

## 1. Моделі і методи зберігання даних

*Зберігання даних* — це процес зберігання і організації інформації для подальшого використання.

*Модель зберігання даних* — це концептуальний підхід до представлення та організації даних, який визначає, як дані будуть структуровані, як вони будуть взаємодіяти та як будуть відображені. Модель зберігання даних визначає логічну структуру даних, не зважаючи на те, де і як ці дані зберігаються фізично. Приклади моделей зберігання даних включають реляційні бази даних, NoSQL бази даних, ієрархічні бази даних, графові бази даних тощо.

*Метод зберігання даних* — це конкретна реалізація або технічний підхід до фізичного зберігання даних в реальній системі, що описує, як саме дані будуть зберігатися на рівні операційної системи чи фізичного обладнання. Приклади методів зберігання даних включають в себе використання реляційних баз даних (таких як MySQL, PostgreSQL), NoSQL баз даних (MongoDB, Cassandra), файлових систем, хмарних зберігальних рішень (Amazon S3, Google Cloud Storage) і тощо.

Отже, модель зберігання даних описує логічну структуру та взаємодію з даними, тоді як метод зберігання даних визначає конкретний технічний спосіб, яким ці дані будуть зберігатися та доступні на фізичному рівні. У багатьох випадках обидва аспекти взаємодіють і визначають загальний підхід до зберігання та роботи з даними в інформаційній системі.

Перелік методів та моделей зберігання даних:

- *Реляційні бази даних (RDBMS)* — один з найпоширеніших методів зберігання даних. Дані організовані у вигляді таблиць, які мають рядки і стовпці. Таблиці можуть бути пов'язані за допомогою ключів, і такі системи баз даних використовують мову SQL для запитів і маніпулювання даними.
- *NoSQL бази даних* призначені для роботи з нереляційними даними. Вони можуть використовувати різні моделі даних, такі як ключ-значення, документ-орієнтовані, стовпцеві або графові бази даних. NoSQL бази даних часто використовуються для великих обсягів даних або ситуацій, де структура даних може змінюватися.
- *Файлова система* використовується для зберігання інформації на рівні операційної системи. Дані можуть бути організовані у вигляді файлів та папок.
- *Хмарне зберігання* — зберігання даних може відбуватися в хмарних обчисленнях, де дані зберігаються на серверах, доступ до яких можливий через Інтернет. Це може включати в себе використання різних послуг, таких як Amazon S3, Google Cloud Storage або Microsoft Azure Blob

Storage.

- *Інші моделі* — деякі спеціалізовані моделі зберігання даних включають ієрархічні бази даних, мережеві бази даних, об'єктно-орієнтовані бази даних та інші.
- *Методи зберігання даних на рівні апаратного забезпечення* включають у себе методи зберігання на твердотільних накопичувачах (SSD), жорстких дисках (HDD), оптичних дисках (CD, DVD, Blu-ray) та інших фізичних носіях.

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

*Інформаційні системи (ІС)* — це комплексні системи, які включають в себе сполучення людей, обладнання, програмного забезпечення та процесів для збору, зберігання, обробки, аналізу та надання інформації з метою підтримки прийняття рішень та виконання операцій у різних областях та організаціях.

Зазвичай, інформаційні системи класифікують за призначенням, обсягом обробки даних, рівнем інтеграції та методами обробки даних.

Класифікація інформаційних систем:

### **1. За призначенням:**

1. *Виробничі (операційні) інформаційні системи* — використовуються для автоматизації операцій та обробки транзакцій в організаціях.
2. *Менеджерські інформаційні системи* — спрямовані на обробку інформації для менеджерів та прийняття управлінських рішень.
3. *Експертні системи* — моделюють рішення експертів у певній області і використовуються для надання порад та рекомендацій.

