

## 18. МЕТОД ПЕРЕНЕСЕННЯ ДАНИХ З HTML-ДОКУМЕНТА В БАЗУ ДАНИХ

Лозовська Д.С. Науковий керівник – доц. каф. “Комп’ютерні інтелектуальні системи та мережі”, к.т.н Мартинюк О.М.

При переведенні статичного сайту в динамічний, в основі якого лежить база даних (БД), виникає необхідність перенесення інформації з html-документів в БД.

Для цього можна використовувати метод аналізу структури html-документа. Якщо у нас є безліч html-документів з однаковою структурою, наприклад сторінки з інформацією про кафедри, то має сенс застосовувати викладений метод. Спочатку html-документ можна розділити на теги (Т) і дані або контент (С) і отримати структуру документа у вигляді послідовності, що складається з Т і С. Вручну прописується шаблон html-документа і відповідно до цього шаблону з кожного документа вибирається контент. Шаблон може описувати або структуру таблиці, або просто абзаци. Вибрані дані розміщуються в txt-файл і розділяються табуляцією. Потім з txt-файлу формується файл з sql-запитом на створення таблиці з числом полів відповідним числу стовпців в початковій таблиці. При необхідності число стовпців в таблиці, що створюється в БД, можна змінювати в програмі. Тип даних всіх стовпців задається як varchar, тобто текстові дані. Порожні ячейки заповнюються символами \N. Отриманий документ запускається на виконання в БД і, таким чином, в БД створюється таблиця з даними, які були отримані з html-документа.

Як розвиток методу передбачається автоматизувати процес створення шаблону html-документа. На сьогоднішній день актуальність методу полягає також у тому, що при необхідності зібрати інформацію про підрозділи (наприклад, кафедри) або будь-яку інформацію з різних джерел в текстових doc-документах в таблиці, легко потім ці файли переконвертувати в html-формат і поновити або занести зібрану інформацію в БД.