МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

А. Ю. Берко, О. М. Верес, В. В. Пасічник

МОДЕЛІ БАЗ ДАНИХ ТА ЗНАНЬ

ПІДРУЧНИК

СЕРІЯ «КОМП'ЮТИНҐ»

За загальною редакцією д.т.н., професора, Лауреата державної премії України у галузі науки та техніки В. В. Пасічника

Видавництво "Магнолія-2006"
Львів

УДК 004.65

Відтворення цієї книги або будь-якої її частини заборонено без письмової згоди видавництва. Будь-які спроби порушення авторських прав будуть переслідуватися у судовому порядку.

Рецензенти:

Литвин В.В. – д-р техн. наук, професор, завідувач кафедри інформаційних систем та мереж Національного університету «Львівська політехніка»;
Турбал Ю.В. – д-р техн. наук, професор, професор кафедри «Комп'ютерних наук та прикладної математики» Національного університету водного господарства та природокористування;
Лупенко С.А. – д-р техн. наук, професор, академік Академії наук вищої школи України, Profesor Politechniki Opolskiej (Польща).

Затверджено Науково-методичною радою Національного університету "Львівська політехніка" як підручник для студентів спеціальності 122 «Комп'ютерні науки», 124 «Системний аналіз» та 126 «Інформаційні системи і технології»

Берко А.Ю., Верес О.М., Пасічник В.В.
Моделі баз даних та знань: підручник. — Львів : «Магнолія-2006», – 466 с. — (Серія «Комп'ютинґ»).

ISBN 978-617-574-295-2

Підручник містить матеріал, необхідний для вивчення основних теоретичних засад, функціональних можливостей та практичного застосування сучасних систем і технологій баз даних, розроблення прикладних засобів та комп'ютерних інформаційних систем різного спрямування на основі реляційних баз даних.

Призначається для студентів, що навчаються за освітньо-професійними програмами першого (бакалаврського) рівня вищої освіти спеціальностей 122 «Комп'ютерні науки», 124 «Системний аналіз» та 126 «Інформаційні системи і технології» галузі знань 12–Інформаційні технології та споріднених галузей знань, пов'язаних з вивченням комп'ютерних інформаційних технологій. Може бути використаний викладачами в якості дидактичного матеріалу, а також для самостійного вивчення і підвищення кваліфікації.

УДК 004.65

© А.Ю.Берко, О.М.Верес, В.В.Пасічник, 2024

ISBN 978-617-574-295-2 © Видавництво "Магнолія-2006"

ЗМІСТ

ПЕРЕДМОВА НАУКОВОГО РЕДАКТОРА СЕРІЇ ПІДРУЧНИКІВ ТА НАВЧАЛЬНИХ ПОСІБНИКІВ «КОМП'ЮТИНҐ» 10

ВСТУП 16

Розділ 1. Основи комп'ютерного опрацювання даних 20

1.1. Інформаційні системи та інформаційні технології 20
1.1.1. Передумови виникнення напряму інформаційних систем та технологій. 21
1.1.2. Основні поняття та визначення 22
1.1.3. Етапи еволюції інформаційних систем 25
Запитання для повторення та контролю знань 35
Завдання для самостійного розв'язування 35
1.2. Інформація і дані 38
1.2.1. Поняття та розуміння інформації 38
1.2.2. Поняття та властивості даних 39
Запитання для повторення та контролю знань 42
Завдання для самостійного розв'язування 42
Висновки до розділу 1 44

Розділ 2. Моделі баз даних 45

2.1. Архітектура баз даних 45
2.1.1. Моделі даних 46
2.1.2. Архітектура баз даних CODASYL 47
2.1.3. Архітектура баз даних ANSI 48
Запитання для повторення та контролю знань 50
Завдання для самостійного розв'язування 50
2.2. Фізичні моделі даних 52
2.2.1. Одиниці фізичної структури даних 53
2.2.2. Способи реалізації фізичної структури даних 54
Запитання для повторення та контролю знань 56
Завдання для самостійного розв'язування 56
2.3. Концептуальна модель бази даних. Метод "сутність – зв'язок" 58
2.3.1. Сутності та атрибути 59
2.3.2. Види та властивості зв'язків 60

2.3.3. Порядок побудови моделі “сутність-зв'язок” ……..…….....….. 61
Запитання для повторення та контролю знань .....……..…….. 65
Завдання для самостійного розв’язування .....……..... 65
2.4. Даталогічна концептуальна модель бази даних. Логічні одиниці даних .. 68
2.4.1. Даталогічний етап побудови баз даних ...……..……...…….. 68
2.4.2. Даталогічні одиниці даних .....…..……..…….....……. 69
Запитання для повторення та контролю знань .....……..…….. 72
Завдання для самостійного розв’язування .....……..... 73
2.5. Даталогічна концептуальна модель. Види моделей баз даних 74
2.5.1. Класичні моделі баз даних ...…….....……...……...…….. 75
2.5.2. Моделі баз даних побудовані на основі класичних ...……..…….. 76
2.5.3. Моделі баз даних NoSQL ……. 78
Запитання для повторення та контролю знань .....……..…….. 98
Завдання для самостійного розв’язування .....……..... 98
Висновки до розділу 2 .....……..... 102

Розділ 3. Реляційна модель даних ….....……..... 103

3.1. Концепція реляційної моделі бази даних ….....……..... 104
3.1.1. Математичні основи реляційної моделі баз даних 104
3.1.2. Інформаційне відношення .....…….....……...…….. 105
3.1.3. Властивості відношень реляційної бази даних ……..……..….. 108
Запитання для повторення та контролю знань .....……..……..….. 111
Завдання для самостійного розв’язування .....……..……..... 111
3.2. Побудова схем реляційних баз даних .....……..……..... 114
3.2.1. Аномалії реляційної бази даних .....…….....……...……..... 115
3.2.2. Обмеження в реляційних базах даних ……..……..…….. 117
Запитання для повторення та контролю знань .....……..…….. 120
Завдання для самостійного розв’язування .....……..……... 120
3.3. Функціональні залежності в реляційних базах даних ...…….. 122
3.3.1. Поняття та зміст функціональних залежностей у відношеннях..... 122
3.3.2. Аксіоматика та властивості функціональних залежностей .....……..….. 124
Запитання для повторення та контролю знань .....……..……. 127
Завдання для самостійного розв’язування .....……..……... 128
3.4. Ключі у відношеннях реляційних баз даних ...……...…..... 130
3.4.1. Поняття та зміст ключів у відношеннях 130
3.4.2. Порядок визначення ключів у відношеннях .....……...……. 132
3.4.3. Функції та призначення ключів ……..…….....…….. 133
Запитання для повторення та контролю знань .....……..……. 135
Завдання для самостійного розв’язування .....……..……... 136
3.5. Нормалізація відношень ……..... 137
3.5.1. Поняття та призначення процесів нормалізації 138
3.5.2. Метод нормалізації відношень за Коддом .....……...……. 139
3.5.3. Порядок проектування відношень за методом нормалізації .....….. 141