### **2. За обсягом обробки даних:**

1. *Малих систем (Small Scale Systems)* — використовуються в невеликих організаціях або для вирішення обмежених завдань.
2. *Середніх систем (Medium Scale Systems)* — призначені для середніх компаній та можуть вирішувати більший обсяг завдань.
3. *Великих систем (Large Scale Systems)* — використовуються великими корпораціями та урядовими установами для широкомасштабних завдань.

### **3. За рівнем інтеграції:**

1. *Автономні системи* — різні компоненти не пов'язані між собою і працюють незалежно.  
*Інтегровані системи* — різні компоненти пов'язані і спільно працюють.

*Інформаційно-пошукові системи (ІПС)* — це підгрупа інформаційних систем, які спрямовані на збір, зберігання і надання доступу до інформації.

Основні характеристики ІПС:

- Збирання інформації — ІПС забезпечують механізми для збору різноманітної інформації з різних джерел.
- Зберігання — інформацію зберігається таким чином, щоб вона була легко доступною та відновлюваною.
- Пошук інформації — одним з головних функціональних елементів ІПС є можливість ефективного пошуку інформації за допомогою різних критеріїв.
- Надання доступу — ІПС надають користувачам можливість отримувати доступ до необхідної інформації відповідно до їхніх потреб.

Інформаційно-пошукові системи можуть бути загального призначення або спеціалізованими для конкретних областей, наприклад, наукових досліджень, бібліотек, підприємств тощо.

### 3. Організація пошуку. Пошукові машини

*Організація пошуку в Інтернеті* — це складний процес, який включає в себе різноманітні компоненти, щоб забезпечити ефективний та точний пошук інформації для користувачів. Пошукові машини грають ключову роль у цьому процесі, здійснюючи індексацію веб-сторінок та надаючи результати пошуку на основі ключових слів чи фраз, які користувач вводить.

Організація пошуку в Інтернеті включає ряд ключових елементів, які спільно працюють для забезпечення ефективного та точного пошуку інформації. Основні компоненти цього процесу включають:

- Індексація:
  - *Спайдери (пауки)* — автоматизовані програми, які переглядають веб-сторінки та збирають інформацію для подальшої індексації.
  - *Індекс* — структурована база даних, що містить інформацію про те, які слова входять в які сторінки та де вони знаходяться.
- Ранжування результатів:
  - *Алгоритми ранжування* — визначають, як сторінки будуть впорядковані в результатах пошуку. Основні фактори — релевантність та авторитет веб-сайту.
- Обробка запитів:
  - *Семантичний аналіз* — розуміння смислу запитань та їхніх зв'язків.
  - *Локалізація* — надання результатів, які відповідають місцезнаходженню користувача.
- Пошукові функції:
  - *Фільтрація* — виключення небажаних результатів.
  - *Ключові слова* — пошук за конкретними словами чи фразами.
  - Пошук за зображенням, відео, новинами тощо.

*Пошукові машини* — це комп'ютерні програми або сервіси, які призначені для пошуку та виведення користувачам інформації в Інтернеті. Їхнє завдання полягає у скануванні та індексації веб-сторінок, а потім у наданні

результатів пошуку, які найбільше відповідають введеним користувачем ключовим словам чи фразам. Пошукові машини використовують складні алгоритми ранжування для визначення порядку результатів, щоб надати користувачам найбільш релевантну інформацію.

Найвідоміші пошукові машини:

- *Google* — найпопулярніша пошукова система, яка використовує складний алгоритм ранжування PageRank для визначення авторитетності веб-сайтів.
- *Bing* — пошукова система, розроблена Microsoft, заснована на індексі Bing та використовує різні алгоритми ранжування.
- *Yahoo* — хоча багато з її пошукових послуг надаються через Bing, Yahoo має власний внутрішній алгоритм ранжування для окремих продуктів.

Пошукові машини постійно вдосконалюють свої алгоритми для забезпечення користувачам найбільш точних та релевантних результатів пошуку.

#### 4. Створення і типи індексів

*Індекс* — це структура даних, яка прискорює швидкий доступ до записів у базі даних.

