTime();

        ...
    }
}
```

b) Integration testing

Jednotlivé moduly sú skombinované a otestované ako celok. Zamerané na presun dát medzi modulmi.

Ak medzi niektorými už existuje prepojenie/funkcionalita a ostatné nie sú hotové, môžeme

existovať existujúce. Cieľom je odhaliť chyby v interakciách medzi modulmi, keďže tie sú zväčša písané rozdielnymi ľuďmi.

[AMS-92] [Zabezpečenie komunikácie medzi klientom a serverom](#) Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 1 hour		
Original estimate:	1 day		

Story point estimate:	8
Sprint:	AMS Sprint 3
Rank:	0 i000iv:

Description

Vytvoriť wrapper pre fetch. Je potrebné, aby fungoval aj na development prostredí (development cross-origin requests).

[AMS-91] Úprava dokumentácie zo stretnutí a šprintov tak, aby mohli ísť na stránku tímu Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 23/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day		
Original estimate:	1 day		

Story point estimate:	8
Sprint:	AMS Sprint 3
Rank:	0 j000kt:

Description

Denníky zo stretnutí, upratanie šprintov, exporthy úloh, šprint review, ...

Comments

Comment by [Táňa Poláková](#) [23/Nov/21]

Denníky zo stretnutí, šprint reviews aj exporthy úloh sú hotové a sú zatiaľ umiestnené na tímovom drive.

[AMS-90] [Integrácia MongoDB a Elasticsearch](#) Created: 09/Nov/21 Updated: 16/Nov/21 Resolved: 16/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	0 minutes
Σ Time Spent:	1 day, 1 hour	Time Spent:	1 day, 1 hour
Σ Original Estimate:	1 day	Original estimate:	1 day

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-97	Vytvorenie docker-compose file s Mong...	Subtask	Done	Jakub Hlavačka
	AMS-98	Napisat dokumentáciu do README.md	Subtask	Done	Jakub Hlavačka
	AMS-99	Skonzultovat komentý v storke	Subtask	Done	Jakub Hlavačka
	AMS-100	Vytvorit config pre articles_index	Subtask	Done	Jakub Hlavačka
	AMS-101	vytvorit komprimovanu collection z ma...	Subtask	Done	Jakub Hlavačka
Story point estimate:	8				
Sprint:	AMS Sprint 3				
Rank:	0ji0008w:				

Description

výskum možností prepojenia, kontrola nastavení - aby neukladal elastic text, iba ID.

Comments

Comment by [Jakub Hlavačka](#) [12/Nov/21]

Neskor potrebne zabezpecenie elasticsearch <https://www.elastic.co/guide/en/elastic-stack-get-started/current/get-started-docker.html>

Comment by [Jakub Hlavačka](#) [12/Nov/21]

Skonzultovat skalovanie. (Pocet shards, pocet nodes, ...).

Comment by [Jakub Hlavačka](#) [12/Nov/21]

Pozriet sa na <https://hevodata.com/learn/integrating-elasticsearch-and-mongodb/>

Comment by [Jakub Hlavačka](#) [12/Nov/21]

https://github.com/FIIT-TEAM8/elasticsearch_mongo

Comment by [Jakub Hlavačka](#) [16/Nov/21]

Casovane meranie: total time in **seconds: 804.9638478755951**, cize cca **13 minut**.


Comment by [Jakub Hlavačka](#) [16/Nov/21]

Elasticsearch vyuziva nejaku jednoduchsiu kompresiu uz defaulte, nasiel som, ze podporuje este jeden sposob, lenze ten zredukuje ulozisko o nejake ~3%... Samotny index ma 126MB z 308MB dat, ked obsahuje jeden shard a jednu repliku. Tym padom ma jeden shard, v ktorom sa drzi LUCENE index cca ~60MB

[AMS-89] [Integrácia scraperu a vylepšeného parseru](#) Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 22/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day		
Original estimate:	5 hours		

Attachments:	 improved_parser_measurments.txt
Story point estimate:	5
Sprint:	AMS Sprint 3
Rank:	0 i000kr:

Description

Je potrebné vylepšiť parser z predošlého šprintu (konkrétne z tasku “Parsovanie tagov pomocou CSS Selector”) a integrovať ho do nášho scraperu.

Comments

Comment by [Jakub Müller](#) [23/Nov/21]

Parser bol integrovaný bez nejakých problémov, keďže bol vytvorený pomocou knižnice Scrapy, ktorá je použitá aj pri samotnom scraperi.

Comment by [Jakub Müller](#) [23/Nov/21]

Vylepšenie parseru sa týkalo odstránenia nepotrebných dát z tagov paragrafov a headingov, ktoré boli vyselektované pomocou XPATH selektora. Tieto nepotrebné dáta predstavovali atribúty daných tagov (teda napríklad classy, ID-čka a rôzne iné špecifikácie), ktoré nepotrebujeme, pretože nijako neovplyvňujú formu nášho výsledného HTML súboru.

Comment by [Jakub Müller](#) [23/Nov/21]

Na odstránenie týchto atribútov bolo vyskúšaných viacero spôsobov. Bolo skúmané, či sa nedajú jednoducho odignorovať už pri samotnom selektovaní tagov pomocou **XPATH** selektora. Toto sa ukázalo ako nemožné. Následne bolo skúšané odstraňovanie atribútov pomocou **regex**-ov a aj pomocou **listových operácií** v Pythone.

Nakoniec sa ukázalo, že riešenie pomocou listových operácií bolo najlepšie. Následne sa tiež kontrolovalo, či daný tag má vôbec nejaký obsah a ak nie, tak sa celý odstránil.

Comment by [Jakub Müller](#) [23/Nov/21]

Toto riešenie bolo testované na rovnakom súbore ako pôvodný parser (1_articles.jl, ktorý mal veľkosť približne 3,5GB). Pôvodný parser dokázal zredukovať daný súbor o cca **92%** a vylepšený parser ho dokázal zredukovať o **93%**, teda o približne 1% viac, pričom sa nestratili žiadne potrebné informácie z daných HTML súborov. Tento vylepšený parser taktiež dokázal spracovať tieto články v približne rovnakom čase ako pôvodný a to za cca **3 minúty**.

Detailné výsledky testovania vylepšeného parseru sú uvedené v prílohe **improved_parser_measurments.txt**

[AMS-88] [Rozšířit úložisko pomocou nepriradeného disku](#) Created: 09/Nov/21 Updated: 23/Nov/21 Resolved: 14/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	2 hours		
Original estimate:	5 hours		

Story point estimate:	5
Sprint:	AMS Sprint 3
Rank:	0 j0008y:

Comments

Comment by [Dominik Horvath](#) [14/Nov/21]

Disk is mounted under **/data**

Comment by [Dominik Horvath](#) [23/Nov/21]

Disk je mounntnuty pod /data na virtualnom stroji timu. Obsahuje 20GB volneho miesta a bol naformatovany na suborovy system ext4.

[AMS-87] [Zabrániť zacykleným buildom](#) Created: 07/Nov/21 Updated: 22/Nov/21 Resolved: 15/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Dominik Horvath	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	3 hours		
Time Spent:	2 hours		
Original estimate:	5 hours		

Story point estimate:	5
Sprint:	AMS Sprint 3
Rank:	0 i000bu:

Description


Github actions timeout setting aby nam nebezali buildy zbytocne dlho

0

 **[AMS-74] [Sekvenčné scrapovanie](#)** Created: 02/Nov/21 Updated: 08/Nov/21 Resolved: 06/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	7 hours		
Original estimate:	7 hours		

Attachments:	 data.zip  timing.txt
Sprint:	AMS Sprint 04
Rank:	0 i000f3:

Comments

Comment by [Dominik Horvath](#) [06/Nov/21]

Sekvenčne scrapovanie bolo skusane pomocou viacerých prístupov, a to BeautifulSoup knižnica, LXML knižnica spolu s [XPATH |<https://www.w3.org/TR/1999/REC-xpath-19991116/>]selector standardom a taktiež [python-fork reader-view](#) pluginu obsiahnutom v prehliadaci firefox.

BeautifulSoup knižnica nebola dostacujuca. Na jednu stranu bola velmi pomala + vyzadovala zlozity kod na extrahovanie len tych atributov, ktore su pre nas zaujimave.

Fork readerview pluginu fungoval perfektne, jeho vytvorene vzorky fungovali 99% pripadov tak ako v prehliadaci, avsak bol velmi pomaly - datovu sadu 1_articles.zip spracovaval asi 4 hodiny, kde sa nedostal ani do polovice clankov. Nasledne som program vypol.

LXML knižnica spolu s XPATH selectorom fungovala najlepsie. Bola vytvorena tato vyhľadavacia query:

```
"//p | //h1 | //h2 | //h3 | //h4 | //h5 | //h6"
```


ktora zoberie všetky stanovené HTML tagy v pôvodnom poradí, ako boli v dokumente. Knížnica bola veľmi rýchla a redukcia dát významná (príloha timing.txt)

V prílohe **data.zip** sa nachádza spracovaných prvých 100 záznamov z datovej sady 1_articles.zip a tiež zdrojový kód. V priečinku **parsed** sa nachádza 100 HTML súborov, jeden pre každú spracovanú stránku. V súbore **articles_sample_links.txt** sa pre možnosť kontroly nachádzajú odkazy na pôvodné články. HTML súbory majú v názve číslo riadku, na ktorom sa v súbore **articles_sample_links.txt** nachádza odkaz k pôvodnému článku.

[Jakub Müller](#) [Táňa Poláková](#) FYI

[AMS-73] [Vyhládanie RSS dvojčat'a k HTML článku - všeobecný postup](#) Created: 02/Nov/21 Updated: 23/Nov/21 Resolved: 09/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	4 hours		
Time Spent:	4 hours		
Original estimate:	1 day		

Sprint:	AMS Sprint 04
Rank:	0 i0008z:

Comments

Comment by [Dominik Horvath](#) [23/Nov/21]

Bohuzial, nic v podobnej sfere v dnesnej dobe neexistuje. Existujuce citacky RSS streamov potrebujú priamy odkaz na stream konkrétneho webu, nedokážu získať stream URL len z "base" URL domeny. Taktiež každý portál obsahuje vlastné štruktúry pre RSS stream URL, ich tvar nie je standardizovaný.

V dnesnej dobe je taktiež náradne najštrkompletnej RSS stream, čiže RSS stream obsahujúci aj telo samotného článku, nielen metadata o nom. Vzhľadom na to, že väčšina RSS streamov obsahuje len odkaz na samotný článok, tento postup by nám neposkytol žiadnu výhodu oproti terajšiemu prístupu.

[AMS-72] [Technický návrh developerského prostredia](#) Created: 02/Nov/21 Updated:
25/Feb/22 Resolved: 24/Feb/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	2 hours, 59 minutes		
Time Spent:	5 hours, 1 minute		
Original estimate:	1 day		

Story point estimate:	8
Sprint:	AMS Sprint 4
Rank:	0 i000nr:

Comments

Comment by [David Silady](#) [07/Dec/21]

Je nutné vytvoriť inicializačné skripty pre DEV databázy s dummy dátami.

docker-compose má obmedzené scripting ability, ale je možné podľa prepínačov nastavovať systémové premenné

Aktuálny návrh:

Pri spustení docker-compose sa špecifikuje, na ktorom containeri sa bude robiť development, to si developer naclonuje k sebe manuálne.

Ostatné containery sa spustia so system variable DEV=true, kde v prípade, že sa v DEV prostredí majú správať inak ako v PROD (napr. databázy), vykonajú inú/upravenú verziu inicializácie.

[AMS-71] [Naplnenie MongoDB pri scrapovaní](#) Created: 02/Nov/21 Updated: 23/Nov/21 Resolved: 14/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day		
Original estimate:	1 day		

Sprint:	AMS Sprint 04, AMS Sprint 3
Rank:	0 i0008y:i

Description

Fields: meno, link, datum vydania, region, jazyk, telo clanku

Comments

Comment by [Adam Šípka](#) [23/Nov/21]

Podarilo sa mi upraviť scraper tak, aby sa získané dáta hneď ukladali do Mongo databázy. Momentálne sa jedná len o lokálnu databázu umiestnenú na mojom osobnom zariadení. Nebude problém neskôr zmeniť, stačí v súbore settings.py upraviť nasledovné premenné (+ pridať heslo):

