

Організація баз даних для комплексної автоматизованої обробки інформації в інформаційно-аналітичних підрозділах бібліотек

У статті розглядається питання розробки та ефективної організації баз даних у контексті їх використання в інформаційно-аналітичних підрозділах бібліотек. Наводяться основні вимоги до створення ефективної інформаційної бази даних, розглядаються питання проектування інформаційних систем баз даних. Як приклад системи бази даних розглядається розроблена у Службі інформаційно-аналітичного забезпечення (СІАЗ) автоматична система для збереження, редагування та пошуку спеціалізованої інформації.

Для забезпечення сучасних потреб бібліотек при формуванні інформаційного простору необхідна наявність комплексної автоматизації технологічних процесів, еволюції послуг та перетворень в усіх сферах діяльності бібліотек: ведення каталогів, обробки надходжень, реєстрації читачів, довідково-бібліографічної роботи та ін. У роботі інформаційно-аналітичних підрозділів бібліотек головну роль відіграють засоби для пошуку, обробки та збереження інформації, що сьогодні, в умовах глобальної комп'ютеризації суспільства, мають на меті автоматизувати всі процеси виробництва інформаційно-аналітичних продуктів.

Комплексна автоматизована обробка інформації забезпечує об'єднання в єдиний комплекс усіх технічних засобів, новітніх технологій, методологій і різноманітних процедур щодо обробки інформації. Реалізація комплексної автоматизованої обробки інформації є метою при створенні та організації систем баз даних. Завдяки широкому спектру вбудованих засобів та набору операцій, що можуть використовувати дані системи при обробці, пошуку та збереженні інформації, їх упровадження у процеси роботи організації є неза-мінним та необхідним.

Система баз даних – це комп'ютеризована система збереження однотипових записів. Саму ж базу даних можна розглядати як аналог електронної картотеки, тобто сховище або контейнер для деякого набору файлів даних, які занесені в комп'ютер. Користувачам цієї системи

надається можливість виконувати (або передавати системі запити на виконання) множини різних операцій над такими файлами, наприклад:

- додавати нові порожні файли у базу даних;
- включати нові дані в існуючі файли;
- одержувати дані з існуючих файлів;
- видаляти дані з існуючих файлів;
- змінювати дані в існуючих файлах;
- видаляти файли з бази даних.

Інформаційні бази даних організації включають увесь комплекс статистичних показників, які характеризують діяльність організації в цілому і її підрозділів, а також фактологічний матеріал про всі фактори, які впливають на стан і тенденції розвитку організації. Так, у роботі підрозділів бібліотек, пов'язаних із випуском інформаційно-аналітичних продуктів, інформацією для занесення в базу даних можуть бути дані про відділи підрозділу, показники кількості опублікованих статей та іншої аналітичної продукції у відділах підрозділу, кількість замовників, з якими співпрацює підрозділ, до того ж вказується статистика ресурсів, що використовується, інформація про співробітників, що задіяні у роботі над певним проектом, та ін.

Між фізичною базою даних (тобто даними, що реально зберігаються у комп'ютері) та користувачами системи існує рівень програмного забезпечення, який можна називати по-різному: диспетчер бази даних (database manager), сервер бази даних (database server) або система управління базами даних, СУБД (DataBase Management System – DBMS). Усі запити користувачів на одержання доступу до бази даних обробляються СУБД. Усі засоби додавання файлів (або таблиць), вибірки і оновлення даних у цих файлах або таблицях також надає СУБД. Основне завдання СУБД – надати користувачу бази даних можливість працювати з нею, не заглиблюючись в усі деталі роботи на рівні апаратного забезпечення. Тобто СУБД дає змогу кінцевому користувачу розглядати базу даних як об'єкт більш високого рівня порівняно з апаратним забезпеченням, а також надає в його розпорядження набір операцій, що виражаються в термінах мови високого рівня (наприклад, набір операцій, які можна виконувати за допомогою мови SQL).

Системи баз даних володіють характерними перевагами відносно організованої іншим чином інформації:

- для баз даних характерним є однократне введення та багатократне використання інформації, введена інформація застосовується для вирішення багатьох проблем, забезпечується її багаточільове і сумісне використання;

– бази даних існують незалежно від конкретних прикладних програм, що забезпечує уніфікацію засобів організації даних і незалежність прикладних програм від організації даних;

– базам даних властива структурованість, що відображає певну предметну область;

– бази даних дають змогу встановити мінімально необхідний рівень надлишковості даних (тобто дані не дублюються при їх використанні різними користувачами);

– у базі даних забезпечується дотримання стандартів подання даних, що спрощує їх створення та обслуговування;

– у базах даних забезпечується централізоване управління інформаційними ресурсами, синхронна підтримка даних для всіх додатків, включаючи мови запитів і засоби захисту.