*Індекс створюється* на певному полі або наборі полів у таблиці бази даних. Цей індекс містить відсортовані значення цих полів, що дозволяє швидше здійснювати пошук та сортування за цими значеннями. Кожне значення в індексі пов'язане з конкретним записом або записами в реальній таблиці бази даних.

Типи індексів:

- *Унікальний індекс* — гарантує унікальні значення та швидкий пошук записів за цими значеннями.
- *Кластерний індекс* — визначає порядок даних у самій таблиці, забезпечуючи швидкий доступ за індексованим полем.
- *Некластерний (некластеризований) індекс* — не визначає порядок даних в самій таблиці.

Індекси дозволяють ефективно знаходити та отримувати доступ до конкретних записів у базі даних, що сприяє швидкості виконання запитів. Унікальні індекси дозволяють контролювати унікальність значень в індексованому полі. Створення та ефективне використання індексів у базі даних є важливою частиною оптимізації даних та запитів, що дозволяє покращити продуктивність інформаційних систем.

#### 5. Проблеми індексування

Індексування в базах даних та інформаційних системах вносить значний внесок у швидкість виконання запитів та оптимізацію доступу до даних. Проте, існують певні проблеми та виклики, які можуть виникнути при

використанні індексів:

- *Велика кількість індексів* — створення багатьох індексів на різних полях може призвести до значного збільшення обсягу пам'яті, яку вони займають. Необхідно уважно вибирати, на яких полях створювати індекси та уникати надмірного їх використання.
- *Оновлення даних* — індексація може ускладнити операції додавання, оновлення та видалення записів, оскільки потрібно підтримувати консистентність індексів. Системи повинні бути добре налаштовані для оптимізації процесів оновлення та управління індексами.
- *Пам'ять та зберігання* — індексація може займати значну кількість пам'яті, що особливо важливо для великих баз даних. Важливо збалансувати вигоди індексації та її вплив на загальний обсяг зберігання та використання пам'яті.
- *Селективність індексів* — якщо індекс включає в себе низькоселективні дані (значення, які часто повторюються), він може стати менш ефективним. Оптимізація індексів, включаючи аналіз та переоцінку їхньої ефективності, може допомогти вирішити це.
- *Запити на великих обсягах даних* — при обробці великих обсягів даних індекси можуть втрачати ефективність. Використання різних видів індексів та оптимізація бази даних може поліпшити продуктивність.

Загальною метою є балансування вигод від індексації та витрат, які вона може призвести. Дбайливе проектування та налаштування індексів дозволяє оптимізувати продуктивність системи.

## **6. Запити до пошукових машин**

*Запити до пошукових машин* — це ключовий спосіб взаємодії користувача із пошуковими системами для отримання необхідної інформації. Запити можуть бути сформульовані у вигляді окремих слів, фраз, або складатися з декількох критеріїв. Важливо враховувати деякі основні аспекти при створенні запитів:

- *Ключові слова* — користувачі формулюють запити до пошукових машин, використовуючи конкретні та відповідні слова чи фрази, які визначають тему або область їхнього інтересу.
- *Логічні оператори* — при створенні запитів можна використовувати логічні оператори, такі як AND, OR, NOT, для уточнення чи розширення пошуку, залежно від того, які критерії важливі для користувача.
- *Виключення непотрібної інформації* — щоб отримати більш точні результати, можна використовувати від'ємні оператори (наприклад, NOT чи -) для виключення непотрібних елементів.
- *Фільтрація за форматом* — користувачі можуть вказувати формат документів, який вони шукають, використовуючи ключові слова, які вказують на PDF, відео, або інші формати.