```
# db server and port (local for now)
MONGODB_SERVER = "localhost"
MONGODB_PORT = 27017
```

Databáza je tvorená z 3 kolekcií (articles, crimemaps a errorlinks). Okrem automatického indexu ***_id***, ktoré tvorí mongo automaticky, som v každej kolekcií nastavil field ***link*** na unikátny index, aby sa zabránilo duplikátom.

[AMS-70] [Rozšírenie google news scrapperu na lokáciu UK](#) Created: 02/Nov/21 Updated: 02/Nov/21 Resolved: 02/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 hour		
Original estimate:	1 hour		

Sprint:	AMS Sprint 04
Rank:	0ji000e7:

Comments

Comment by [Dominik Horvath](#) [02/Nov/21]

Scrapper je pripraveny na parsovanie clankov z lokacie UK, v jazyku anglictina. Staci inicializovat triedu na parsovanie takto:

```
gnews_parser = GnewsParser()
gnews_parser.setup_search("covid", '2021-09-01', '2021-09-02', locale="en-gb")
```

0

 **[AMS-66] [Ako mongoDB pracuje s vopred komprimovanim textom](#)** Created:

01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0 i000db:

Comments

Comment by [Jakub Hlavačka](#) [01/Nov/21]

MongoDB podporuje rovnaku kompresiu ako ZIP, pricom v mongoDB budu data v ramci jedneho zaznamu celistve, takze sme zvolili tento sposob.



MongoDB podporuje zstd, snappy, zlib a gzip

0

 **[AMS-65] [vyskusat nad datach](#)** Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 image-20211029-170742.png  image-20211029-170825.png
Sprint:	AMS Sprint 04
Rank:	0 i000d3:

Comments

Comment by [Jakub Hlavačka](#) [01/Nov/21]

Zazipované html tagy bez title, keywords, link a published




Compresovane html tagy bez title, keywords, link a published



Zatial vyzerá najrozumnejšie mongoDB alebo ten ZIP.



0

 **[AMS-64] [realne kniznice](#)** Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0ji000cv:

Comments

Comment by [Jakub Hlavačka](#) [01/Nov/21]

v mongoDB sa pouziva zlib. V ramci ZIP riesenia boli vyskusane zlib a ziplib kniznice v Pythone.

0

 **[AMS-63] [elastic search a komprimacia](#)** Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0ji000cn:

Comments

Comment by [Jakub Hlavačka](#) [01/Nov/21]
Vytvorene ako nova storka...

0

 **[AMS-62] [analýza vyhľadavania nad komprimovaným textom](#)** Created:

01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0ji000cf:

Comments

Comment by [Jakub Hlavačka](#) [01/Nov/21]

Indexovanie sa bude robiť nad nekomprimovaným textom, pričom indexy budú fungovať nad ID rows v mongoDB, kde už bude komprimovaný text.

0

 **[AMS-61] [bezstratova kompresia text](#)** Created: 01/Nov/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0 i000c7:

Comments



Comment by [Jakub Hlavačka](#) [01/Nov/21]

Najlepsi huffman, komprimovat viacerymi sposobmi da vzdy rovnaky vysledok ako da posledna komprimacia...

[AMS-60] [Pamäťová zložitosť elastic search na vzorke dát](#) Created: 01/Nov/21 Updated: 12/Nov/21 Resolved: 12/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	6 hours, 30 minutes		
Original estimate:	6 hours		

Attachments:	 image-20211108-145633.png  main.py
Sprint:	AMS Sprint 04, AMS Sprint 3
Rank:	0 i0008x:

Description

Zobrat dump od [Adam Šípka](#) a spravit index v elastic search, pricom jednotlivé values pre terms budu ID zo zaznamov v MongoDB

Comments

Comment by [Jakub Hlavačka](#) [07/Nov/21]

Neskôr by bolo vhodné použiť asi toto <https://hevo.com/learn/integrating-elasticsearch-and-mongodb/>

Comment by [Jakub Hlavačka](#) [08/Nov/21]

Mam zatiaľ vytvorený docker-compose, kde je mongodb a elasticsearch s ich vlastnými volumes. Mongo má col clanky z articles_1.zip . K tasku prikladám script, pomocou ktorého som čítal z mongodb a snažil sa vytvárať index

Posting list sa mi zatiaľ nepodarilo naimplementovať, pretože elasticsearch analyzer má obmedzenie pre tokenizáciu

Aby som aspoň hrubým odhadom zistil pamäťovú zložitosť vytvoreného indexu v elasticsearch, tak som vkladal jeho kľuč bola id z mongodb... Pricom po nejakom 650 clanku vybehol nasledovný error

Navrhujem dalsi postup:

- vyriesit problem s obmedzenim tokenizacie
- vytvarat uz rovno index (posting list), ktory budeme vyuzivat v AMS
- vytvorit dalsi task na analyzu nastroja z <https://hevo.com/learn/integrating-elasticsearch-and-mongo>
- Alokovat viac casu !!!

EDIT: pridane navrh na dalsi tak s nastrojom.






Comment by [Jakub Hlavačka](#) [12/Nov/21]

Z 308MB clankov spravil elasticsearch 244MB.

[AMS-59] [Parsovanie tagov pomocou lxml, readability, trafiletura](#) Created: 30/Oct/21 Updated: 09/Nov/21 Resolved: 09/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	6 hours		
Original estimate:	3 hours		

Attachments:	 2_sk_article.html  lxml_measurement - Copy.txt  p_lxml_measurement.txt  readability_measurement.txt  trafiletura_measurement.txt
Sprint:	AMS Sprint 04
Rank:	0 i00093:

Description

pozriet aj ine jazyky, pozriet dokumentaciu, preco kniznica suvisi s jazykom

Comments

Comment by [Táňa Poláková](#) [07/Nov/21]

Readability - funguje iba na zaklade nazvov tagov a tried, nie na zaklade jazyka textu - funkcia summary() extrahuje hlavny obsah HTML stranky. Ako prve vymaze vsetky script alebo style tagy. Nasledne maze tagy na zaklade regexov:

```
{{ 'unlikelyCandidatesRe':  
re.compile('combx|comment|community|disqus|extra|foot|header|menu|remark|rss|shoutbox|sidebar|sponsor|ad-  
break|agegate|pagination|pager|popup|tweet|twitter',re.I), 'okMaybeItsACandidateRe':  
re.compile('and|article|body|column|main|shadow',re.I),}}  
  
if (  
    REGEXES["unlikelyCandidatesRe"].search(s)
```

```
    and (not REGEXES["okMaybeItsACandidateRe"].search(s))
    and elem.tag not in ["html", "body"]
):
    log.debug("Removing unlikely candidate - %s" % describe(elem))
    elem.drop_tree()
```

Potom pridava score tagom na zaklade urcitych kriterii:

```
def class_weight(self, e):
    weight = 0
    for feature in [e.get("class", None), e.get("id", None)]:
        if feature:
            if REGEXES["negativeRe"].search(feature):
                weight -= 25

            if REGEXES["positiveRe"].search(feature):
                weight += 25

            if self.positive_keywords and self.positive_keywords.search(feature):
                weight += 25

            if self.negative_keywords and self.negative_keywords.search(feature):
                weight -= 25

    if self.positive_keywords and self.positive_keywords.match("tag-" + e.tag):
        weight += 25

    if self.negative_keywords and self.negative_keywords.match("tag-" + e.tag):
        weight -= 25

    return weight

def score_node(self, elem):
    content_score = self.class_weight(elem)
    name = elem.tag.lower()
    if name in ["div", "article"]:
        content_score += 5
    elif name in ["pre", "td", "blockquote"]:
        content_score += 3
    elif name in ["address", "ol", "ul", "dl", "dd", "dt", "li", "form", "aside"]:
        content_score -= 3
    elif name in [
        "h1",
        "h2",
        "h3",
        "h4",
        "h5",
        "h6",
        "th",
        "header",
        "footer",
        "nav",
    ]:
        content_score -= 5
    return {"content_score": content_score, "elem": elem}
```


Vyberie sa najlepsí kandidát na základe score a potom sa hľadajú susedné elementy, ktoré by s najlepším kandidátom mohli súvisieť.

Comment by [Táňa Poláková](#) [08/Nov/21]

Readability - pri 20-tich článkoch nesedel jeden extrahovaný s pôvodným článkom. Extrahovaný hlavný obsah bol nezmysel - iba svg obrázky.

V kóde nemajú veľkú váhu headings, čo by bolo možné upraviť pre naše potreby. Všeobecne mi ale príde, že v kóde je bordel - veľa zakomentovaných vecí.

Comment by [Táňa Poláková](#) [08/Nov/21]

Trafilatura - všetkých 20 článkov obsahovalo hlavný content - aj ten, ktorý readability nezvládol. Zaujímavosťou bolo, že v kóde použili readability a na základe dĺžky extrahovaného textu sa rozhodli, či použijú svoj alebo readability algoritmus.

Rozdiel je, že trafilatura vracia iba holý text, bez tagov.

Comment by [Táňa Poláková](#) [08/Nov/21]

Readability funguje rovnako pre anglické aj slovenské články.

Príklad:

(zdroj: <https://www.cas.sk/clanok/2603841/brutalna-dvojnashobna-vrazda-v-lednickych-rovniach-kedy-sa-zacne-proces-s-obzalovanim/>)

[2 sk article.html](#)[□]

Trafilatura tiež.

[AMS-57] [Zavesenie klienta](#) Created: 27/Oct/21 Updated: 07/Nov/21 Resolved: 07/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	1 hour, 30 minutes		
Time Spent:	1 day, 3 hours, 30 minutes		
Original estimate:	1 day, 5 hours		

Story point estimate:	13
Sprint:	AMS Sprint 04
Rank:	0 i000bj:

Description

Zavesit' klienta bud' cez Express alebo priamo cez NGINX. Testovacie API calls, redux, ...

0

 **[AMS-56] [NGINX konfigurácia](#)** Created: 27/Oct/21 Updated: 31/Oct/21 Resolved: 31/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0 i000bb:

0

 **[AMS-55] [Kontajner Express server](#)** Created: 27/Oct/21 Updated: 31/Oct/21 Resolved: 31/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0ji000b3:

0

 **[AMS-54] Kontajner Flask API** Created: 27/Oct/21 Updated: 30/Oct/21 Resolved: 30/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Unassigned
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	AMS Sprint 04
Rank:	0ji000av:

[AMS-53] [Rozšíriť docker compose](#) Created: 27/Oct/21 Updated: 31/Oct/21 Resolved: 31/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	6 hours	Remaining Estimate:	6 hours
Σ Time Spent:	7 hours	Time Spent:	7 hours
Σ Original Estimate:	1 day, 5 hours	Original estimate:	1 day, 5 hours

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-54	Kontajner Flask API	Subtask	Done	
	AMS-55	Kontajner Express server	Subtask	Done	
	AMS-56	NGINX konfigurácia	Subtask	Done	
Story point estimate:	13				
Sprint:	AMS Sprint 04				
Rank:	0 j0009n:				

Comments

Comment by [Dominik Horvath](#) [31/Oct/21]

Vytvorene 2 github repozitare: [node server](#) a [flask server](#). Oba repozitare obsahuju github actions procedury nakonfigurovane tak, aby sa spustili pri pushnuti / mergnuti do main branche.

Github actions vykonaju build kontajneru z aktualnej main branch, nahra hotovy image na docker-hub a nasledne nacita najnovsie docker image aj na timovy virtualny stroj, kde ich rovno spusti a nahradi stare kontajnery.

Nove cesty v nginx reverse proxy:

- /api => flask server
 - /ams => node server
-

[AMS-51] [Pridat' info o projekte a opisat' členov tímu](#) Created: 27/Oct/21 Updated:
02/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	3 hours		
Time Spent:	1 hour		
Original estimate:	4 hours		

Story point estimate:	4
Sprint:	AMS Sprint 04
Rank:	0 i000a7:



[AMS-50] [Request URL s odpoveďou inou ako 200 uložiť do súboru](#) Created:
27/Oct/21 Updated: 23/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	2 hours		
Time Spent:	1 hour		
Original estimate:	3 hours		

Story point estimate:	3
Sprint:	AMS Sprint 04
Rank:	0 i0009p:

Description

Čokoľvek, čo počas scrapovania dostane odpoveď inú ako 200 je potrebné uložiť do súboru. Okrem URL treba uložiť aj zločin.

Comments



Comment by [Adam Šípka](#) [23/Nov/21]

Úspešne sa mi podarilo splniť task. URL aj zločiny sa počas behu programu ukladajú do dictionary. Ten je po ukončení programu uložený v JSON formáte.

[AMS-49] [Lokálna MongoDB so získanými dátami](#) Created: 27/Oct/21 Updated: 23/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	3 hours		
Time Spent:	5 hours		
Original estimate:	1 day		

Attachments:	 mongodb zistenia_01.txt  mongodb zistenia_02.txt
Story point estimate:	8
Sprint:	AMS Sprint 04
Rank:	0 i0009r:

Description

Naše získané dáta (všetky fields) dať do lokálnej MongoDB, zistiť, aké sú možnosti práce s ňou, koľko miesta zaberá na disku - uvidíme, ako naše dáta Mongo zredukuje. Zistiť, ako je tvorené ID.

Comments

Comment by [Adam Šípka](#) [23/Nov/21]

Veľkosť pôvodných dát: 28,3 GB (3,53 + 4,75 + 4,94 + 4,56 + 4,66 + 5,87)

Veľkosť databázy (bez extra kompresie): 9,028870144 GB

Veľkosť databázy (dodatočná kompresia): 5,791744000 GB

Počet záznamov: 228,774 (articles_all)

Bez komprimovania (úplne basic databáza):

"db" : "articledata", meno databázy

"collections" : 1, počet kolekcií (niečo ako tabuľka v SQL databázach)

"views" : 0,

