

BAZY DANYCH

1

GRZEGORZ SURÓWKA

2018

PLAN WYKŁADU

1. **WPROWADZENIE**
2. DBMS, MODELE DANYCH, MODEL RELACYJNY
3. DIAGRAMY E/R
4. E/R → MODEL RELACYJNY
5. ALGEBRA RELACJI
6. ZŁĄCZENIA
7. NORMALIZACJA BAZY DANYCH
8. TRANSAKCJE
9. OPTYMALIZACJA ZAPYTAŃ
10. WYDAJNOŚĆ
11. PODZAPYTANIA SQL
12. GRUPOWANIE
13. WIDOKI, KURSORY, WYZWALACZE, SEKWENCJE
14. SQL

ŹRÓDŁA

- J.D. Ullman, J. Widom, Podstawowy kurs systemów baz danych, wyd. III, Helion 2011
- R. Elmasri, S.B. Navathe, Fundamentals of Data Base Systems, Addison-Wesley
- SQLzoo.net
- www.graspsql.com
- SQLfiddle.com
- www.mysql.com
- Helion: "bazy", "DBMS", "SQL", "MySQL"

PLAN

- POJĘCIA PODSTAWOWE
- BAZY RELACYJNE
- ŚRODOWISKO PRACY
- PRZYKŁAD

informacja: wiedza o obiektach (przedmioty, fakty, procesy itp.) zależna od kontekstu

dane: reprezentacja informacji, odpowiednia do przetwarzania/przesyłana

przetwarzanie danych: wykonywanie operacji na danych (logiczne, arytmetyczne, sortowania, wyszukiwania, łączenia, zestawiania, itp.)

przetwarzanie danych \leftrightarrow przetwarzanie informacji

BAZY DANYCH - KONCEPCJE

baza danych

- zorganizowana kolekcja (magazyn) danych powiązanych tematycznie posiadająca określoną strukturę, która pozwala na wydajne przeszukiwanie informacji
- aplikacja komputerowa, która może szybko wyszukiwać informacje

model bazy danych

- reguły, zgodnie z którymi dane umieszcza się w strukturach
- dozwolone operacje na bazie i na danych w bazie
- zbiór zasad, którymi należy się posługiwać podczas tworzenia bazy danych

wymagania logiczne

- poprawność (zgodność z rzeczywistością i ograniczeniami)
- trwałość danych [*persistence*] (niezależność od sprzętu, systemu operacyjnego, aplikacji, nieulotność)
- spójność [*integrity*] (odporność na anomalie i błędy)
- bezpieczeństwo danych [*security*] (autoryzacja, uprawnienia)
- metadane (dane o danych, strukturach, użytkownikach)
- elastyczność modelu (duże dane, dużo danych)
- wspieranie transakcji (bezkolizyjność, spójność)
- wydajność [*performance*] (szybkość wydobywania informacji, współbieżność [*concurrency*])
- zarządzanie (systemem bazodanowym i danymi)
- interfejs programistyczny

przetwarzanie transakcyjne

- dostęp do bazy za pomocą transakcji o własnościach ACID
- synchronizacja transakcji (2PL, znaczniki czasowe, wielowersyjność danych)
- odtwarzanie spójności bazy (punkty kontrolne, wycofywanie operacji, logi)
- archiwizacja i odtwarzanie bazy

fizyczne struktury zapisu i metody dostępu

- niezależność danych od nośnika
- pliki: uporządkowane, haszowe, zgrupowane
- indeksy: drzewiaste, bitmapowe
- metody: połowienie binarne, haszowanie statyczne i dynamiczne, połączenia, sortowanie, grupowanie
- optymalizacja: składniowa, kosztowa

modele danych

- modele pojęciowe: związków-encji, UML
- modele logiczne: jednorodny, hierarchiczny, obiektowy, sieciowy, relacyjno-obiektowy, relacyjny

narzędzia programistyczne

- języki dostępu do danych w bazie (SQL)
- interfejsy (np. JDBC)
- narzędzia modelowania i projektowania