- *Фразовий пошук* — запити можуть включати фрази, що заключені в подвійні лапки, для отримання точних відповідей чи результатів, які містять цю фразу у конкретному порядку.
- *Питальні слова* — використання питальних слів у запиті допомагає уточнити або зазначити конкретні запитання, на які користувач сподівається знайти відповідь.
- *Локалізація запиту* — зазначення локальних елементів у запиті дозволяє користувачам отримати інформацію, специфічну для певного регіону чи місцезнаходження.
- *Використання дужок* — користувачі можуть використовувати дужки для групування частин запиту та визначення порядку виконання операцій у складних запитах.
- *Використання сайтового фільтра* — додавання оператора site: до запиту дозволяє отримати результати лише з конкретного сайту чи домену.

## 7. Якість роботи пошукачів

Якість роботи пошукових систем визначається кількома ключовими аспектами, які впливають на здатність системи забезпечувати точні, релевантні та корисні результати для користувачів. Основні фактори, які визначають якість роботи пошукових систем, включають:

- *Релевантність результатів* — здатність пошукової системи видачі результатів, які найточніше відповідають запитам користувачів. Релевантність є критичним параметром, і пошукова система повинна адаптуватися до широкого спектру запитів.
- *Швидкодія* — швидкість, з якою пошукова система може видавати результати користувачам. Важливо для задоволення очікувань користувачів та забезпечення ефективності в реальному часі.
- *Актуальність індексації* — частота та швидкість оновлення індексів, що визначають наявність свіжої інформації у видачі результатів. Важливо для користувачів, які шукають останні новини, події чи інші швидкозмінювані дані.
- *Природна мова та семантичний аналіз* — здатність розуміти природну мову, враховувати семантичний контекст та навіть враховувати схожі та спільні терміни. Дозволяє покращити розуміння користувачних запитів та надавати більш широкий та точний спектр результатів.
- *Мультиплатформеність* — здатність працювати ефективно на різних пристроях та платформах, включаючи комп'ютери, мобільні пристрої та інші гаджети. Важливо для задоволення потреб сучасних користувачів, які використовують різні пристрої для пошуку інформації.
- *Захист від спаму та маніпуляцій* — заходи, які використовуються для попередження поширення неправдивої чи шкідливої інформації. Важливо для забезпечення безпеки та довіри до результатів пошуку.

Загальна якість роботи пошукових систем залежить від успішної інтеграції цих факторів, щоб забезпечити користувачам зручний та ефективний пошук інформації.

## 8. Посилальне ранжування (PageRank)

*Посилальне ранжування*, відоме також як PageRank, є алгоритмом ранжування веб-сторінок, який був розроблений засновниками Google. Цей алгоритм був ключовим фактором, який допоміг Google здобути перевагу над іншими пошуковими системами у ранні роки їхнього існування.

Основні принципи PageRank:

- *Концепція посилань* — PageRank використовує концепцію посилань між веб-сторінками. Якщо сторінка А посилається на сторінку В, то це може вказувати на важливість сторінки В.
- *Вага посилань* — не всі посилання важливі рівнозначно. PageRank враховує вагу сторінок, які посилаються, тобто якщо на сторінку посилається велика кількість інших важливих сторінок, це може позитивно вплинути на рейтинг.
- *Ітераційний процес* — алгоритм PageRank працює ітераційно. Він починає з припущення, що всі сторінки мають однаковий ранг, а потім в кожній ітерації оновлює цей ранг на основі сторінок, які посилаються на неї.
- *Важливість власної сторінки* — PageRank також бере до уваги важливість сторінки самої по собі. Якщо сторінка має важливі посилання, її власний ранг може зрости.
- *Демпфер* — ймовірність того, що користувач буде переходити за посиланням. Введений щоб уникнути нескінченного циклу і гарантувати стійкість алгоритму.

Алгоритм PageRank дозволяє розташовувати сторінки відповідно до їхньої важливості в мережі, враховуючи як кількість посилань, так і їхню важливість. Цей підхід допомагає вивести на передній план авторитетні та релевантні сторінки під час пошуку.