Запитання для повторення та контролю знань 145
Завдання для самостійного розв'язування 145
3.6. Подальша нормалізація відношень. Нормальні форми вищих порядків . 147
3.6.1. Багатозначні залежності у відношеннях 147
3.6.2. Аксіоматика та властивості багатозначних залежностей 148
3.6.3. Четверта та п'ята нормальні форма відношень 150
3.6.4. Нормальні форми вищих порядків 151
Запитання для повторення та контролю знань 153
Завдання для самостійного розв'язування 154
Висновки до розділу 3 156

Розділ 4. Засоби маніпулювання даними в реляційній моделі 157

4.1. Поняття реляційної алгебри. Теоретико – множинні операції 159
4.1.1. Операції над відношеннями 159
4.1.2. Теоретико-множинні операції 160
Запитання для повторення та контролю знань 164
Завдання для самостійного розв'язування 164
4.2. Спеціальні реляційні операції 166
4.2.1. Операція проекції відношень 167
4.2.2. Операція селекції відношень 168
4.2.3. Операція натурального з'єднання відношень 169
4.2.4. Операція умовного з'єднання відношень 170
4.2.5. Операція ділення відношень 171
Запитання для повторення та контролю знань 174
Завдання для самостійного розв'язування 174
4.3. Операції над інформаційним наповненням відношень 176
4.3.1. Призначення та міст операцій зміни інформаційного наповнення відношень 176
4.3.2. Операція додавання кортежу до відношення 177
4.3.3. Операція вилучення кортежів з відношення 178
4.3.4. Операція зміни значень атрибутів у кортежах 179
Запитання для повторення та контролю знань 181
Завдання для самостійного розв'язування 182
4.4. Операції над схемами відношень 183
4.4.1. Операція визначення нового атрибута у відношенні 184
4.4.2. Операція вилучення атрибута з відношення 185
4.4.3. Операція зміни опису атрибута відношення 186
Запитання для повторення та контролю знань 189
Завдання для самостійного розв'язування 190
4.5. Реляційне числення зі змінними-кортежами 191
4.5.1. Основні елементи та терміни реляційного числення зі змінними-кортежами 192
4.5.2. Вирази та формули реляційного числення зі змінними-кортежами 193
4.5.3. Застосування реляційного числення зі змінними-кортежами 194

Запитання для повторення та контролю знань 196
Завдання для самостійного розв'язування 197
4.6. Відповідність формул реляційного числення зі змінними-кортежами та операцій реляційної алгебри 199
4.6.1. Функціональність формул реляційного числення зі змінними-кортежами 200
4.6.2. Теоретико-множинні операції 200
4.6.3. Спеціальні реляційні операції 201
4.6.4. Операції над інформаційним наповненням відношень 202
4.6.5. Операції над схемами відношень 203
Запитання для повторення та контролю знань 204
Завдання для самостійного розв'язування 205
4.7. Реляційне числення зі змінними на доменах 207
4.7.1. Основні елементи та терміни реляційного числення зі змінними на доменах 208
4.7.2. Вирази та формули реляційного числення зі змінними на доменах 209
4.7.3. Застосування реляційного числення зі змінними на доменах 210
4.7.4. Еквівалентність формул реляційного числення зі змінними на доменах та інших засобів маніпулювання даними 210
4.7.5. Вирази еквівалентні до теоретико-множинних операцій реляційної алгебри 211
4.7.6. Вирази еквівалентні до спеціальних реляційних операцій реляційної алгебри 212
4.7.7. Вирази еквівалентні до операцій зміни стану відношень 213
4.7.8. Вирази еквівалентні до спеціальних реляційних операцій над схемами відношень 213
Запитання для повторення та контролю знань 215
Завдання для самостійного розв'язування 215
Висновки до розділу 4 217

Розділ 5. Мови запитів 219

5.1. Мова SQL 220
5.1.1. Коротка історія мови SQL 220
5.1.2. Структура мови SQL 222
5.1.3. Базові поняття мови SQL. Об'єкти бази даних SQL 223
5.1.4. Типи даних SQL 226
5.1.5. Оператори 229
5.1.6. Засоби створення об'єктів бази даних – SQL-DDL 231
5.1.7. Засоби опрацювання даних – SQL-DML 235
5.1.8. Засоби управління даними – SQL-DCL 236
Запитання для повторення та контролю знань 239
Завдання для самостійного розв'язування 240
5.2. Мова Query-By-Example 242
5.2.1. Коротка історія мови Query-By-Example 242

5.2.2 Структура запитів Query-By-Example 243
5.2.3. Пошук та вибір даних засобами Query-By-Example 249
5.2.4. Оновлення даних засобами Query-By-Example 254
Запитання для повторення та контролю знань 257
Завдання для самостійного розв'язування 257
Висновки до розділу 5 259