```
"objects" : 228774, počet všetkých dokumentov vo všetkých kolekciách
"avgObjSize" : 125179.16999746475, priemerná veľkosť 1 dokumentu
"dataSize" : 28637739437, veľkosť všetkých dokumentov (neskomprimovaných)
"storageSize" : 9025990656, priestor alokovaný všetkým kolekciám v databáze vrátane voľného miesta (bez indexov)
"freeStorageSize" : 2019328,
"indexes" : 1, počet indexov (1 automaticky vytvorený - niečo ako id, unikátny pre každý dokument)
"indexSize" : 2879488, priestor alokovaný pre indexy
"indexFreeStorageSize" : 163840,
"totalSize" : 9028870144, súčet indexSize a storageSize (veľkosť databázy)
"totalFreeStorageSize" : 2183168,
"scaleFactor" : 1,
"fsUsedSize" : 511983788032, využitá veľkosť disku kde mongo ukladá veci
"fsTotalSize" : 984432504832, celková veľkosť disku kde mongo ukladá veci
"ok" : 1
```

Databáza kde bola kolekcia vytvorená s dodatočnou kompresiou dát:

```
db.createCollection('collectionName', {storageEngine: {wiredTiger: {configString:
'block_compressor=zlib'}}})
"db" : "articlescompressed",
"collections" : 1,
"views" : 0,
"objects" : 228774,
"avgObjSize" : 125179.16999746475,
"dataSize" : 28637739437,
"storageSize" : 5788897280,
"freeStorageSize" : 1695744,
"indexes" : 1,
"indexSize" : 2846720,
"indexFreeStorageSize" : 118784,
"totalSize" : 5791744000,
"totalFreeStorageSize" : 1814528,
"scaleFactor" : 1,
"fsUsedSize" : 517487702016,
"fsTotalSize" : 984432504832,
"ok" : 1
```

vytvorenie databázy a kolekcie cez mongo shell (s extra kompresiou)

```
use dbName
```

```
db.createCollection('collectionName', {storageEngine: {wiredTiger: {configString:
'block_compressor=zlib'}}})
```

import dát cez command line

