

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОСТРОЗЬКА АКАДЕМІЯ»

Мікула М. П., Коцюк Ю. А., Мікула О. М.

ОРГАНІЗАЦІЯ БАЗ ДАНИХ ТА ЗНАНЬ

*Навчальний посібник
для студентів спеціальності
«Комп'ютерні науки»*

Острог
Видавництво Національного університету «Острозька академія»
2021

УДК 004(075.8)
ББК 30
М 65

*Рекомендовано до друку вченою радою
Національного університету «Острозька академія»
(протокол № 2 від 29 вересня 2021 року)*

Рецензенти:

Тішков Б. О., кандидат економічних наук, доцент, завідувач кафедри інформаційних систем ДВНЗ «Київський національний економічний університет ім. В. Гетьмана»;

Мельник А. М., кандидат технічних наук, доцент кафедри комп'ютерних наук Західноукраїнського національного університету;

Новоселецький О. М., кандидат економічних наук, доцент, декан економічного факультету Національного університету «Острозька академія».

Мікула М. П., Коцюк Ю. А., Мікула О. М.

М 65 Організація баз даних та знань: навчальний посібник для студентів спеціальності «Комп'ютерні науки». Острог: Видавництво Національного університету «Острозька академія», 2021. 194 с.

ISBN 978-617-8041-08-3

DOI 10.25264/978-617-8041-08-3

У навчальному посібнику розглянуті головні концепції організації і проектування баз даних та баз знань, основні моделі їх побудови, архітектури СУБД і експертних систем. З урахуванням процесів цифровізації розкриті напрямки розвитку технологій баз даних та баз знань, а також питання, які пов'язані з їх адмініструванням та захистом. Всі головні положення навчального посібника розглядаються з використанням прикладів і ілюструються відповідними графічними матеріалами. Зміст навчального посібника укладений у відповідності з чинною програмою курсу «Організація баз даних та баз знань».

Навчальний посібник призначений для студентів, викладачів та осіб, які цікавляться розробкою баз даних та додатків на основі баз даних, а також може бути використаний в якості основи для курсу лекцій по відповідним дисциплінам.

УДК 004(075.8)
ББК 30

ISBN 978-617-8041-08-3

© Мікула М. П., Коцюк Ю. А., Мікула О. Н., 2021
© Вид-во НаУ «Острозька академія», 2021

ЗМІСТ

ВСТУП	5
-------------	---

СИСТЕМИ БАЗ ДАНИХ

1.1. Числова та нечислова обробки даних	8
1.2. Обмеження файлових систем	12
1.3. Бази даних. Системи управління базами даних	14
1.3.1. Визначення та класифікація	14
1.3.2. Властивості і основні функції СУБД	15
1.4. Архітектура системи управління базами даних	18

МОДЕЛІ ДАНИХ СУБД ЯК ІНСТРУМЕНТАРІЙ ПРЕДСТАВЛЕННЯ КОНЦЕПТУАЛЬНОЇ МОДЕЛІ

2.1. Основні визначення та поняття	22
2.2. СУБД першого покоління	23
2.3. Реляційні СУБД	30
2.4. Об'єктно-орієнтовані СУБД	36
2.5. Об'єктно-реляційні СУБД	38
2.6. Тенденції розвитку та використання СУБД	39

МОВА СТРУКТУРОВАНИХ ЗАПИТІВ SQL

3.1. Загальна характеристика мови SQL	43
3.2. Вибірка даних запитом SQL	44
3.3. Використання реляційних і булевих операторів для створення предикатів	52
3.4. Використання спеціальних операторів в умовах	57
3.5. Узагальнення даних за допомогою агрегатних функцій	64
3.6. Формування виводу у запитах	70
3.7. Запит до багатьох таблиць як однієї	76
3.8. Об'єднання таблиці з собою	79
3.9. Використання підзапитів	85
3.10. Співвіднесені підзапити	90
3.11. Використання оператора EXISTS	94

