

About The crawled data

To crawl data there is python script in the "CrawledData\app\Python" folder in the Laravel Project. With this script you can connect to the <https://www.rrpcanada.org/#/available-products-services> site and get list of data from <https://api.rrpcanada.org/supplies> API .

After that scraped data inserted to collection in the Mongodb Database. The collection name is TheCrawledData .

In the Laravel project there is model "CrawledData" and controller "CrawledDataController" to select data from Mongodb collection and to show data on the welcome page of Laravel.

To Config

- download php mongodb driver form <https://pecl.php.net/package/mongodb/1.10.0/windows>
- copy php_mongodb.dll to C:\xampp\php\ext
- Add the following line to your php.ini file: extension=php_mongodb.dll
- Install jenssegers by composer require jenssegers/mongodb

control the speed of scraping hits within this Python script:

When the user send response to view Laravel page at the first old data in the collection will be deleted and after that data scrap by python script and insert to database and show to user.

Why: because if we want to scrap on the period of time for example every 10 seconds the goal site maybe downgrade or run anti scrap filter, so , I decided to use scraper every time the webpage load and show new data to user.