BAZA RELACYJNA

baza danych: składa się z tabel

tabela: pojemnik na dane, składa się z wierszy i kolumn, w jednej bazie może być wiele tabel, pomiędzy tabelami mogą występować powiązania logiczne (relacje)

kolumna: cecha (atrybut) obiektu

wiersz: rekord, krotka, pojedynczy obiekt
z wartościami wszystkich opisujących go cech

pole: część tabeli przechowująca jednostkowe dane

wydobywanie informacji z bazy: zwracanie krotek (rekordów), które spełniają pewien warunek (nieraz nie jest potrzebna cała krotka, ale jej wybrane pola)

ŚRODOWISKO PRACY

baza danych:

Xampp

- Apache
- **MySQL**
- FileZilla
- Tomcat
- ...



<http://localhost/phpMyAdmin>

narzędzie do łatwego zarządzania bazą danych MySQL
(napisane w php)

baza danych: biblioteka

tabela: uczniowie

imie	nazwisko	klasa
Paweł	Mazur	1b
Jan	Nowak	3c
Andrzej	Kowalski	3a
Jan	Nowak	3c

(w klasie jest dwóch uczniów o tych samych danych)

tabela: książki

tytuł	autor	rok
Lalka	Prus	2000
Potop	Sienkiewicz	2002
Pan Tadeusz	Mickiewicz	2005
Pan Tadeusz	Mickiewicz	2005

(są dwa egzemplarze tej samej książki)

problem: jak zapisać, że uczeń dostał konkretny egzemplarz?

→ książki musimy ponumerować unikatowymi numerami

→ dokładamy cechę (kolumnę) id

tabela: książki

idksiążki	tytuł	autor	rok
1	Lalka	Prus	2000
2	Potop	Sienkiewicz	2002
3	Pan Tadeusz	Mickiewicz	2005
4	Pan Tadeusz	Mickiewicz	2005

tabela: uczniowie (ta sama sytuacja)

iducznia	imie	nazwisko	klasa
1	Paweł	Mazur	1b
2	Jan	Nowak	3c
3	Andrzej	Kowalski	3a
4	Jan	Nowak	3c

klucz główny: (podstawowy, primary key)
jedno, lub więcej pól, których wartość
jednoznacznie (unikatowo) identyfikuje
każdy rekord w tabeli

system obsługi bazy danych powinien pilnować,
aby wartość id **nie została zdublowana**
(zdublowanie oznaczałoby stratę spójności danych)

→ aby powiedzieć bazie, że wartości pewnej
kolumny muszą być unikalne trzeba
oznaczyć tą kolumnę jako klucz główny

relacja: powiązanie logiczne występujące pomiędzy tabelami realizowane za pomocą klucza podstawowego i tzw. klucza obcego (klucz obcy to klucz podstawowy w innej tabeli) albo specjalnej tabeli (tabela łącząca)

typy: jeden-do-jeden, jeden-do-wielu, wiele-do-wielu

- bazę najlepiej jest zaprojektować od razu tak, jak ma wyglądać ostatecznie (czyli przed wprowadzeniem do niej pierwszego rekordu)
- dokładanie kolumny do niepustej tabeli wiązałoby się z poprawianiem każdego rekordu

tabela: wypozyczenia

idwypozyczenia	iducznia	idksiazki
1	3	1
2	3	4
3	12	9
4	5	7

- nie powinno być powtórzeń (redundancji)
tzn. informacja, która już jest w (innej) tabeli
nie powinna być kopiowana do następnej
- warto dołożyć: datę wypożyczenia, datę zwrotu

ćwiczenie praktyczne

localhost/phpMyAdmin

Databases → Create Database: "quiz", utf8_polish_ci

Create Table: "pytania", 7 kolumn

(id, pytanie, odpowiedź A, B, C, D, poprawna odpowiedź)

A_I: Auto Increment (automatyczne zwiększanie o jeden)

Collation: metoda porównywania napisów

- wprowadzić jeden rekord np. żółć, aaa, bbb, ccc, ddd, a
- usunąć tabelę: Drop
- zaimportować gotową tabelę z pliku
- zmodyfikować pola w phpMyAdmin
(generuje kwerendę i pokazuje ją)