```
mongoimport --db dbName --collection collectionName --file fileName
```

Mongo vygeneruje unikátne id pre každý dokument sám, ale asi bude lepšie použiť url ako unikátne id, aby sme zabránili duplikátom. Alebo sa to vyrieši ešte pri scrapovaní dát, navštívené linky už nebudeme scrapovať.

[AMS-48] [Analýza bezstratovej kompresie textu](#) Created: 27/Oct/21 Updated: 01/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	2 hours	Remaining Estimate:	2 hours
Σ Time Spent:	6 hours	Time Spent:	6 hours
Σ Original Estimate:	1 day	Original estimate:	1 day

Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-61	bezstratova kompresia text	Subtask	Done	
	AMS-62	analyza vyhľadavania nad komprimovany...	Subtask	Done	
	AMS-63	elastic search a komprimacia	Subtask	Done	
	AMS-64	realne kniznice	Subtask	Done	
	AMS-65	vyskusat nad datach	Subtask	Done	
	AMS-66	Ako mongoDB pracuje s vopred komprimo...	Subtask	Done	
Story point estimate:	8				
Sprint:	AMS Sprint 04				
Rank:	0j0009h:				

Description

Analýza bezstratovej kompresie text; analýza vyhľadávania nad komprimovaným textom; elastic search a komprimácia; encoding tetxtu; reálne knižnice; vyskúšať na dátach; ako MongoDB pracuje s vopred komprimovaným textom.

[AMS-47] [Parsovanie tagov pomocou regex](#) Created: 27/Oct/21 Updated: 02/Nov/21 Resolved: 02/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	6 hours		
Time Spent:	2 hours		
Original estimate:	1 day		

Story point estimate:	8
Sprint:	AMS Sprint 04
Rank:	0 i0009b:

Description

Pomocou regexov je potrebné vyskúšať rozparsovať tagy, ktoré obsahujú text. Dôležité je, aby mali správne poradie.

Comments



Comment by [Táňa Poláková](#) [30/Oct/21]

Analýza ukázala, že nie je efektívne pracovať s regexami. Úloha sa ukončila a nahradila ju nová <https://tim8-2021.atlassian.net/browse/AMS-59>

[AMS-46] [Parsovanie tagov pomocou CSS Selector](#) Created: 27/Oct/21 Updated: 08/Nov/21 Resolved: 08/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Σ Remaining Estimate:	0 minutes	Remaining Estimate:	0 minutes
Σ Time Spent:	1 day, 5 hours	Time Spent:	6 hours
Σ Original Estimate:	1 day, 4 hours	Original estimate:	5 hours

Attachments:	 example.html  scrapy_xpath_measurments.txt.txt				
Sub-tasks:	Key	Summary	Type	Status	Assignee
	AMS-74	Sekvenčné scrapovanie	Subtask	Done	Dominik Horvath
Story point estimate:	5				
Sprint:	AMS Sprint 04				
Rank:	0ji000bn:				

Description

Pomocou CSS Selector je potrebné vyskúšať parsovať tagy, ktoré obsahujú text

Comments

Comment by [Jakub Müller](#) [08/Nov/21]

Pri analýze CSS selektora z knižnice *scrapy* sa ukázalo, že pre náš účel nebude dostačujúci. Je to z toho dôvodu, že tento selektor dokáže selektovať iba jednotlivé tagy zvlášť a my potrebujeme všetky tagy, ktoré obsahujú text (teda `<p>` a všetky `<h>` tagy) a v pôvodnom poradí.

Preto sme sa namiesto CSS selektora z knižnice *scrapy* zamerali na selektor *XPATH* z tejto knižnice.

Tento selektor dokáže vyselektovať všetky tagy s textom (teda paragrafy a headingy) v takom

poradí, v
akom sú uvedené v pôvodnom HTML súbore.

Výsledný parser bol testovaný na vzorke dát 1_articles.jl, ktorá má cca 3.7GB a obsahuje cca 27 000 článkov. Celú túto vzorku dát dokázal parser spracovať za necelé 3 minúty, pričom dokázal znížiť veľkosť výsledného súboru o cca 92%.

Detailnejšie výsledky parseru sú uvedené v prílohe **scrapy_xpath_mearuments.txt** a príklad výsledného HTML súboru, z ktorého bol vyselektovaný iba text je v prílohe **example.html**

[AMS-44] [Stránka tímu](#) Created: 18/Oct/21 Updated: 22/Nov/21 Resolved: 01/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0ji0001y:zv

[AMS-43] [Otvorenie portov na serveri](#) Created: 18/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0008v:

Description

Otvorené:
443, 80, 22 (asi aj 53)
Je lepšie ostatné ani neotvárať. (Pochybujem, že by to vôbec šlo bez ďalších komplikácií - Docker)

[AMS-39] Vyriešiť, aby fiitkar nemohol urobiť "sudo su" na virtuálnom stroji (aby nemal šancu sa zmeniť na roota) Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 19/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 hzzzzz:i

Description

a nech sa ani nevie prepnut ani na ubuntu

Comments

Comment by [Táňa Poláková](#) [17/Oct/21]

<https://www.thegeekdiary.com/how-to-disable-sudo-su-for-users-in-sudoers-configuration-file/>

[AMS-35] [Prieskum databáz](#) Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0j00012:

Description

Všetko v containeroch. Analýza MongoDB, ... kompatibilita s elastic search

[AMS-34] [Vytvorit' github projekt](#) Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0j00001:i

Description

Viac repozitárov pod jedným projektom

[AMS-32] [Prihláška na TP CUP](#) Created: 13/Oct/21 Updated: 02/Nov/21 Resolved: 02/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

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


Sprint:	
Rank:	0j00067:



[AMS-31] [Set up elastic search v docker container](#) Created: 13/Oct/21 Updated: 22/Nov/21 Resolved: 16/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 docker-compose.yml  image-20211016-144056.png
Sprint:	
Rank:	0 i00000:i

Comments

Comment by [Jakub Hlavačka](#) [16/Oct/21]

For more information on this comment, please click on the comment icon in the right-hand column of the comment list.

Pre spustenie je potrebne byt v adresari /home/fiitkar/docker-folder a odpalit command: *sudo docker-compose up*

Prepinac -d zabezpeci, ze sa proces spusti na pozadi.

Command na zastavenie a zmazanie docker kontajneru: *sudo docker-compose down*

Nerobit: *sudo docker-compose down -v* , pretoze to zmaze volume, v ktorom su ulozene naindexovane data.