Розділ 6. Створення баз даних 261

6.1. Проектування баз даних 262
6.1.1. Послідовність етапів проектування баз даних 262
6.1.2. Процедури проектування баз даних реляційного типу 264
6.1.3. Метод декомпозиції проектування баз даних реляційного типу 265
6.1.4. Синтетичний метод проектування реляційних баз даних 266
Запитання для повторення та контролю знань 269
Завдання для самостійного розв'язування 269
6.2. Проектування реляційних баз даних з використанням семантичних моделей: ER-діаграми 271
6.2.1. Обмеженість реляційної моделі при проектуванні баз даних 272
6.2.2. Семантичні моделі даних 273
6.2.3. Основні типи елементів моделі "сутність-зв'язок" 274
6.2.4. Діаграма сутностей і зв'язків 276
6.2.5. Типи обмежень та їхнє подання в моделі "сутність-зв'язок" 279
6.2.6. Проблеми ER-моделювання 281
6.2.7. EER-моделювання 281
6.2.8. Нормальні форми ER-діаграм 285
6.2.9. Отримання реляційної схеми з ER-діаграми 286
Запитання для повторення та контролю знань 290
Завдання для самостійного розв'язування 290
6.3. Проектування баз даних засобами UML 295
6.3.1. Основні поняття діаграм класів UML 295
6.3.2. Категорії зв'язків 297
6.3.3. Обмеження цілісності і мова OCL 303
6.3.4. Отримання схеми реляційної бази даних з діаграми класів UML 306
Запитання для повторення та контролю знань 309
Завдання для самостійного розв'язування 310
6.4. Цілісність даних 312
6.4.1. Поняття цілісності даних 312
6.4.2. Цілісність сутностей 313
6.4.3. Цілісність посилань 314
6.4.4. Цілісність доменів 315
6.4.5. Цілісність переходів 316
6.4.6. Цілісністьбази даних 316
6.4.7. Цілісністьтранзакцій 317
6.4.8. Обмеження цілісності 318

6.4.9. Засоби підтримання цілісності даних 322
Запитання для повторення та контролю знань 324
Завдання для самостійного розв'язування 324
6.5. Захист баз даних 327
6.5.1. Поняття безпеки даних 328
6.5.2. Види захисту даних 329
6.5.3. Управління доступом до баз даних 330
6.5.4. Захист від втрат та руйнувань 333
6.5.6. Журналізація даних 335
6.5.6. Копіювання/відновлення баз даних 335
Запитання для повторення та контролю знань 338
Завдання для самостійного розв'язування 338
Висновки до розділу 6 342

Розділ 7. Класифікація баз даних 345

7.1. Розподілені бази даних 346
7.1.1. Основні поняття 347
7.1.2. Різновиди розподілених систем 347
7.1.3. Переваги розподілених баз даних 347
7.1.4. Проблеми розподілених баз даних 352
7.1.5. Розподілена система управління базами даних System R* 354
7.1.6. Інтегровані або федеральні системи і мультибази даних 362
Запитання для повторення та контролю знань 362
Завдання для самостійного розв'язування 363
7.2. Паралельні бази даних 365
7.2.1. Основні поняття паралельної обробки даних 366
7.2.2. Паралельні архітектури 367
7.2.3. Розподіл даних 370
7.2.4. Паралельна обробка запитів 371
7.2.5. Розпаралелювання операцій реляційної алгебри 372
7.2.6. Паралелізм між операціями реляційної алгебри 376
Запитання для повторення та контролю знань 376
Завдання для самостійного розв'язування 377
7.3. Системи баз даних, що ґрунтуються на правилах 379
7.3.1. Екстенсіональна та інтенсіональна частини бази даних 379
7.3.2. Активні бази даних 380
7.3.3. Дедуктивні бази даних 380
Запитання для повторення та контролю знань 382
Завдання для самостійного розв'язування 383
7.4. Об'єктно-орієнтовані бази даних 384
7.4.1. Зв'язок об'єктно-орієнтованих СУБД зі загальними поняттями об'єктно-орієнтованого підходу 385
7.4.2. Об'єктно-орієнтовані моделі даних 387
7.4.3. Мови програмування об'єктно-орієнтованих баз даних 389

7.4.4. Мови запитів об'єктно-орієнтованих баз даних 391
7.4.5. Приклади об'єктно-орієнтованих СУБД 394
Запитання для повторення та контролю знань 396
Завдання для самостійного розв'язування 396
7.5. Бази даних в Інтернеті 398
7.5.1. Вступ до Internet та Web 399
7.5.2. Середовище World Wide Web 401
7.5.3. Використання середовища Web як платформи додатків баз даних ... 405
7.5.4. Платформа Microsoft Web Solution Platform 411
Запитання для повторення та контролю знань 418
Завдання для самостійного розв'язування 419
7.6. Бази знань 422
7.6.1. Поняття комп'ютерних знань 422
7.6.2. Моделі подання знань 424
7.6.3. Отримання знань 429
7.6.4. Інтелектуальний аналіз даних 433
7.6.5. Склад і структура баз знань 436
7.6.6. Засоби побудови та опрацювання баз знань 439
Запитання для повторення та контролю знань 443
Завдання для самостійного розв'язування 444
Висновки до розділу 7 447
ПРЕДМЕТНИЙ ПОКАЖЧИК 451
СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ 459

Передмова наукового редактора серії підручників «КОМП'ЮТИНҐ»

Шановний читачу!

Розпочинаючи роботу над вступною статтею до другої черги крупного ініціативного освітньо-наукового проекту щодо видання серії україномовних посібників, підручників та монографій, які виходять у світ під узагальненим гаслом «COMPUTING», число книг якої вже перескочило чотири десятки найменувань, хочу щиро подякувати Тобі Вельмишановний Читачу, авторам, видавцю, фінансовим меценатам та спонсорам, котрі в результаті творчої, синергетичної взаємодії забезпечили успішність цього унікального вітчизняного інноваційного проекту.

Пройшло майже десятиліття як у 2008-му році до друку було подано рукопис видання першого підручника серії «КОМП'ЮТИНГ» - «Сховища даних» за авторством професорів Н.Б. Шаховської та В.В. Пасічника. Наприкінці 2017 року завершився перший етап проекту, світ побачили понад чотири десятки найменувань навчальних посібників, підручників та наукових монографій, що загалом стало не пересічною подією в українському освітньому, науковому та професійно-бізнесовому ІТ середовищі.

Нині можна впевнено стверджувати, що започаткований не чисельною групою ініціативних високо фахових викладачів та науковців кафедри інформаційних систем та мереж Львівської політехніки проект серії україномовних посібників, підручників та наукових монографій, об'єднаних проблематикою комп'ютингу – успішно реалізований.

Автори та видавці «видали на гора» перших декілька десятків оригінальних україномовних видань, а це Вельмишановний Читачу в наш час, зовсім не мало.

Започаткування наступного (другого) етапу реалізації видавничого освітньо-наукового проекту співпадає з імплементацією сутнісних змін в системі середньої та вищої освіти України загалом і її інформаційно-технологічного сектору зокрема.

