

ПРОГРАМНИЙ ЗАСІБ ДЛЯ ЗБИРАННЯ ТА КЛАСИФІКАЦІЇ ДАНИХ ДИНАМІКИ РОЗВИТКУ ТЕРИТОРІЙ

Броварник О.О., Бреславський Д.В.

Національний технічний університет

«Харківський політехнічний інститут», м. Харків

Задачі соціальної динаміки, що визначають стан та тенденції розвитку різних соціумів, останні часи привертають увагу дослідників. На цьому шляху одним з перспективних трендів є поєднання даних соціально-гуманітарних наук з методами математичного моделювання та інформаційних технологій. Одним з найважливіших засобів при цьому є використання останніх для збирання великого обсягу даних, що описують фактологічну базу досліджень. Така згенерована база може використовуватись як для безпосередньої статистичної обробки даних, що в ній містяться, так й для двох інших важливих наукових завдань. Перше з них – це використання створеної бази даних для висування гіпотез щодо прогнозування окремих трендів розвитку соціальних систем. Друге – застосування фактологічної інформації для верифікації розроблених методів математичного моделювання та бенчмаркінгу розроблених для нього програмних засобів.

Доповідь присвячено опису розробленого програмного засобу, реалізованого у вигляді вебзастосунку. Спроековано структуру його бази даних, що включає можливість опису соціально-економічної інформації у контексті її історичного змінювання. Розроблений вебзастосунок надає можливість переглядати, редагувати та додавати нові дані. На сучасному, першому етапі розробки, вебзастосунок містить інформаційну базу щодо розташування та термінів утворення міст України.

Для створення вебзастосунку обрано мову PHP з використанням фреймворку Laravel. Для зберігання даних використовується система управління базами даних PostgreSQL, що дозволяє ефективно структурувати зібрану інформацію за допомогою прийнятої схеми бази даних.

Для первинного наповнення бази даних використано алгоритми автоматичного збирання та обробки інформації про тисячі населених пунктів України. Використано мову програмування Python, бібліотеки BeautifulSoup та Requests.

Створений алгоритм дає можливість проаналізувати HTML-сторінки онлайн-енциклопедії «Вікіпедія» та інших ресурсів й зібрати інформацію про доступні населені пункти без значних часових та трудових витрат. HTML-сторінки відрізняються за структурою, тому для пошуку необхідної інформації використано декілька регулярних виразів, які значно збільшили кількість зібраної інформації. Проведено тестування бази даних та вебзастосунку у цілому показало його задовільну працездатність.

Отримані та структуровані дані застосовано для побудови інформаційної картини урбаністичного розвитку України, що буде використано для розробки математичних моделей його динаміки.