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

Інформація є важливим та всепроникливим елементом в сучасному світі, її розуміння визначає успіх у багатьох сферах діяльності. Інформацію можна розглядати як категорію, що охоплює набір даних, які набувають значення та корисності через обробку та інтерпретацію. Основні аспекти цієї категорії включають:

- *Дані як вихідний матеріал* — дані є сухими фактами та цифрами, які, необроблені, не несуть значення чи корисності, наприклад список чисел: 3, 7, 12, 5.
- *Трансформація в інформацію* — інформація виникає після обробки та

інтерпретації даних, надаючи їм контекст та значення, наприклад середнє значення списку чисел: 6.75.

- Характеристики інформації:
  - *Структура* — інформація має структуру та організацію, які роблять її зрозумілою та придатною для використання.
  - *Контекст* — значення інформації визначається її контекстом та тим, як вона відноситься до конкретного запитання чи завдання.
- Знання:
  - *Значення*: Інформація надає значення, роз'яснення чи освітлення конкретної ситуації чи явища.
  - *Корисність*: Інформація корисна для прийняття рішень, вирішення проблем, розвитку навичок та багато інших аспектів.
- Форми виявлення:
  - *Текст* — слова та речення, що надають інформацію у віртуальній чи паперовій формі.
  - *Графіка* та мультимедіа — зображення, відео та інші візуальні елементи для подачі інформації.

*Знання* — це вищий рівень розуміння, що виникає, коли інформація інтегрується в систему знань особи або організації. Знання передбачає розуміння, висновки та використання інформації для розв'язання завдань та прийняття рішень.

Ієрархія від даних через інформацію до знань — є ключовою у сучасному світі, де обробка та використання інформації та знань є вирішальними для розвитку технологій, науки та бізнесу. Розуміння цих понять визначає успіх у використанні інформаційних ресурсів та прийнятті обґрунтованих рішень.

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

Ефективний пошук інформації в мережі Інтернет визначається не лише активністю користувачів, але й використанням різноманітного програмного та апаратного забезпечення. Існують різні види програм та апаратних засобів, що використовуються для організації ефективного пошуку інформації в Інтернеті.

Програмне забезпечення:

- *Пошукові машини* — забезпечують індексацію та пошук веб-сторінок у мережі, наприклад Google, Bing, Yahoo.
- *Веб-браузери* — надають інтерфейс для перегляду веб-сторінок та виконання пошуку, наприклад Google Chrome, Mozilla Firefox, Microsoft Edge.
- Антивірусне програмне забезпечення — захищає від шкідливих програм та шкідливих веб-сайтів, наприклад Norton, McAfee.
- *Додатки для соціальних мереж* — забезпечують можливість пошуку та



обміну інформацією в соціальних мережах, наприклад Facebook, Twitter, LinkedIn.

- *Плагіни та розширення браузерів* — додають функціональність до браузера, полегшуючи пошук та фільтрацію результатів, наприклад Adblock Plus, Grammarly, LastPass.
- *Системи управління закладками та історією* — допомагають зберігати та організовувати важливі сторінки та інформацію, наприклад Rocket, Raindrop.io.

Апаратне забезпечення:

- *Комп'ютер або мобільний пристрій* — надають основний інструмент для роботи в Інтернеті та використання програм, наприклад ПК, ноутбуки, смартфони.
- *Мережеве обладнання* — забезпечують з'єднання з Інтернетом та розподіл мережевого трафіку, наприклад маршрутизатор, комутатор, модем.
- *Сервери пошукових систем* — забезпечують індексацію та обробку запитів від користувачів, наприклад сервери Google, сервери Bing.
- *Засоби зберігання даних* — забезпечують простір для зберігання індексів та кешів для покращення ефективності пошуку, наприклад жорсткі диски (HDD та SSD), сховища в хмарі.
- *Пристрої введення та виведення* — дозволяють користувачам взаємодіяти з пошуковими системами та результатами, наприклад клавіатура, монітор, принтер.