ЕТАПИ ПРОЕКТУВАННЯ РЕЛЯЦІЙНИХ БАЗ ДАНИХ

4.1. Попередні відомості	98
4.2. Системний аналіз предметної області та опис інформаційних об'єктів	100

4.3. Проектування інфологічної моделі предметної області	104
4.4. Вибір СУБД	109
4.5. Даталогічне проектування	109
4.6. Фізичне проектування	110

ІНФОЛОГІЧНЕ ПРОЕКТУВАННЯ СТРУКТУРИ БАЗИ ДАНИХ

5.1. Попередні відомості	116
5.2. Інфологічне проектування на основі алгоритму нормалізації	119
5.3. Інфологічне проектування на основі семантичного моделювання	125

РЕАЛІЗАЦІЯ ПРОЕКТУ ЗАСОБАМИ СУБД

6.1. Основні об'єкти СУБД Access	132
6.2. Порядок створення БД у MS Access	134
6.3. Типи даних полів	136
6.4. Способи створення таблиць	138
6.5. Використання зовнішніх даних – імпортування таблиць	139
6.6. Створення схеми бази даних	140

АДМІНІСТРУВАННЯ БАЗ ДАНИХ. ЗАХИСТ БАЗ ДАНИХ

7.1. Адміністрування бази даних	143
7.2. Захист даних у базах даних	144
7.3. Захист даних від несанкціонованого доступу в СУБД Access	149

ОРГАНІЗАЦІЯ СХОВИЩ ДАНИХ

8.1. Загальні положення	154
8.2. Архітектура сховищ даних	155
8.3. Структура даних сховища даних	156

ОРГАНІЗАЦІЯ БАЗ ЗНАНЬ ЕКСПЕРТНИХ СИСТЕМ

9.1. Експертні системи. Основні поняття та архітектура	163
9.2. Класи експертних систем та формалізація бази знань	166
9.3. Експертні системи підтримки прийняття рішень	169

ДОДАТОК 1. Теми лабораторних робіт	173
---	------------

ДОДАТОК 2. Варіанти індивідуальних завдань	180
---	------------

ДОДАТОК 3. Словник термінів	185
--	------------

СПИСОК ЛІТЕРАТУРИ	191
--------------------------------	------------

ВСТУП

У наш час складно знайти підприємство чи компанію без інформаційної системи (ІС), побудованої на використанні бази даних (БД). Приклади таких ІС: облік кадрів, бухгалтерський облік, складський облік, електронний документообіг і багато-багато інших. Ці системи обслуговують в першу чергу поточну, вживається також термін оперативну, діяльність окремих підрозділів (бухгалтерію, кадри, склад, маркетинг і т.д.). Для малих підприємств чи окремих підприємців вони залишаються основним засобом ведення обліку для оцінки поточного стану. Для великих підприємств або корпорацій традиційні інформаційні системи з БД – це також джерела даних для оперативного управління. Але виникла потреба в узагальненій інформації, отримуваної від різних ІС з БД, придатної для аналітичного аналізу. Для рішення цієї проблеми була розроблена і реалізована концепція Сховищ Даних, важливого засобу якісно іншого класу інформаційних систем – систем підтримки прийняття рішень (СППР). До СППР також відносяться системи, які використовують бази знань – експертні системи.

Основне призначення навчального посібника полягає у системному відображенні основних питань, які зустрічаються при проектуванні реляційних баз даних, а також в розгляді організації сховищ даних (СД) та баз знань (БЗн) інформаційних систем підтримки прийняття рішень.

Існують навчальні видання, присвячені як окремим питанням теорії та практики використання БД та СД, так і такі, що претендують на охоплення технології БД або СД в цілому. Останні відрізняються великим обсягом і докладним викладом теорії – це література, розрахована на професіоналів. Є книги із серії «для чайників». Такі видання не сприяють збільшенню кількості фахівців в області баз даних. Даний посібник розрахований на непідготовлених читачів, зміст і рівень представлених матеріалів відповідає вимогам стандарту до дисципліни «Організація баз даних і знань» для спеціальності 122 «Комп'ютерні науки».