Інформаційна база даних має відповідати таким вимогам:

– структура бази даних має допускати легке розбиття її на складові частини, які розміщуються в окремих вузлах мережі, забезпечувати простоту доступу до будь-якої підбази, захист від несанкціонованого доступу до тих чи інших даних і високу продуктивність у роботі з даними;

– структура інформаційної бази має забезпечувати адекватність змісту зовнішньої (документальної) і внутрішньої (комп'ютерної) форми збереження інформації про об'єкти чи процеси, з якими працює виконавець;

– структура інформаційної бази і схема її розподілення по вузлах локальної обчислювальної мережі має забезпечувати можливість єдиного або одночасного процесу корегування декількох однакових баз даних, що зберігаються у різних вузлах;

– структура інформаційної бази має бути мінімально надлишковою і одночасно зручною для архівування даних.

У теорії проектування інформаційних систем предметну область розглядають у вигляді трьох подань:

– подання предметної області у тому вигляді, як вона реально існує;

– як її сприймає розробник бази даних;

– як вона може бути описана за допомогою символів.

У зв'язку з цим при формуванні архітектури системи бази даних виділяють три рівні:

– *зовнішній рівень (рівень користувача, логічний)*. Він пов'язаний зі способами подання даних для окремих користувачів;

– *внутрішній рівень (фізичний)*. Цей рівень пов'язаний зі способами збереження інформації на фізичних пристроях;

– *концептуальний рівень* є проміжним рівнем між двома першими.

На зовнішньому рівні база даних подається у такому вигляді, у якому її бачить певний користувач. Наприклад, користувач із відділу кадрів може розглядати базу даних як набір записів з інформацією про відділи та співробітників і нічого не знати про записи з інформацією про замовників продукції, що випускається відділами.

У кожного користувача є своя власна мова для роботи із СУБД:

– для прикладного програміста це або одна із відомих мов програмування (наприклад, PL/I, C++ або Java), або спеціальна мова, що використовується в цій системі;

– для кінцевого користувача це або спеціальна мова запитів, або мова спеціального призначення, яка може базуватися на використанні форм та меню, розроблена спеціально з урахуванням вимог користувача.

Мовою, що підтримується практично всіма існуючими сьогодні системами баз даних, є SQL.

Внутрішнє подання бази даних – це подання на низькому рівні всієї бази даних як бази, яка складається з деякої множини екземплярів кожного з існуючих типів внутрішніх записів (записів, що зберігаються). Тобто внутрішнє подання припускає наявність нескінченного лінійного адресного простору. Особливості методів відображення цього простору на фізичні пристрої збереження значною мірою залежать від операційної системи, що використовується.

Концептуальне подання – це подання даних у тому вигляді, якими вони є насправді, а не в тому, якими їх має розглядати користувач у рамках, наприклад, певної мови або апаратного забезпечення, що використовується. Концептуальне подання складається з деякої множини екземплярів кожного з існуючих типів концептуальних записів. Наприклад, воно може складатися з набору екземплярів записів, які містять інформацію про відділи, наборів екземплярів записів, які містять інформацію про замовників, набору екземплярів записів, які містять інформацію про інформаційно-аналітичні продукти, що випускаються, тощо. На концептуальному рівні будується модель предметної області з використанням концептуальної схеми, що в свою чергу використовує концептуальну мову визначення даних. Широко розповсюдженою моделлю концептуального рівня є модель «сутність – зв'язок» або ER-модель (Entity Relationship Model).

При побудові моделі предметної області визначається форма подання інформації. Якісне управління процесами опрацювання електронних документів неможливе без детальної інформації про них. Дані про документи описують різні елементи – як прості, так і складні.

Кожний елемент (об'єкт, явище чи процес) має певні властивості (атрибути). Наприклад, документ описується обсягом, датою створення, має назву, тематику та автора. Відомості можуть бути різних рівнів та мати складну ієрархічну структуру. Дані про автора можуть включати його ім'я та адресу, список його публікацій тощо. У свою чергу деякі властивості можуть становити окремий елемент інформації зі складною внутрішньою структурою. Окремі елементи інформації, які мають власну структуру та властивості, є сутностями, а їх властивості – атрибутами.

Зв'язки між об'єктами в моделі «сутність – зв'язок» можуть мати тип «один до одного», «один до багатьох» або «багато до багатьох».

Охарактеризуємо основні етапи, на які розбивається процес проектування бази даних інформаційної системи.

Концептуальне проектування – збір, аналіз та редагування вимог до даних. Для цього проводиться:

- дослідження предметної області, вивчення її інформаційної структури;

- виявлення всіх фрагментів, кожний з яких характеризується поданням користувача, інформаційними об'єктами та зв'язками між ними, процесами над інформаційними об'єктами;

- моделювання та інтеграція всіх подань.