В українській середній школі реалізовується амбітний проект «Нова українська школа». Ми стаємо учасниками процесів запровадження викладання у школі предмету «Інформатика», починаючи з другого класу і вихід на профільну інформатично-технологічну освіту в старшій школі (10-12 –ий класи).

При цьому ми уважно відслідковуємо базові тенденції розвитку шкільної інфоматичної освіти у світовому вимірі. У цьому контексті не можемо обійти увагою роботу групи освітян в рамках зазначеного проекту з розроблення системи сучасних вітчизняних шкільних освітніх стандартів з інформатики. Проект реалізовувався у 2016- 2017 роках на замовлення МОН України. Не буде зайвим навести ряд

відомостей щодо зарубіжного досвіду запровадження шкільних освітніх стандартів з інформатики, які подані у підготовленому у вересні 2017 р. Оксаною Володимирівною Пасічник у підсумковому аналітичному матеріалі.

«Шкільний предмет інформатика у різних країнах має як різне змістове наповнення, так і різне місце у навчальних планах та стандартах. Це пов'язано із динамічною історією викладання інформатики, а також зі змінами, які спостерігаються останніми роками. На початку впровадження цього навчального предмету в освітніх системах різних країн наприкінці 1980-х років, основна увага приділялась алгоритмічному мисленню та розвитку вмінь складати програми. З середини 1990-х до початку 2000-х років відбувалось зміщення у напрямку вивчення прикладних програм та опанування курсу інформаційних та комунікаційних технологій. У зв'язку з тим, що зазначені технології набували все більшого поширення та ставали інтуїтивно зрозумілішими, вивчення окремого предмету в школі втрачало актуальність, тож на початок 2010-х, в освітніх системах багатьох країн цей предмет був відсутнім, інтегрованим, або викладався факультативно.

Водночас, у переліку ключових компетентностей переважної більшості національних стандартів шкільної освіти присутня цифрова грамотність чи ІКТ-компетентність, як наскрізне вміння користуватися цифровими технологіями. Все очевиднішою стає необхідність вивчення не стільки конкретних програмних засобів чи технологій, як принципів їх функціонування та вміння створювати власне цифрове середовище, в тому числі програмуючи його під свої потреби.

За результатами дослідження, проведеного Вугт та Брумельгуса, інтеграція інформаційної грамотності в інші предмети в Нідерландах призвела до зникнення предмету на початку 2000-х років, а у 2013 році знову почало розглядатись питання щодо повернення предмету в шкільні програми, надаючи йому оновленого змісту та ролі.

Слід зауважити, що у контексті шкільного предмету, інформатика у різних системах освіти також зустрічається під назвами ІКТ (інформаційні та комунікаційні технології у Чехії, Таїланді, Ісландії), комп'ютерні науки (США, Канада, програма міжнародного бакалаврату International Baccalaureate) та комп'ютинг (Великобританія, Сингапур). У деяких країнах виділяється окремий предмет під назвою програмування або кодування (Бельгія, Данія, Естонія, Іспанія, Фінляндія, Угорщина, Нідерланди, Норвегія, Польща, Португалія, Словаччина). У ряді країн також використовуються такі терміни, як алгоритмічні застосунки (Греція), алгоритмічне розв'язування задач (Словаччина), розроблення алгоритмів та моделей даних (Угорщина), алгоритми та робототехніка (Tecnología, Programación y Robótica, Іспанія), інформатика та цифрові науки (Informatique et sciences du numérique, Франція). В Австралії та Новій Зеландії предмет запроваджується із назвою цифрові технології.

До прикладу, в Хорватії предметна галузь об'єднана із технологіями під загальною назвою технології та інформатика, в межах якої виділяється 4 складові: технічна творчість, світ техніки, інформаційні та комунікаційні технології та

комп’ютерне розв’язування задач, а у Словаччині та Болгарії цей предмет об’єднано в одну групу із математикою. У Швейцарії інтегрований модуль має назву медіа та інформатика.

У національному звіті Королівської академії інженерії “Шлях інформатики у школах Великобританії: вимкнути чи перезавантажити?” окремою рекомендацією є уникнення використання терміну ІКТ (ICT), і розділення предметної області на такі аспекти, як цифрова грамотність (Digital Literacy), інформаційні технології (Information Technology) та інформатика (Computer Science). Наводиться цікава аналогія вивчення інформатики з мовною освітою, у якій можна виділити читання та письмо (базова грамотність), вивчення мови (принципи функціонування) та літератури (приклади застосування).

Також хотілося би згадати термін, для якого, на жаль, наразі не вдається знайти україномовний відповідник – це computational thinking: розв'язання задач, проектування систем та розуміння людської поведінки, виходячи з базових понять інформатики. Вперше цей термін використав Сеймур Пеперт (математик, інформатик, педагог) ще у 1996 році, але актуалізація його дослідження та використання відбулася після виходу у світ статті Джанет Вінг у 2006 році.

Описове визначення та пояснення цього терміну виглядає так: процеси мислення, пов'язані з формулюванням проблеми та представленням її вирішення у такий спосіб, щоб його можна було реалізувати за допомогою комп'ютера.

Це поняття інкапсулює алгоритмічне мислення, декомпозицію, виявлення шаблонів та абстракцію. Декомпозицією є вміння розкладати задачу на складові елементи, оперуючи якими можна чітко пояснити процес виконавцеві, або записати його для власного подальшого використання. Виявлення шаблонів передбачає вміння знаходити тенденції, схожі елементи або відмінності, котрі дозволять робити прогнози. Узагальнення шаблонів та абстрагування означає вміння відкидати відомості, які не є значущими для розв’язання певної задачі, а також підсумовувати необхідну інформацію. Розроблення алгоритму є вмінням створювати покрокову стратегію для розв’язання задачі.

Часто до цих основних елементів додають такі характеристики, як творчість, експериментування, відлагодження, наполегливість та співпраця. Таке складене синтетичне означення обчислювального мислення виходить на ключову позицію при визначенні цілей вивчення предмету інформатики в освітніх системах різних країн.