Comment by [Jakub Hlavačka](#) [16/Oct/21]

[Dominik Horvath](#) mozes indexovat data na masine.

0

 **[AMS-30] Redukcia a kategorizacia zoznamu zločinov** Created: 13/Oct/21 Updated: 23/Nov/21 Resolved: 23/Nov/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		


Sprint:	
Rank:	0 i0005r:



[AMS-29] [Získanie dát pre prototyp](#) Created: 13/Oct/21 Updated: 23/Nov/21 Resolved: 26/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
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

Attachments:	 image-20211019-151711.png										
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>AMS-30</td><td>Redukcia a kategorizácia zoznamu zloč...</td><td>Subtask</td><td>Done</td><td>Dominik Horvath</td></tr></tbody></table>	Key	Summary	Type	Status	Assignee	AMS-30	Redukcia a kategorizácia zoznamu zloč...	Subtask	Done	Dominik Horvath
Key	Summary	Type	Status	Assignee							
AMS-30	Redukcia a kategorizácia zoznamu zloč...	Subtask	Done	Dominik Horvath							
Sprint:											
Rank:	0 hzzzzz:r										

Comments

Comment by [Jakub Hlavačka](#) [19/Oct/21]

<https://github.com/JKBGIT1/scraper>

V readme je sposob akym to spustit.

Treba osetrit veci, ktore su v komentoch spider.py ako TODO

Comment by [Jakub Hlavačka](#) [19/Oct/21]

© 2014-2021 by Jakub Hlavačka. All rights reserved. This software is licensed under the GNU GPL license.

Asi to ide moc rychle...

Comment by [Jakub Hlavačka](#) [19/Oct/21]

Zjavne bude potrebne zabezpecit, aby to nerobilo requesty na server, kde uz dostal timeout...

Comment by [Dominik Horvath](#) [23/Nov/21]

Scrapper program bol upravený a rozšírený o stahovanie dát pomocou gNewsParseru vytvoreného v <https://tim8>

[AMS-21] [Vytvorit Slack server](#) Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 11/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0j00004:

Comments

Comment by [Táňa Poláková](#) [11/Oct/21]

Vytvorila som Team na platforme Microsoft Teams

[AMS-20] [Analyza parametrov Google News](#) Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 13/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i00002:

Comments

Comment by [Adam Šípka](#) [13/Oct/21]

<http://books.gigatux.nl/mirror/googlehacks/0596008570/googlehks2-CHP-4-SECT-3.html>

0

 **[AMS-17] [API / scrapovanie](#)** Created: 05/Oct/21 Updated: 10/Oct/21 Resolved: 10/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 output.csv
Sprint:	
Rank:	0 i0001y:z

Comments

Comment by [Dominik Horvath](#) [10/Oct/21]

Vytvoreny custom parser: <https://github.com/Dominik1799/gnewsParser>

Vsetky kniznice, ktore sme nasli, pracovali len ako wrapper nad RSS streamom. Vyuzivali funkcie, ktore su pre nas problem zbytocne, a boli preto velmi pomale.

Vlastny parser uskutocnil 1827 requestov a ziskal 35772 zaznamov za 93 sekund processing time a cca 20 minut realneho casu. Vyhľadavacia query ziskala vsetky clanky v casovom období od 10.10.2016 do 10.10.2021 na klucove slova "Jeffrey Epstein".

0

 **[AMS-16] Získanie vzorky RSS** Created: 05/Oct/21 Updated: 08/Oct/21 Resolved: 08/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Subtask	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 google_news_rss.py
Sprint:	
Rank:	0ji0001y:y

Description

získanie RSS a zistenie, či je to efektívne alebo nie

+ časovanie všetkeho, čo sa vykona

Comments

Comment by [Jakub Hlavačka](#) [08/Oct/21]

RSS vrati 100 clankov na danu query.

Total time: 0.046875

Malo by to byt v sekundach.

Ak sa vyberieme touto cestou, tak bude zjavne potrebne sa dostat k dalsim clankom a nie len prvym 100, ktore to vyhodi...

[AMS-15] [Realny scrapping - vytvorenie vzorky](#) Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 10/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Dominik Horvath
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
	AMS-16	Ziskanie vzorky RSS	Subtask	Done	Jakub Hlavačka
	AMS-17	API / scrapovanie	Subtask	Done	Dominik Horvath
Sprint:					
Rank:	0ji00009:				

Description

Ziskanie vzorky: Jeffrey epstein, poslednych 5 rokov.

[AMS-13] [Zistit, ktore miestnosti su na karticky](#) Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 18/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0001y:x

Description

6-te poschodie - coworking.
Treba sa 5 krát pipnuť a napísať pánovi (neviem komu).
Je tam stále restricted access - treba poznať správnych ľudí.
Vraj sa tam na tím FIIT-WIX škaredo pozerali.
Netreba tam mať rúško?

[AMS-11] [Lofi pouzivatel'skeho rozhrania - main page](#) Created: 05/Oct/21 Updated: 20/Feb/22 Resolved: 20/Feb/22

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 Screenshot_93.png
Sprint:	
Rank:	0 i0001y:zi

Description

<https://www.figma.com/file/zouHpeljsUKIQPuKYDeRB2/Lo-Fi-Prototypes?node-id=0%3A1>

[AMS-10] [Revizia poziadaviek](#) Created: 05/Oct/21 Updated: 22/Nov/21 Resolved: 13/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0j00001:

[AMS-9] [Spísať denník z prvého stretnutia s vedúcim](#) Created: 03/Oct/21 Updated: 22/Nov/21 Resolved: 05/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None


Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0000m:

[AMS-8] [Pripady pouzitia](#) Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 05/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Dominik Horvath
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 Snímka obrazovky 2021-10-05 124128.png
Sprint:	
Rank:	0 i0000q:

Description

1. User opens AMS web
2. User enters search query - a name of a person
3. System returns list of news articles that contain this name
4. user clicks on one of the results
5. system redirects the user to the news source

