

Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут ім. Ігоря Сікорського”

Факультет прикладної математики
Кафедра програмного забезпечення комп’ютерних систем

Лабораторна робота № 1
з дисципліни «Бази даних. Частина 2»
«Вивчення базових операцій обробки XML-документів»

Виконав:

студент групи КП-73

Булаєвський Ігор Олегович

Перевірив:

Київ - 2020

Метою роботи є здобуття практичних навичок створення програм, орієнтованих на обробку XML-документів засобами мови Python.

Завдання роботи полягає у наступному:

1. Виконати збір інформації зі сторінок Web-сайту за варіантом.
2. Виконати аналіз сторінок Web-сайту для подальшої обробки текстової та графічної інформації, розміщеної на ньому.
3. Реалізувати функціональні можливості згідно вимог, наведених нижче.

Екранні форми:

```
Select a task to do

1 - Crawl bigmir.net
2 - Find page from bigmir.net with the smallest number of images
3 - Crawl sokol.ua
4 - Create XHTML table with fridges from sokol.ua
5 - Exit
```

Рис 1. Початкове меню

```
2020-02-27 21:31:13 [scrapy.core.engine] INFO: Spider closed (Maximum visited pages number exceeded)
bigmir.net was crawled, results are saved to output/bigmir.xml
Press ENTER to continue...
```

Рис 2. Результат виконання 1 пункту меню

```
Page with the smallest number of images: https://www.bigmir.net/ (22 images)
Press ENTER to continue...
```

Рис 3. Результат виконання 2 пункту меню

```
2020-02-27 21:33:11 [scrapy.core.engine] INFO: Spider closed (finished)
sokol.ua was crawled, results are saved to output/sokol.xml
Press ENTER to continue...
```

Рис 4. Результат виконання 3 пункту меню

```
XHTML table was created, results are saved to output/table.xhtml
Would you like to view it in browser? [y/n]
```

Рис 5. Результат виконання 4 пункту меню

Image	Name	Price	Description
	ZANUSSI ZRT23100XA	7999 грн	Тип: Двухкамерный Тип управления: Механическое Общий полезный объем, л: 228 Морозильное отделение: сверху Система охлаждения: статическая Система разморозки: капельная Класс энергопотребления: A+ Полки: ударопрочное стекло
	ELECTROLUX EJ2801AOW2	8299 грн	Тип: Двухкамерный Тип управления: Механическое Общий полезный объем, л: 265 Морозильное отделение: сверху Система охлаждения: статическая Система разморозки: капельная Класс энергопотребления: A+ Полки: ударопрочное стекло
	ZANUSSI ZRX51100WA	4299 грн	Тип: Однокамерный Тип управления: Механическое Общий полезный объем, л: 42 Морозильное отделение: сверху Система охлаждения: статическая Класс энергопотребления: A+ Полки: металлическая решетка

Рис 6. Візуалізація створеної XHTML таблиці