У сучасному інформаційному суспільстві цифрова компетентність не обмежується базовими навичками роботи з типовими програмами, а проникнення інформаційних, комунікаційних та цифрових технологій у всі сфери діяльності людини означає, що розуміння принципів функціонування комп’ютерних систем та програм є основоположною складовою грамотності у 21 столітті. Це, а також прогнозований дефіцит ІТ-фахівців, який оцінюється у 756 тисяч осіб до 2020 року, спричинили до того, що зміст, роль та місце предмету інформатика, нині активно трансформується, і це відображається у відповідних стандартах освіти.

Зокрема, у 2010 році було розпочато європейську кампанію з реалізації маніфесту е-навичок, одним із кроків якої є курикулярна реформа, що полягає в інтегруванні програмування у шкільні програми в 15 країнах, оновленні програм з інформатики, розробці стратегій оцифрування процесів функціонування шкіл, впровадженні у навчальні програми питань ІТ-безпеки та авторського права, тіснішій співпраці з професійною спільнотою.

Навесні 2009 року університет Альберти (Канада) додав інформатику як один з вибіркових вступних предметів, цей крок підтримали ряд університетів: Калгарі, Летбриджа, МакЮен. Це спричинило підвищення інтересу до предмету інформатики, як вибіркового на шкільному рівні.

З 2013 року відзначається європейський тиждень кодування, набула значного розмаху глобальна акція Години коду на основі ініціативи Code.Org, яка має на меті популяризацію інформатики у середніх школах.

Відповідно до даних звіту «Обчислювальне майбутнє. Комп'ютерне програмування та кодування – пріоритети, навчальні програми та Європейські ініціативи», що оновлений наприкінці 2015 року на основі опитувань 21 міністерства освіти європейських країн, у 16 країнах інформатика та кодування включені у навчальні стандарти на національному, регіональному чи локальному рівні. Спостерігається тенденція до виділення окремого предмету (якщо його не було раніше) та його вивчення на раніших освітніх етапах (в тому числі у початковій школі). При цьому 12 країн виділяють окремий предмет кодування/комп'ютингу на регіональному чи шкільному рівні, 13 країн інтегрують кодування у загальний курс інформаційних та комунікаційних технологій.

Сталий досвід присутності окремого інформативного предмету в навчальних планах з кінця 1980-х років мають такі країни, як Україна, Ізраїль, Польща, Литва, Словаччина, Кіпр. У школах багатьох країн протягом тривалого часу інформатика або не викладалась взагалі, або мала вигляд вибіркового предмету, спрямованого на оволодіння базовими інформаційними та комунікаційними технологіями. У деяких країнах паралельно існувала можливість обрати поглиблений курс програмування.

Підсумовуючи аналітичне опрацювання широкого спектру матеріалів національних стандартів шкільної освіти щодо стану та перспектив вивчення предмету інформатики можна сформулювати наступні висновки.

- Залежно від особливостей національних систем шкільної освіти, предмет може бути окремим чи інтегрованим, обов'язковим або вибірковим, імовірно з різною реалізацією цієї варіативності на різних рівнях освіти.
- У більшості країн, які оновлювали стандарти освіти останніми роками, спостерігаємо зміщення від суто прикладних вмінь використовувати програмні засоби до глибшого розуміння принципів роботи та розробки програм та систем.
- Робиться важливий наголос на відповідальне та безпечне використання інформаційних технологій, а також розуміння потенційних переваг та ризиків їх поширення та використання з різними цілями.

Випускник школи XXI століття повинен бути компетентним учасником інформаційного суспільства, усвідомлюючи його особливості, принципи функціонування, переваги та ризики швидкозмінних інформаційних технологій, а також вміти творчо розв'язувати проблемні ситуації за допомогою цифрових пристроїв та програм».

Наша увага як ініціаторів видавничого освітньо-наукового проекту до шкільної складової у системі цифрової грамотності членів сучасного інформаційного суспільства є цілком закономірною. Адже, шкільна освіта закладає фундамент як для здобуття вищої освіти, так і загалом подальшого суспільного буття.

В умовах становлення та стрімкого розвитку суспільства знань і цифрових навичок невідворотною є якісна зміна підходів до визначення змісту вищої освіти, методів і форм оцінювання якості підготовки фахівців, набуття кваліфікації, яка відповідала б вимогам стейкхолдерів. Основне протиріччя сучасної системи освіти слід трактувати як протиріччя між високими темпами приросту обсягів знань і обмеженими можливостями їх сприйняття та засвоєння особою, членом цього ж суспільства знань. Це протиріччя невідворотно примушує як теоретиків, так і педагогів-практиків відмовитися від ідеального образу усебічно розвиненої особистості і перейти до нового ідеального образу особи з максимально розвинутими здібностями, яка вдосконалюється шляхом саморегуляції та самоосвіти впродовж всього життя. Інтеграція вищої освіти України у світову освітню систему є об'єктивним процесом, який сутнісно трансформує національну систему вищої освіти, ключовими складовими якої стають освітні програми та комплекси відповідних ступеневих компетентностей. Компетентнісний підхід до побудови освітніх програм у вищій школі, що реалізується в Україні відповідно до стратегії реформування національного освітнього простору складає основу студентоцентрованого навчання, базовим і системо твірним елементом якого є студент, слухач, учень, що прагнуть набути нових знань, умінь та навичок.

Інформатична освіта університетського рівня в цьому контексті не є виключенням. Саме таку парадигму ми прагнемо реалізувати в рамках другого етапу проекту видання книг освітньо-наукової серії КОМП'ЮТИНГ.

У вищій школі України з 2016 року імплементується нова номенклатура галузей знань та спеціальностей. В контексті наших професійних інтересів мова йде про галузь знань 12-Інформаційні Технологііі з шістьма спеціальностями: 121 – програмна інженерія; 122 – комп'ютерні науки; 123 – комп'ютерна інженерія; 124 – системний аналіз; 125 – кібербезпека; 126 – інформаційні системи та технології.