Особливості посібника:

– теорія баз даних та баз знань викладається лише в тому обсязі, який необхідний як основа, база для розуміння особливостей цих технологій. Виклад теоретичних питань супроводжується великою кількістю прикладів, що полегшує сприйняття матеріалу;

– всі аспекти технології БД та БЗ розглядаються з точки зору їх застосування на практиці;

– фактично відсутня орієнтація на якусь конкретну систему управління базами даних (СУБД) при розгляді питань проектування БД, при цьому надаються відомості про типові можливості СУБД;

– посібник орієнтований на навчальний курс, що передбачає виконання лабораторних робіт та індивідуального завдання і містить приклади завдань.

Навчальний посібник заснований на матеріалах лекцій та практичних занять, які протягом кількох років проводяться авторами на Економічному факультеті Національного університету «Острозька академія».

Матеріал книги розбитий на глави, кожна з яких відповідає одній темі. Кожна глава представляє собою більш-менш завершений фрагмент, який можна вивчати окремо від інших глав.

В *Главі 1* розглядаються структури даних інформаційних задач, визначаються недоліки файлових систем при використанні їх в якості БД, наводяться визначення основних термінів, а також загальні відомості про СУБД даних та її архітектуру.

Глава 2 присвячена опису моделей даних, використовуваних в СУБД. Основна увага приділяється реляційній моделі, але також наводяться відомості і про інші моделі: про мережні, ієрархічні, об'єктно-реляційні та об'єктно-орієнтовані. Проаналізовано переваги та недоліки цих моделей. Визначені тенденції розвитку та використання СУБД.

Глава 3 займає введення в мову SQL – структуровану мову запитів до баз даних. Синтаксис команд і приклади використання відповідають стандарту SQL-92, який підтримується більшістю реляційних СУБД. У цьому розділі основна частина глави присвячена можливостям команди SELECT: використання операторів, предикатів, функцій агрегування, запитів на кількох таблицях, підзапитів, самоз'єднанню таблиць.

Глава 4 стосується питань проектування реляційних баз даних. Викладена методологія проектування реляційних баз даних, наведені етапи проектування та аналізується їх зміст. Теоретичні положення супроводжуються практичним виконанням типових робіт, відповідних етапу. Підсумком вивчення навчального матеріалу являються знання методології проектування реляційної БД і досвід використання її на практиці – розроблена схема бази даних «Бібліотека».

Глава 5 присвячена інфологічному проектуванню структури БД на основі алгоритму нормалізації та технології семантичного моделювання. Значна частина матеріалу присвячена нормалізації відношень методом нормальних форм. Описується застосування методу сутність-зв'язок (entity-relation method, ER-метод), розглядаються правила перетворення ER-діаграми в схему БД, визначення первинних ключів і типів даних атрибутів. Опис супроводжується прикладами, які ілюструють повний цикл проектування реляційної бази даних від опису предметної області до розробки ІЛМ та даталогічної моделі.

Глава 6 охоплює питання, пов'язані зі створенням фізичної бази даних засобами СУБД ACCESS: послідовність дій, виконуваних для створення закінченого функціонального застосунку БД предметної області; основні об'єкти СУБД; використовувані типи даних полів; способи створення таблиць; використання зовнішніх даних та створення схеми БД.

Глава 7 присвячена адмініструванню та захисту баз даних. Розглянуті основні завдання адміністрування бази даних, методи забезпечення цілісності даних, засоби забезпечення безпеки та фізичного захисту даних.

В Главі 8 викладається організація сховищ даних. Розглянуті питання архітектури та структури СД. Розкрита сутність розмірної (Dimensional) – методології проектування сховищ даних, та способи структурування даних «зірка» та «сніжинка».

