

Scraper: Bratislava - faktury

Scraper: Bratislava - faktúry

Morph.io	https://morph.io/soit-sk/bratislava_faktury
GitHub	https://github.com/soit-sk/bratislava_faktury
stav	⚠️ funkčný ale nedokončený
TODO	❌ opraviť zle vyparované a/alebo nevyparované údaje: https://github.com/soit-sk/bratislava_faktury/issues/1

Popis, motivácia

Prehľad dokumentov z webu HI. mesta Bratislava je použiteľný napr. pre potreby PR BA v kombinácii s [OpenSpending.org](https://www.openspending.org/) a pod.

Konkrétne dopĺňujúce informácie m zmlúvam:

- vstupy: podstránky pod http://www.bratislava.sk/register/vismo/zobraz_dok.asp?id_org=700026&id_ktg=1001&p1=15331
 - kúpne zmluvy: http://www.bratislava.sk/register/vismo/zobraz_dok.asp?id_org=700026&id_ktg=1154&archiv=0&p1=15331
 - zmluvy na stavebné práce: http://www.bratislava.sk/register/vismo/zobraz_dok.asp?id_org=700026&id_ktg=1125&archiv=0&p1=15331
 - ...
- výstup: jedna tabuľka
 - kategória zmlúv (id_ktg zo zdrojového URL)
 - id HTML stránky: http://www.bratislava.sk/register/vismo/dokumenty2.asp?id_org=700026&id=24396&p1=15331
 - id PDF dokumentu: http://www.bratislava.sk/register/VismoOnline_ActionScripts/File.ashx?id_org=700026&id_dokumenty=24397
 - dátum
 - titulok
 - popis
 - URL na PDF

Nedokončené

Hneď v niekoľkých prvých riadkoch je vidno ešte nedokončené veci:

Miesto	PSČ	Cena [EUR]	Názov	Interné čís.	Ulica	Predmet fakturácie	Schválil Skratka	IČO	Schválil Skratka Schválil	Číslo
Považská Bystrica	017 01	6 076.00	SAROUT E s.r.o.	1130008869	Mládežnícka 326	dopravné značenie	Eva Elefantová Zamestnanec	316064 58	Eva Elefantová Zamestnanec Eva Elefantová	313267
821 06	Kazanská 48	radu kri žovatiek, 12 /2013	4 162.32	1130008867	35758805	servis a údržba televízneho doh	Eva Elefantová Zamestnanec Eva Elefantová	NOPE a.s.	Eva Elefantová Zamestnanec Eva Elefantová Eva Elefantová Zamestnanec	40550413

K tomuto som otvoril issue: https://github.com/soit-sk/bratislava_faktury/issues/1 .