Часовий відтинок, що пройшов з моменту виходу у світ першої книги освітньо-наукової серії КОМП'ЮТИНГ характеризується бурхливими процесами становлення, формування та стрімкого зростання інформаційного суспільства у всіх без винятку напрямах, профілях та аспектах. Не є виключенням в цьому потоці і освітня галузь, у якій відбувається активний пошук нових парадигм, концепцій, підходів. Якщо наприкінці 2010 року в україномовній «освітній» літературі та

«наукових» публікаціях набув популярності феномен ІКТ (англ. ICT), який не зовсім коректно був привнесений в україномовне термінологічне середовище як інформаційно-комунікаційні технології замість оригінального подання, яке слід перекладати як «інформаційні та комунікаційні технології» (див. першоджерело, в якому вперше використано концепт), то для кінця 2017 року характерною ознакою та особливістю європейського освітнього простору є нове «народження» та запровадження до освітніх стандартів різних освітніх та освітньо-кваліфікаційних рівнів галузі комп'ютингу термінологічного концепту ТІС (англ. IST), який трактується як технології інформаційного суспільства (information societies technology).

Навіть перший побіжний погляд на здавалося б незначні «косметичні» уточнення трактування понять та концептів наштовхує на думку щодо їх глибинної тектонічної природи та сутнісних методологічних змін, які за цим невідворотно послідують. Зверну увагу Шановного Читача на принципову інноваційну ознаку концепту ТІС (англ. IST), в якому інформаційні технології переводяться до розряду загальнолюдських та загально цивілізаційних. Пророче бачення видатного українського вченого В.І.Вернадського щодо високотехнологічного підґрунтя ноосферного шару сучасного світу знаходять своє методологічне, наукове та виробничо-технологічне втілення.

Основним завданням, яке ставиться перед виданнями освітньо-наукової серії КОМП'ЮТИНГ на другому (нинішньому) етапі її розгортання, є якомога повніше, змістовніше та яскравіше подання нашої професійної галузі КОМП'ЮТИНГу в трактуванні ТЕХНОЛОГІЙ ІНФОРМАЦІЙНОГО СУСПІЛЬСТВА, розуміючи, що такими технологіями у сучасному світі є саме комп'ютерні інформаційні технології. Цей процес цілком відповідає нинішньому тренду, який фіксується появою таких інформаційно-технологічних інновацій, якими є хмарні обчислення (Cloud Computing), інтернет речей (IoT), розумні міста (SC), розумні регіони (SR), системи штучного інтелекту, технології кібербезпеки та великих за обсягом даних (Big Data), а також нового освітньо-наукового напрямку «Науки про дані» (Data Science), який покликаний опікатися дослідженням процесів аналізу даних та знань.

Нове «дихання», яке отримує ініційований групою небайдужих ІТ фахівців, освітян, бізнесменів, науковців видавничий проект серії КОМП'ЮТИНГ має надійне творче підґрунтя та добру верифіковану часом традицію. Бажаю Тобі Шановний Читачу ще тривалий час мати потребу та можливість успішно розвиватись в галузі КОМП'ЮТИНГу, активно опрацьовуючи все нові і нові видання цієї без перебільшення унікальної освітньо-наукової серії.

З глибокою повагою,
д.т.н., професор Володимир ПАСІЧНИК

Вступ

Сучасні комп'ютерні інформаційні технології дають можливість вирішувати проблеми, практично, в усіх галузях та напрямах людської діяльності. Будь-які задачі – від ведення домашнього господарства до керування польотами космічних кораблів чи технологічними процесів, можуть вирішуватись засобами комп'ютерних систем. Спільною рисою в усіх випадках є потреба в накопиченні, зберіганні, підтримці і опрацюванні значних обсягів інформації. Проблема організації інформаційних ресурсів в комп'ютерних системах видається простою лише на перший погляд. При незначній кількості значень, якими треба оперувати, дійсно, можна застосувати певні тривіальні вирішення, такі як простий перебір, перегляд, візуальна оцінка, тощо. Наприклад, особистий список номерів телефонів, що складає декілька десятків значень, можна просто вивчити напам'ять, зберегти в записнику, чи пам'яті мобільного телефону. Дещо складнішим виглядає завдання віднайти потрібний номер у списку всіх абонентів телефонної мережі міста. Якщо він записаний у довільному порядку, то затрати часу і ресурсів на знаходження потрібного номеру можуть бути дуже значними. Тому при роботі з великими обсягами застосовуються певні прийоми і методи, що дозволяють полегшити їх опрацювання, зокрема впорядкування списку абонентів за певною ознакою (алфавітом чи категорією) значно спрощує пошук. Бази даних є одним зі способів вирішення проблем використання великих та дуже великих обсягів інформації в комп'ютерних системах.

За період від початку 60-х років XX-го сторіччя в галузі комп'ютерних наук, що вивчає бази даних було зроблено багато теоретичних досліджень, розроблено велику кількість технологій та засобів опрацювання даних. Завдяки цьому сьогодні бази даних є невід'ємною частиною комп'ютерних інформаційних систем. Сучасні бази даних – це складні системи, які забезпечують не лише накопичення та зберігання певної множини значень, а надають користувачеві широкі можливості з їх ефективного застосування і володіють значним набором функцій попереднього опрацювання, перетворення та аналізу даних.

Автори книги професори Володимир Пасічник, Андрій Берко та доцент Олег Верес є представниками київсько-львівської наукової школи з реляційних баз даних та знань започаткованої у 1980-х роках академіком Анатолієм Олександровичем Стогнієм. Підручник створено на основі результатів тривалих наукових досліджень та багаторічного досвіду авторів у викладанні предметів пов'язаних з теорією та практикою баз даних і знань, сховищ даних, систем та технологій управління базами даних для студентів, що навчаються за низкою напрямів у Національному університеті "Львівська політехніка" та інших вищих навчальних закладах Львова. Чинниками, які спонукали авторів до написання підручника стали, зокрема, незначна кількість україномовних підручників в галузі інформаційних систем і технологій, необхідність формування та запровадження української термінології з баз даних, а також прагнення

систематизувати і поширити власні ідеї та погляди щодо методики викладання і змісту навчального курсу.

У підручнику **“Моделі баз даних та знань”** подано основні фундаментальні поняття та принципи, необхідні для розуміння суті та предмету баз даних та знань, вивчення основних закономірностей, понять, принципів, методів і засобів створення, опрацювання та застосування баз даних. Зміст підручника відповідає вимогам стандартів вищої освіти України першого (бакалаврського) рівня вищої освіти за галуззю знань 12 “Інформаційні технології“ спеціальності 122 “Комп’ютерні науки“ (затвердженого та введено в дію наказом МОН України від 10.07.2019 р. № 962,) спеціальності 124 “Системний аналіз” (затвердженого та введено в дію Наказом МОН України від 13.11.2018 № 1245) та спеціальності 126 “Інформаційні системи і технології” (затвердженого та введено в дію Наказом МОН України від 12.12.2018 № 1380) щодо забезпечення нормативного змісту підготовки здобувачів вищої освіти, сформульований у термінах результатів навчання.