Глава 9 стосується питань архітектурної організації експертних систем та баз знань. Викладені в загальній формі методи формалізації бази знань експертних систем, детально розглянута технологія формалізації бази знань на основі продукції. Розкрита структура експертної системи підтримки прийняття рішень.

Наприкінці кожної глави надається список запитань для самоконтролю до теми, завдання для підготовки до лабораторних робіт та список використаних джерел.

У додатку 1 наведені теми лабораторних робіт і навчальні питання для опрацювання, *у додатку 2* – теми індивідуальних завдань.

Матеріал посібника охоплює найбільш важливі питання організації і проектування баз даних та організації баз знань і орієнтований на практичне використання даних технологій.

СИСТЕМИ БАЗ ДАНИХ

1.1. Числова та нечислова обробки даних

Дані – це цифрові і графічні відомості про об'єкти навколишнього світу. Під терміном обробка даних розуміється послідовність дій, які необхідно виконати для рішення певної задачі.

Числова обробка. ЕОМ – електронна обчислювальна машина, її найменування говорить про те, що вона була створена з метою звільнити людину від важких та нудних (рутинних) обчислень. Отже перша область застосування обчислювальної техніки – виконання чисельних розрахунків, які занадто довго або взагалі неможливо робити вручну. Розвиток цієї галузі сприяв інтенсифікації методів чисельного рішення складних математичних задач, появи мов програмування, орієнтованих на зручний запис чисельних алгоритмів, становленню зворотного зв'язку з розробниками нової архітектури ЕОМ. Характерною особливістю даної сфери застосування обчислювальної техніки є наявність складних алгоритмів обробки, які застосовуються до простих по структурі даних, обсяг яких порівняно невеликий. Розглянемо не складний приклад обчислення по формулі:

$$S = \sum_{i=1}^{10} \frac{1}{i^2} = 1 + \frac{1}{4} + \frac{1}{9} + \dots + \frac{1}{100}.$$

Обчислювальний процес в даному випадку є циклічним, тобто містить багаторазово повторюваний оператор обчислення суми, а саме 10 повторень. Алгоритм процесу представлений на рис. 1.2. При його програмуванні буде використано всього дві прості змінні: i та S . Перша змінна визначає кількість повторень i тому буде описана типом цілий, друга змінна S буде описана типом дійсний. Таким чином, прості змінні в системах програмування характеризуються певним типом їх значень, які задаються програмістом в залежності від завдання. Зміна характеру завдань обумовлює необхідність використання інших видів даних.

Розглянемо другий приклад задачі обчислювального характеру: необхідно вичислити суму виду

$$\text{SUM} = \sum_{i=1}^4 \sum_{j=1}^4 a_{ij}, \text{ де } a_{ij} - \text{дійсні числа.}$$

Із постановки задачі зрозуміло, що для її рішення необхідно використовувати такий вид даних як масив – фіксована кількість упорядкованих однотипних компонентів (елементів), кожному з яких відповідає свій індекс(и). Тип компонентів буде визначати тип масиву, кількість індексів у компонента – роз-

Навчальне видання

**Микола Павлович МІКУЛА,
Юрій Анатолійович КОЦЮК,
Олеся Миколаївна МІКУЛА**

ОРГАНІЗАЦІЯ БАЗ ДАНИХ ТА ЗНАНЬ

*Навчальний посібник для студентів
спеціальності «Комп'ютерні науки»*

Комп'ютерна верстка Наталії Крушинської

Формат 100x70/16. Ум. друк. арк. 15,76. Обл.-вид. арк. 15,06. Наклад 100 пр. Зам. № 68–21.
Папір офсетний. Друк цифровий. Гарнітура «Cambria».

Оригінал-макет виготовлено у видавництві
Національного університету «Острозька академія»,
Україна, 35800, Рівненська обл., м. Острог, вул. Семінарська, 2.
Свідоцтво суб'єкта видавничої справи РВ № 1 від 8 серпня 2000 року.

Виготовлено ФОП Свиначук М. В.
Тел. (+38068) 68 35 800, e-mail: 35800@ukr.net.