

Практичне заняття № 4

Тема: Моделі процесів вилучення, обробки даних, зберігання, представлення і використання в інформаційних системах

Мета заняття:

Ознайомлення студентів з основними моделями та технологіями процесів вилучення, обробки, зберігання, представлення і використання даних в інформаційних системах. Формування навичок застосування цих процесів в умовах реальних інформаційних систем.

Основні питання заняття:

1. Процеси вилучення даних:

- Методи збору та організації даних.
- Технології вилучення даних із різних джерел (бази даних, текстові файли, веб-сервіси).

2. Обробка даних:

- Алгоритми та методи обробки даних (фільтрація, агрегація, аналіз).
- Підготовка даних для зберігання та подальшого використання.

3. Зберігання даних:

- Моделі зберігання даних (реляційні, нереляційні бази даних).
- Системи управління базами даних (СУБД).

4. Представлення даних:

- Формати представлення даних (таблиці, графіки, діаграми).
- Методи візуалізації даних.

5. Використання даних в інформаційних системах:

- Процеси доступу до даних і їх аналіз у реальному часі.
- Програмні інтерфейси для роботи з даними (API).

6. Зв'язок між процесами вилучення, обробки, зберігання, представлення та використання даних:

- Взаємодія етапів в інформаційних системах та забезпечення ефективності їх роботи.

Практичні завдання:

1. Моделювання процесу вилучення даних:

- Студенти створюють модель процесу збору даних з різних джерел (наприклад, веб-скрапінг або підключення до бази даних) з використанням мови Python або іншої мови програмування.

2. Обробка даних:

- Студенти отримують набір даних (наприклад, CSV або JSON файл) і виконують базову обробку даних: очищення, фільтрацію, агрегацію.

3. Створення бази даних для зберігання даних:

- Використовуючи СУБД (наприклад, MySQL або SQLite), студенти створюють структуру бази даних для зберігання оброблених даних, проєктують таблиці, встановлюють зв'язки між ними.

4. Представлення даних:

- Студенти створюють просту візуалізацію оброблених даних за допомогою бібліотек для візуалізації, таких як Matplotlib або Tableau.

5. Інтерфейс для використання даних:

- Студенти розробляють простий API для доступу до збережених даних за допомогою фреймворку Flask або Django.

6. Інтеграція всіх етапів:

- Студенти поєднують всі етапи в єдину інформаційну систему, демонструючи процес вилучення, обробки, зберігання, представлення та використання даних в одному проєкті.

Очікувані результати:

1. Студенти повинні освоїти основні принципи і технології процесів вилучення, обробки, зберігання, представлення та використання даних в інформаційних системах.
2. Розвинути навички роботи з різними типами даних, їх обробкою та зберіганням.
3. Отримати практичний досвід у створенні інтегрованих інформаційних систем, які охоплюють весь цикл роботи з даними.
4. Навчитись застосовувати різноманітні інструменти для обробки, зберігання та візуалізації даних.

Література:

1. Бен-Гур А., Костенко О.І. «Основи обробки даних в інформаційних системах» — К.: Наукова думка, 2015.
2. Чен І.І. «Методи та технології обробки даних» — М.: МЦНМО, 2016.
3. Ширшов М.Є. «Системи управління базами даних» — К.: Академвидав, 2017.
4. Кузнецов І.А. «Інформаційні системи: теорія і практика» — К.: Вища школа, 2018.
5. Технології обробки даних у сучасних СУБД, статті та ресурси на сайті <https://www.sqlservercentral.com/>.

Ресурси:

- **Python** — для обробки даних і роботи з API: <https://www.python.org>
- **Matplotlib** — бібліотека для візуалізації даних: <https://matplotlib.org>
- **SQLite** — СУБД для створення бази даних: <https://www.sqlite.org>
- **Flask** — фреймворк для розробки API: <https://flask.palletsprojects.com>
- Онлайн-курси на платформах Coursera, edX та Udeму з теми обробки даних та інформаційних систем.