Ефективне фахове застосування будь яких практичних знань, прикладних засобів і технологій, у тому числі і в галузі формування, опрацювання та адміністрування ресурсів даних в інформаційних системах та мережах, неможливе без розуміння фундаментальних принципів, понять і закономірностей які сформульовано у вигляді відповідних теоретичних результатів. Саме тому вивчення предмету організації баз даних та знань пропонується починати з опанування теоретичних основ баз даних. У даній частині підручника у стислій, компактній формі викладено основні теоретичні принципи, положення та закономірності, на яких ґрунтується більшість сучасних технологій та систем опрацювання баз даних.

Основну увагу підручника спрямовано на вивчення теоретичних понять з галузі реляційних баз даних, оскільки цей підхід до організації баз даних сьогодні є достатньо ґрунтовно дослідженим і широко вживаним в різноманітних інформаційних системах та технологіях. Переважна більшість сучасних систем та засобів управління базами даних є орієнтованою саме на принципи реляційної моделі баз даних закладені у відповідних теоретичних дослідженнях. Створена на основі цієї моделі мова запитів SQL на сьогодні де-юре та де-факто є стандартом побудови засобів створення та застосування баз даних інформаційних систем. Теоретичні принципи реляційної моделі реалізують у своїх продуктах такі провідні виробники систем управління базами даних як Oracle, IBM, Microsoft, SyBase, Postgress Community та інші. Теорія реляційних баз даних, в свою чергу, стала підґрунтям для розвитку таких напрямів у технологіях баз даних як сховища даних, багатовимірні структури даних, об’єктно-реляційні бази даних, системи і технології оперативного аналізу, та багатьох інших сучасних та перспективних методів і засобів опрацювання даних. Саме тому, на думку авторів, ґрунтовне вивчення теоретичних засад реляційного підходу до баз даних є необхідною складовою підготовки майбутнього фахівця в галузі комп’ютерних наук та інформаційних технологій.

У книзі розглянуто загальні принципи комп’ютерного опрацювання даних, фундаментальні засади інформаційних систем та технологій, моделі баз даних, проектування та нормалізація баз даних, реляційна алгебра та реляційні числення. Основні підходи до організації реляційних баз даних, принципи, визначення, терміни та мето-

дологію їхньої побудови і опрацювання викладено відповідно до концепцій сформульованих у фундаментальних працях Е.Ф. Кодда, К. Дейта, Дж. Ульмана, Т. Коннолі та інших авторитетних науковців визнаних у світі. Зміст матеріалу підручника відповідає міжнародним рекомендаціям Computing Curricula 2020 (CC2020) та Computer Science Curricula 2023 щодо організації вивчення курсу "Інформаційний менеджмент" в частині організації баз даних. Опанування матеріалу підручника забезпечить правильне розуміння процесів створення і застосування баз даних, а також дасть необхідні знання для творчої роботи зі створення нових систем і технологій на основі баз даних.

Книгу поділено на сім розділів, кожен з яких містить змістовно завершений фрагмент навчального матеріалу. Послідовність розділів, забезпечує логіку вивчення теоретичного курсу за принципами “від простого до складного”, та “від загального до детального”.

Розділ 1. *Основи комп’ютерного опрацювання даних*. У розділі подано визначення та зміст фундаментальних понять інформаційних систем та інформаційних технологій, еволюцію систем і технологій опрацювання даних, визначення інформації та даних, їхньої ролі і властивостей.

Розділ 2. *Моделі баз даних.* Розділ присвячено проблемам загальної архітектури баз даних, процесам моделювання баз даних та їхніх елементів, класифікації та стандартизації моделей баз даних, методикам формалізації та моделювання даних, визначенню стандартних одиниці даних, типових видів моделей баз даних, які застосовують для їх зберігання, опрацювання та застосування.

Розділ 3. *Реляційна модель даних.* У цьому розділі розглянуто основні терміни та поняття реляційної моделі, проблеми побудови реляційних баз даних, методи та засоби створення реляційних баз даних та покращення їхніх якісних показників на основі методики нормалізації.

Розділ 4. *Засоби маніпулювання даними в реляційній моделі.* В першій частині розділу вивчаються основні принципи маніпулювання реляційними базами даних на процедурному рівні зі застосуванням операцій реляційної алгебри, зокрема, з такими можливостями як теоретико-множинні операції, спеціальні реляційні операції, операції актуалізації баз даних, операції зміни структур баз даних. Друга частина розділу містить викладення теоретичних принципів непроцедурного опрацювання баз даних на основі математичної логіки. Реляційні числення покладено в основу побудови та функціонування високотехнологічних та інтелектуальних засобів роботи з базами даних.

Розділ 5. *Мови запитів.* Цей розділ містить описання мови запитів SQL і QBE. Розділ починається з короткої історії становлення мови SQL, базових понять мови SQL та системи об’єктів бази даних. У подальшому розглянуто типи даних і операторів мови SQL, а також засоби створення, опрацювання та управління даними. У другій частині розділу подано історію мови запитів QBE, порядок створення запитів QBE, характеристику видів запитів QBE, застосування критеріїв в запитах QBE, а також описання запитів на оновлення баз даних і аналітичних.

Розділ 6. *Створення баз даних.* Розділ присвячено опису послідовності етапів проектування баз даних. Коротко розглянуто деякі особливості проектування реляційних баз даних з використанням семантичних моделей – моделі “сутність-зв’язок”

(ER-модель від *Entity-Relationship*). Подано основні поняття діаграм класів мови UML та можливість застосування цієї діаграмної моделі для проектування реляційних баз даних. У розділі описано аспекти цілісності даних, сутностей, зв'язків і доменів, види обмеження цілісності і засоби підтримання цілісності даних. Розділ завершується розглядом поняття безпеки даних, видів захисту даних, управління доступом до баз даних, захисту від втрат та руйнувань, журналізації даних, копіювання чи відновлення баз даних.

