

## **1. Моделі і методи зберігання даних**

Модель зберігання даних - це стратегії структуризації зберігання даних. Такі як:

- -Реляційна модель - зберігання даних у таблицях та відношень між полями таблиць. Основна технологія що використовує даний спосіб це SQL бази даних. Приклади MySQL, PostgreSQL
- -Документо-орієнтована модель - збереження даних як документів. Популярні приклади таких контейнерів є JSON та XML.
- Також є документо орієнтованна бд MongoDB
- Граф - зберігання даних у графах. Приклад Neo4j
- Ключ-значення - збереження даних де до даних доступуються за ключем та отримують його значення.

Метод зберігання даних - це способи зберігання даних. Наприклад:

- База даних(реляційна/NoSQL)
- Внутрішній/зовнішній носій інформації
- Мережевий носій інформації
- Хмарний носій інформації

## **2. Класифікація інформаційних систем і місце серед них інформаційно-пошукових систем**

Інформаційна система - це організована система збору, зберігання, обробки, передачі та представлення інформації. Інформаційні системи можуть включати комп'ютерні системи та мережі, а також людські ресурси, що використовують ці технології.

Загально прийнята класифікація поділяє системи за:

- рівнем або сферою діяльності — державні, територіальні (регіональні), галузеві, об'єднань, підприємств або установ, технологічних процесів.
- рівнем автоматизації процесів управління — інформаційно-пошукові, інформаційно-довідкові, інформаційно-керівні, системи підтримки прийняття рішень, інтелектуальні ІС тощо.
- ступенем централізації обробки інформації — централізовані ІС, децентралізовані ІС, інформаційні системи колективного використання.

- ступенем інтеграції функцій — багаторівневі ІС з інтеграцією за рівнями управління (підприємство — об'єднання, об'єднання — галузь і т. ін.), багаторівневі ІС з інтеграцією за рівнями планування і т. ін.

Інформаційно-пошукові системи - це комплекс програмно-апаратних засобів, призначених для зберігання, організації та пошуку інформації. Вони дозволяють користувачам ефективно знаходити необхідні дані серед великих обсягів інформації за допомогою ключових слів, запитів або інших методів фільтрації.

Інформаційно-пошукові системи відіграють важливу роль серед інформаційних систем, оскільки їх основна роль є пошук інформації вони можуть бути інтегровані в інші системи або об'єднувати декілька різних систем.

### **3. Організація пошуку. Пошукові машини**

Етапи організації пошуку:

- Індексация - присвоєння індексів словам у документі. Спрощує пошук за словами або фразами
- Ранжування - алгоритми що дозволяють визначити релевантність документів до запиту користувача
- Обробка запиту - розпізнавання ключових слів у запиті та використання їх для пошуку необхідних документів

Пошукові машини - це системи, які використовуються для пошуку інформації. Найбільш відомі приклади : Google, Bing. Ці машини створюють великі бази індексів та документів, де документами виступають веб-сторінки

### **4. Створення і типи індексів**

Типи індексів:

- Первинні індекси - основний ключ у запити
- Вторинні індекси - використовуються для фільтрації результатів первинного індексу
- Складні індекси - комбінація декількох індексів в один
- Бітові індекси - зазвичай логічні операції з булевими полями

- Хеш-індекс
- Індекс дерево
- Повнотекстові - використовують зазвичай увесь запит як індекс

Створення індексу відбувається зазвичай для певних значень. Даний індекс зберігає дані впорядковані за необхідним правилом, це надає можливість швидкого доступу до даних за індексом

## **5. Проблеми індексування**

Основні проблеми індексування:

Обмеженість за типом даних - ефективність індексації залежить від типу даних що індексується. Наприклад для повнотекстових індексів проблемою буде велике використання пам'яті для їх збереження.

Оновлюваність - зазвичай при отриманні нових даних потрібно змінити індекс згідно отриманим даним. Це може бути ресурсо затратно при великих обсягах даних.

Використання диску - чи більше даних тим більше місця займають індекси

## **6. Запит до пошукових машин**

Запит - є основним способом взаємодії користувача з пошуковою машиною. Запит може містити в собі тільки ключові слова так і бути складним реченням, що необхідно розбирати семантично.

Зазвичай користувачю на запит віддається відповідь що відповідає або мові його запиту або його географічному положенню або локалізації його пошукової машини.

Можливості модифікації та спеціалізації запиту залежить від пошувої машини, але основними є використання булевих операцій по типу І АБО, \* - оператор для заміщення невідомих слів, зазначення сайту пошуку, зазначення типу документа, "" конкретизація що пошук повинен відбутися чітко за запитом в дужках

## **7. Якість роботи пошукачів**

Основними критеріями оцінки якості роботи пошукачів є: швидкість пошуку, релевантність відповіді запиту, простота використання, безпека та конфіденційність

## **8. Посилальне ранжування (Page Rank)**

Page Rank - алгоритм ранжування веб-сторінок, створений Google. Основний принцип це оцінка важливості на основі кількісної та якісної оцінки посилань що ведуть на цю сторінку, кожне посилання має різний вплив на цей ранг

## **9. Поняття інформації як категорії, дані і знання**

Інформація - це загальна категорія опису фактів, подій, явищ, тощо, що є релевантним для отримувача.

Дані - це необроблений формат представлення інформації. Вони набувають значення після аналізу та інтерпретації

Знання - результат обробки даних та інформації

## **10. Програмне та апаратне забезпечення для організації пошуку інформації в мережі інтернет**

Пошукові сервіси: Google, Bing, тощо. Ці системи використовують для пошуку інформації

Браузери: Chrome, Firefox. Інтерфейс взаємодії з системами

Мережеві магістралі - основні інтернет магістралі що об'єднані в єдину мережу та надають доступ до інтернету

Спеціалізовані засоби - засоби автоматизації взаємодії з пошуковими системами. Наприклад інструменти для створення запитів використовуючи мови програмування

Спеціалізовані ресурси - сайти та бази даних що мають дані які відносяться до певних категорій, наприклад Google Scholar - пошукова машина для наукових робіт.

Сервери для зберігання даних - фізичні носії де збірається інформації.