[AMS-7] [Prieskum ziskavania dat, praca s kniznicou](#) Created: 01/Oct/21 Updated:
23/Nov/21 Resolved: 05/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Adam Šípka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i00001:

Comments

Comment by [Adam Šípka](#) [23/Nov/21]

Podarilo sa mi vytvorit' prvotný script na zber dát, ktorý pracuje s knižnicou GoogleNews a newspaper3k. Avšak nie vždy sa nám podarilo získať samotný text článku, hlavne ak sa jednalo o iný jazyk ako angličtina. Ďalším problémom bola rýchlosť získavania samotných tiel článkov, ktorá celý proces podstatne spomalila.

[AMS-6] [Specifikacia poziadaviek](#) Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 03/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Hlavačka
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Sprint:	
Rank:	0 i0000j:

Comments

Comment by [Jakub Hlavačka](#) [03/Oct/21]

https://docs.google.com/document/d/18XLz0RXSFB50VAKdrydNxtY9zznVDNJK0wyjOS_NG0s/edit?usp=sh

[AMS-5] [Vysoka architektura](#) Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 03/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None


Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	David Silady
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 ams_high_arch.png
Sprint:	
Rank:	0j0000n:

[AMS-4] [Zaobstarat stroj v škole](#) Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 13/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Táňa Poláková
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 virtualne-stroje-info2021.txt
Sprint:	
Rank:	0 i00000:

Description

- VM na fiit(zdarma)
- 1GB Ram, 1CPU, min 50 GB - 100GB disk
- nie FreeBSD, ale skor Ubuntu

Comments

Comment by [Táňa Poláková](#) [03/Oct/21]

Kontaktovala som Ing. Juraja Petrika (5731@is.stuba.sk), ktorý by podľa úvodnej prezentácie mal mať na starosti virtuálne stroje.

Comment by [Táňa Poláková](#) [05/Oct/21]

[virtualne-stroje-info2021.txt](#)

Žiadosť o SSH key bola odoslaná na meno polakova18

[AMS-3] [Zoznam trestnych cinov, ktore budu pouzite ako queries na google news](#) Created: 01/Oct/21 Updated: 22/Nov/21 Resolved: 04/Oct/21

Status:	Done
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Hlavačka	Assignee:	Jakub Müller
Resolution:	Done	Votes:	0
Labels:	None		
Remaining Estimate:	Not Specified		
Time Spent:	Not Specified		
Original estimate:	Not Specified		

Attachments:	 list_of_crimes.txt
Sprint:	
Rank:	0 i0000t:

[AMS-200] [Prerobit' developerské prostredie](#) Created: 24/Mar/22 Updated: 29/Mar/22

Status:	TO REVIEW
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	1 day, 30 minutes		
Original estimate:	Not Specified		

Story point estimate:	8
Sprint:	AMS Sprint 8
Rank:	0 i000po:

[AMS-193] [Prepojenie servisov](#) Created: 24/Mar/22 Updated: 26/Mar/22

Status:	TO REVIEW
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Dominik Horvath
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	3 hours		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0 i000pl:

Description

prepojiť komponenty na novú infraštruktúru

prepísať IPčku

prepísať usera do elasticu

po update connection urobiť aj update repozitáru so správnymi hodnotami

[AMS-154] [Článok IIT.SRC](#) Created: 19/Feb/22 Updated: 25/Mar/22

Status:	TO REVIEW
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Táňa Poláková	Assignee:	Adam Šípka
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	Not Specified		

Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0 i000pm:

Description

Deadline 31. 03. 2022

https://www.fiit.stuba.sk/sk/aktualita/iitsrc2022-call.html?page_id=6154

Comments


Comment by [Adam Šípka](#) [22/Mar/22]

<https://www.overleaf.com/project/6237756a34fe8d0dcab3b312>

[AMS-184] Implementovať spôsob pamätania si usera Created: 10/Mar/22 Updated: 24/Mar/22

Status:	TO TEST
Project:	Adverse Media Screening
Components:	None
Affects versions:	None
Fix versions:	None

Type:	Story	Priority:	Medium
Reporter:	Jakub Müller	Assignee:	Táňa Poláková
Resolution:	Unresolved	Votes:	0
Labels:	None		
Remaining Estimate:	0 minutes		
Time Spent:	5 hours		
Original estimate:	Not Specified		

Attachments:	 image-20220313-143009.png
Story point estimate:	5
Sprint:	AMS Sprint 8
Rank:	0 i0012s:i

Description

Napr. cez hook

Comments

Comment by [Táňa Poláková](#) [13/Mar/22]

User sa pamätá cez "useContext" a "setContex" v rámci modulu react.

<https://reactjs.org/docs/context.html>

Boli vytvorené dva nové súbory:

Utils/UserContext.js - tu sa definuje defaultný stav pre používateľské údaje, čiže prázdny objekt a taktiež metóda useUser, ktorá umožní jednotlivým komponentom prístup k používateľským údajom.

Components/UserProvider.jsx - slúži ako správca používateľa. Tento komponent bol pridaný do komponentu App a vďaka tomu môže každý child komponent prístupit' k používateľským údajom, keďže vracia useContext definovaný v predchádzajúcom súbore.

Logika prihlásenia:

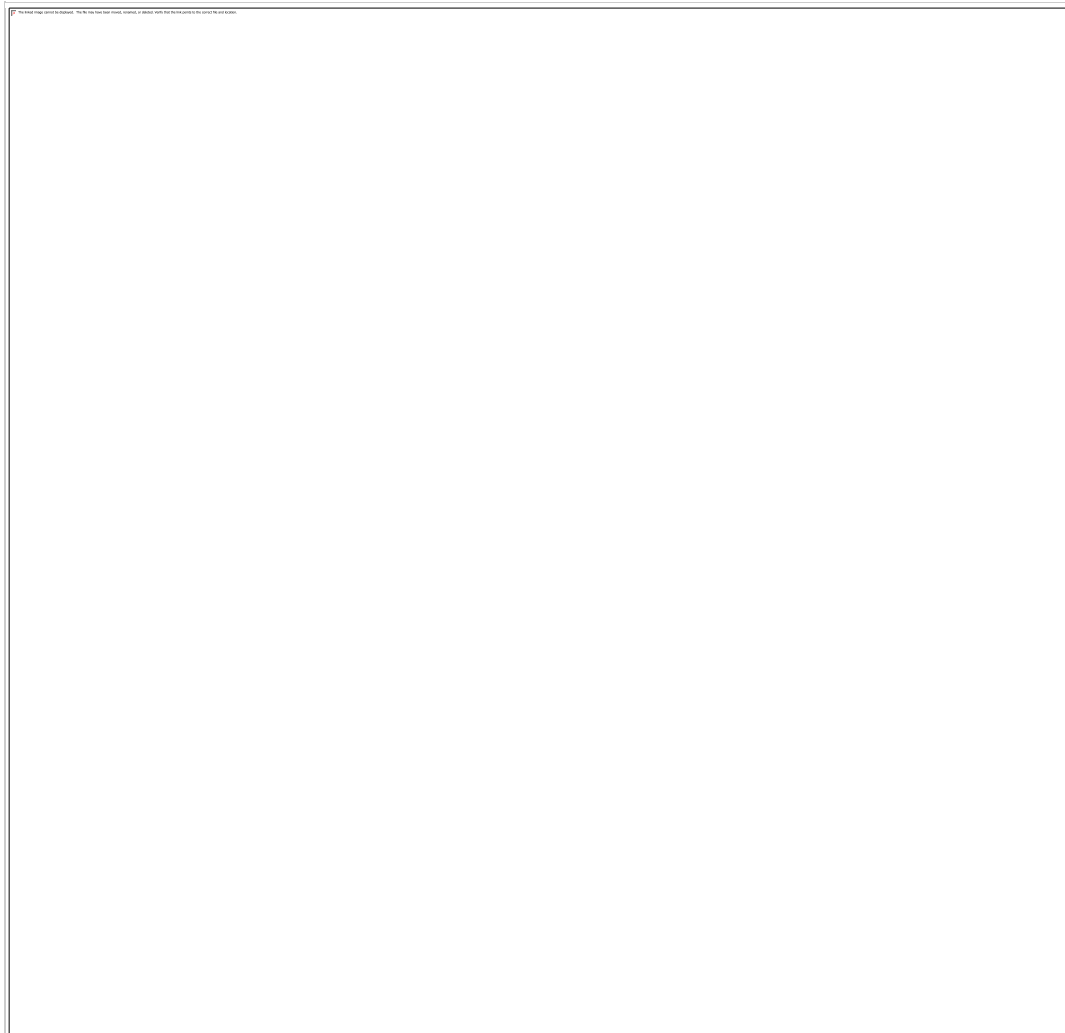
```
import { useUser } from "../Utils/UserContext";

...
const {setAccessToken} = useUser();

const onSubmit = (user_credentials) => {
  token = apiCall(user_credentials);
  setAccessToken(token);
}
```

Keďže sa token zmení, automaticky sa vyplní aj kontext pre používateľa:

Implementácia v Components/UserProvider.jsx



Logika odhlásenia:

```
const onLogout = () => {  
  setAccessToken(null);  
};
```

Comment by [Táňa Poláková](#) [20/Mar/22]

Na základe spoločnej konzultácie bola logika prihlásenia a odhlásenia presunutá do UserProvider. Viac info na stránke dokumentácie.

Generated at Sun Apr 03 17:53:56 UTC 2022 by Tana Polakova using Jira 1001.0.0-SNAPSHOT#100197-sha1:666e16473a14382ff06c420b48977a82c4cbb24c.