Розділ 7. *Класифікація баз даних.* У розділі подано описання особливостей розподілених і паралельні бази даних, технологій розподілених і паралельних баз даних. Розглянуто паралельні архітектури, види паралельності та реалізацію паралельності. У розділі також описано бази даних, в яких правила відіграють суттєво більшу роль ніж в традиційних реляційних, їхні екстенсіональна та інтенсіональна частини, особливості активних і дедуктивних баз даних. Ускладнення інформаційного середовища стимулювало пошук інших моделей баз даних для зручнішого проектування, управління і реалізації складних типів даних. Першою такою моделлю стала описана в розділі об'єктно-орієнтована модель даних, розглянуто мови програмування і запитів об'єктно-орієнтованих баз даних, приклади об'єктно-орієнтованих СУБД. Використання Інтернету як найважливішого інструменту для бізнесу радикально змінило роль і масштаби ринку баз даних. Виникла потреба розроблення стратегії доступу до баз даних з Інтернету. На завершення розділу розглянуто поняття комп'ютерних знань, отримання, подання та опрацювання знань засобами бази знань.

Кожен розділ поділено на теми, які відповідають за обсягом і змістом матеріалу одного лекційного заняття. Тема містить коротке викладення необхідних для вивчення і розуміння предмету теоретичних відомостей, приклади, які ілюструють теоретичне викладення, контрольні запитання для перевірки та самоперевірки засвоєння вивченого матеріалу і набір тестових завдань для контрольних заходів чи іспитів. Кожне тестове завдання ґрунтується на матеріалі теми і може містити одну або декілька правильних відповідей.

Підручник призначений для використання викладачами та студентами різних спеціальностей, пов'язаних з комп'ютерними інформаційними системами та технологіями у процесі вивчення курсів "Організація баз даних та знань", "Бази та банки даних", "Бази даних" тощо. Книгу можуть застосовувати фахівці в галузі баз даних та знань як довідковий підручник у практичній роботі чи для підвищення кваліфікації.

Для сприйняття та розуміння матеріалу підручника читачеві бажано мати базові знання у таких галузях як програмування, алгоритми та структури даних, теорія множин, математична логіка, основи моделювання.

Автори висловлюють подяку своїм колегам – викладачам кафедри інформаційних систем та мереж Національного університету "Львівська політехніка" та студентам університету, які брали участь в апробації матеріалу підручника у навчальному процесі.

108. Shellman M., Afyouni H., Pratt P. J., Last M. Z. A Guide to SQL (MindTap Course List). Cengage Learning; 10th edition, 2020. 320 p.
109. Shields W. SQL QuickStart Guide: The Simplified Beginner's Guide to Managing, Analyzing, and Manipulating Data With SQL. ClydeBank Media LLC, 2019. 251 p.
110. Silberschatz A., Korth H. F., Sudarshan S. Database system concepts. McGraw-Hill, 2011. 247 p.
111. Simsion G., Witt G. Data modeling essentials. Elsevier, 2004. 560 p.
112. Singh S. K. Database Systems: Concepts, Design and Applications, 2nd Edition. Pearson India, 2011. 944 p.
113. Suciu D., Olteanu D., Ré C., Koch C. Probabilistic databases. Springer Nature, 2022. 164 p.
114. Sullivan D. NoSQL for mere mortals. Addison-Wesley Professional, 2015. 552 p.
115. Teorey T. J. Database modeling and design. Morgan Kaufmann, 1999. 366 p.
116. Teorey T. J., Lightstone S. S., Nadeau T., Jagadish H.V. Database modeling and design: logical design. Elsevier, 2011. 352 p.
117. Ullman J. D. Principles of Computer and Knowledge-base Systems. Computer Science Press, 1988. 374 p.
118. Vaswani V. MySQL database usage & administration. McGraw Hill Professional, 2009. 369 p.
119. Wang L., Ranjan R., Chen J., Benatallah B. Cloud Computing: Methodology, Systems, and Applications. CRC Press, 2017. 844 p.
120. Wang X.Y. 100 Interview Questions: SQL & Database (Advanced Topics in Database Book 1). Kindle Edition, 2023. 187 p.
121. Wang X.Y. PostgreSQL: 100 Interview Questions (Advanced Topics in Database Book 2). Kindle Edition, 2023. 242 p.
122. Wang X.Y. MySQL: Interview Questions And Answers (Advanced Topics in Database Book 3). Kindle Edition, 2023. 228 p.
123. Wang X.Y. SQL Server: Interview Questions and Answers (Advanced Topics in Database Book 4). Kindle Edition, 2023. 214 p.
124. Wang X.Y. Oracle Database: Interview Questions and Answers (Advanced Topics in Database Book 5). Kindle Edition, 2023. 221 p.
125. Wang X.Y. MongoDB: Interview Questions and Answers (Advanced Topics in Database Book 6). Kindle Edition, 2023. 236 p.
126. Warren J., Marz N. Big Data: Principles and best practices of scalable realtime data systems. Simon and Schuster, 2015. 328 p.
127. West M. Developing high quality data models. Elsevier, 2011. 408 p.
128. West R., Zacharias M., Assaf W., Aelterman S., Davidson L., D'Antoni J. SQL Server 2019 Administration Inside Out. Microsoft Press, 2020. 992 p.
129. White T. Hadoop: The Definitive Guide . O'Reilly Media, 2012. 657 p.
130. Zhao A. SQL Pocket Guide. O'Reilly Media, Inc., 2021. 356 p.

Навчальне видання

А. Ю. Берко, О. М. Верес, В. В. Пасічник

Моделі баз даних та знань

Підручник

Формат 70x100/16. Друк цифровий. Папір офсетний. Гарнітура PetersburgC.
Умовн. друк. арк. 37,78.

Видавництво «Магнолія 2006»
м. Львів-53, 79053, Україна, тел.: +38 (050) 370-19-57
e-mail: magnol06@ukr.net
https://magnolia.lviv.ua

Свідоцтво про внесення суб'єкта видавничої справи
до Державного реєстру видавців, виготівників і розповсюджувачів видавничої продукції:
серія ДК № 6784 від 30.05.2019 року,
видане Державним комітетом інформаційної політики,
телебачення та радіомовлення України.

Надруковано у друкарні видавця ФОП Марченко Т. В.