По завершенню цього етапу будується концептуальна модель, інваріантна структурі бази даних. Часто вона подається у вигляді ER-моделі.

Логічне проектування – перетворення вимог до даних у структури даних. На виході формується СУБД-орієнтована структура бази даних та специфікації прикладних програм. На цьому етапі часто моделюють бази даних, орієнтованих на застосування до різних СУБД, та проводять порівняльний аналіз моделей.

Фізичне проектування – визначення особливостей збереження даних, методів доступу і т. д.

Відмінність рівнів подання даних на кожному етапі проектування (табл. 1):

Відмінність рівнів подання даних на кожному етапі проектування

Концептуальний рівень: – сутності – атрибути – зв'язки	Подання даних для аналітика
Логічний рівень: – записи – елементи даних – зв'язки між записами	Подання даних для програміста
Фізичний рівень: – групування даних – індекси – методи доступу	Подання даних для адміністратора

Основним призначенням систем баз даних є підтримка розробки та виконання додатків баз даних. Тому на високому рівні систему баз даних можна розглядати як систему, яка складається з двох частин – сервера та клієнта.

Сервер – це сама СУБД. Він підтримує всі основні функції СУБД: визначення даних, маніпулювання даними, захист даних, підтримка цілісності тощо. Зокрема, він надає повну підтримку зовнішньому, внутрішньому та концептуальному рівням.

Клієнти – це різні додатки, які виконуються за допомогою СУБД. Такими є як додатки, написані користувачами, так і вбудовані додатки, що надаються постачальниками СУБД.

У зв'язку з поділом системи на ці дві частини з'являється можливість організувати їх роботу на різних комп'ютерах. Тому існує можливість організації розподіленої обробки даних. Розподілена обробка припускає, що окремі комп'ютери можуть бути поєднані деякою комунікаційною мережею так, щоб виконання одного завдання обробки даних можна було розподілити на декілька комп'ютерів цієї мережі.

Прикладом системи бази даних для обробки, збереження та пошуку інформації може слугувати розроблена у Службі інформаційно-аналітичного забезпечення автоматична система для збереження, редагування та пошуку інформації про персоналії – політтехнологів, які відслідковували певні політичні процеси в Україні, та партії, які брали

участь у певних подіях в Україні. Так, інформація відбиралася з різних інформаційних ресурсів протягом тривалого періоду у відділі СІАЗ та зберігалася у мережі. У процесі розробки цієї автоматичної системи була спроектована база даних та сформована концептуальна модель, що включала такі інформаційні об'єкти (множини однотипних сутностей): персону (політтехнолог), стаття, тема (подія, політичний процес), партія. Для кожного інформаційного об'єкта були виділені його атрибути та встановлені зв'язки з іншими об'єктами. Розроблена система базувалася на архітектурі «клієнт – сервер», де програма-клієнт давала змогу вводити нові файли, які містили інформацію про персоналії та партії, що зберігалися в окремому каталозі в мережі, а програма-сервер давала можливість автоматично включати нові файли у базу даних, вилучати існуючі, редагувати записи, занесені у базу даних, та здійснювати пошук у базі за допомогою вбудованих у систему засобів: форм та меню, що дають змогу сформуванню SQL-запит, без знання без посередньо SQL, спеціальних програмних модулів, що реалізують пошук документа за ключовими словами та пошук у документі відповідного фрагмента. Розроблена система дає можливість реалізувати розподілену роботу над проектом: введення нової інформації у систему виконується за допомогою програми-клієнта та може здійснюватися на декількох комп'ютерах декількома працівниками, а безпосереднє наповнення бази даних цією інформацією здійснюється програмою-сервером, що знаходиться на окремій машині.

У такій системі реалізована можливість сортування записів, передбачено захист інформації від несанкціонованого доступу (включення, вилучення та модифікація даних у базі здійснює програма-сервер, до якої має доступ лише адміністратор бази даних) та забезпечується цілісність даних, що контролюється автоматичною перевіркою даних при їх включенні та вилученні із бази.

Список використаної літератури

1. *Дейт, К. Дж.* Введение в системы баз данных [Текст] / К. Дж. Дейт. – М. : Издательский дом «Вильямс», 2005. – 1328 с.
2. *Конноли, Т., Бегг, К., Страчан А.* Базы данных: проектирование, реализация и сопровождение. Теория и практика [Текст] / Т. Конноли, К. Бегг, А. Страчан. – М. : Издательский дом «Вильямс», 2000. – 1120 с.
3. *Системы управления базами данных и базами знаний* [Текст] / под ред. А. Н. Наумова. – М. : Финансы и статистика, 1991. – 352 с.
4. *Тиори, Т., Фрай, Дж.* Проектирование структур баз данных [Текст] : в 2 т. / Т. Тиори, Дж. Фрай. – М. : Мир, 1985. – Т. 1. – 280 с.