



Міністерство освіти і науки України
Національний технічний університет України
“Київський політехнічний інститут імені Ігоря Сікорського”
Факультет інформатики та обчислювальної техніки
Кафедра інформаційних систем та технологій

Лабораторна робота №4
**ШАБЛОНИ «SINGLETON»,
«ITERATOR», «PROXY», «STATE»,
«STRATEGY»**

Варіант 11

Виконала
студентка групи ІА – 14:
Літвін Юлія

Перевірив:
Мягкий М. Ю.

Завдання:

1. Ознайомитися з короткими теоретичними відомостями.
2. Реалізувати частину функціоналу робочої програми у вигляді класів та їхньої взаємодії для досягнення конкретних функціональних можливостей.
3. Застосування одного з розглянутих шаблонів при реалізації програми

Варіант:

..11 Web crawler (proxy, chain of responsibility, memento, template method, composite, p2p)

Веб-сканер повинен вміти розпізнавати структуру сторінок сайту, переходити за посиланнями, збирати необхідну інформацію про зазначений термін, видаляти не семантичні одиниці (рекламу, об'єкти javascript і т.д.), зберігати знайдені дані у вигляді структурованого набору html файлів вести статистику відвіданих сайтів і метадані.

Хід роботи

Замісник — це структурний патерн проектування, що дає змогу підставляти замість реальних об'єктів спеціальні об'єкти-замінники. Ці об'єкти перехоплюють виклики до оригінального об'єкта, дозволяючи зробити щось *до* чи *після* передачі виклику оригіналові.

DouForumScraper:

Цей клас розширює **BaseForumScraper**.

У ньому є методи, такі як `_scrape_comments_count`, `_scrape_author_info`, `_save_to_html`, `_scrape_page` та `scrape_forum`, які виконують різні завдання скрапінгу.

DouForumScraperProxy:

Цей клас виступає як проксі для **DouForumScraper**.

Він також реалізує той самий інтерфейс, що і **DouForumScraper**, але делегує фактичну роботу збору даних екземпляру **DouForumScraper**.

Цей клас відповідає за додаткові завдання, такі як логування запитів на завантаження.

Методи `download_results_file`, `get_parser_stats`, `scrape_forum` та `undo_scrape` фактично є проксі до відповідних методів у **DouForumScraper**.

Клас `DouForumScraperProxy` обгортає екземпляр **DouForumScraper**.

Він перехоплює виклики методів, таких як `download_results_file`, `get_parser_stats`, `scrape_forum` та `undo_scrape`.

Перед делегуванням виклику реальному **DouForumScraper** виконуються додаткові дії, такі як логування.

Це дозволяє проксі контролювати доступ до справжнього **DouForumScraper** та додавати функціональність без модифікації самого скрапера.

```
class DouForumScraperProxy(ScraperComponent):
    def __init__(self, url, headers):
        self.scraper = DouForumScraper(url, headers)

    4 usages (4 dynamic)
    def download_results_file(self, filename, download_path):
        logger.info(f"Download request: filename={filename}, download_path={download_path}")
        result = self.scraper.download_results_file(filename, download_path)
        logger.info(f"Download result: {result}")
        return result

    4 usages (4 dynamic)
    def get_parser_stats(self):
        logger.info("Requesting parser stats")
        result = self.scraper.get_parser_stats()
        logger.info(f"Parser stats: {result}")
        return result

    4 usages (4 dynamic)
    def scrape_forum(self, output_filename, tag, start_date=None, end_date=None):
        result = self.scraper.scrape_forum(output_filename, tag, start_date=start_date, end_date=end_date)
        return result

    2 usages (2 dynamic)
    def undo_scrape(self):
        self.scraper.undo_scrape()
```