

Wprowadzenie do baz danych

- Wprowadzenie
- Przeczytaj
- Film edukacyjny
- Sprawdź się
- Dla nauczyciela



Tworzymy społeczeństwo informacyjne. Współczesny świat nie mógłby sprawnie funkcjonować bez wydajnych systemów bazodanowych.

Początki komputerowych baz danych sięgają lat 60. XX wieku. W ramach programu Apollo wysłano wówczas ludzi na Księżyc. Użyto do tego komputerów mających moc obliczeniową współczesnych kalkulatorów. Uznaje się, że wtedy właśnie powstał pierwszy system bazodanowy, który zastąpił zbyt wolne przetwarzanie danych w plikach – stawką było przecież życie astronautów!

W sekcji „Przeczytaj” znajdziesz informacje na temat pierwszych komputerowych baz danych. W tym e-materiale omówimy również najważniejsze koncepcje i pojęcia związane z systemami bazodanowymi. Rozpoczniemy od udzielenia odpowiedzi na pytanie: czym są bazy danych?

Więcej informacji o bazach danych znajdziesz w e-materiale: [Zastosowanie baz danych](#).

Twoje cele

- Zdefiniujesz pojęcia: informacja, baza danych, system zarządzania bazą danych, model danych, język SQL.
- Sklasyfikujesz rozwijane przez lata rodzaje modeli danych.

- Opiszysz podstawowe założenia i definicje związane z najczęściej stosowanym obecnie relacyjnym modelem danych.

Przeczytaj

Programy komputerowe **przetwarzają informacje** – takie stwierdzenie nikogo nie dziwi. Jednak co ono dokładnie oznacza? Czym w ogóle jest informacja? Udzielenie odpowiedzi na te pytania nie jest łatwe, ponieważ wymaga dokonania abstrakcyjnego opisu otaczającej nas rzeczywistości.

Definicja: Informacja

Informacją nazywamy **porcję danych**, których celem jest zmniejszenie stopnia niepewności (niewiedzy) ich odbiorcy. Może nią być atrybut opisujący **stan** obiektu, jego **zachowanie** albo istniejące **powiązanie** pomiędzy fragmentami rzeczywistości. Informację należy rozpatrywać w trzech podstawowych kontekstach:

1. **syntaktycznym** (ilość informacji zawarta w analizowanych danych);
2. **semantycznym** (znaczenie informacji – jaką nowość wnosi ona do świadomości odbiorcy);
3. **pragmatycznym** (przydatność informacji, czyli wpływ porcji danych na realizowane przez odbiorcę zadanie).

Odkąd ludzie rozpoczęli tworzenie oprogramowania, stale opracowywali coraz lepsze **pojemniki na dane**, czyli tak zwane **struktury danych**. Chcemy, aby przetwarzanie informacji było możliwie szybkie oraz skuteczne i to najlepiej we wszystkich trzech wspomnianych kontekstach.

Ewolucja przetwarzania informacji

Znane ci są już niektóre pojemniki na dane: **zmienna, tablica, stos, lista, kolejka, kopiec, drzewo binarne** itd. Współczesne systemy bazodanowe to – mówiąc najprościej – bardzo zaawansowane, autonomiczne pojemniki na dane.

Bazy danych przechowują ogromne ilości informacji, udostępniają **wygodne interfejsy** do ich modyfikowania, dbają o **integralność** przechowywanego zbioru, pozwalają realizować tysiące tzw. **zapytań** (zwracając odpowiedź w ułamku sekundy), **bronią dostępu** do danych i spełniają wiele innych zadań.

Oczywiście powstanie tak rozbudowanych, zaawansowanych systemów bazodanowych wymagało wiele czasu. Nastąpiło to w drodze **powolnej ewolucji** rozwiązań, pełnej poświęceń i wspólnej pracy wielu ludzi.

Za pierwszy system bazodanowy powszechnie uznaje się oprogramowanie opracowane w ramach programu Apollo. Wysłanie człowieka na Księżyc wymagało szybkiego przetwarzania dużej ilości danych w czasie rzeczywistym. Stawką było życie astronautów.

Później okazało się, że pierwsze kroki w dziedzinie tzw. hierarchicznego modelu danych, postawione wówczas przez inżynierów IBM, były ogromnym skokiem dla informatyki.

Przegląd modeli danych

Baza danych to – najprościej mówiąc – uporządkowany zbiór informacji. Kiedy zajmujemy się tworzeniem takich wielkich zbiorów danych, to tak naprawdę **modelujemy rzeczywistość**. Tworzymy model, czyli abstrakcyjną, wirtualną reprezentację rzeczywistego świata.

Model danych to integralny zbiór zasad opisujących **stan** reprezentowanych obiektów, ich **zachowania** oraz istniejące **powiązania** logiczne pomiędzy fragmentami symulowanego świata.

Najpopularniejszym sposobem reprezentacji rzeczywistości stosowanym we współczesnych bazach danych jest tzw. **model relacyjny**, którego ideę opisał w 1985 roku **Edgar Frank Codd** w słynnej pracy pt. *Model relacyjny w zarządzaniu bazami danych*. Przedstawione przez autora tzw. **dwanaście postulatów Codda** zrewolucjonizowało sposób projektowania systemów zarządzających bazami danych.

Warto wymienić i krótko scharakteryzować wszystkie rozwijane przez lata modele danych, pamiętając jednak, że w tym cyklu e-materiałów będziemy zajmować się przede wszystkim bazami relacyjnymi.

Typologia modeli danych jest następująca:

- **Model jednorodny** – najprostszy, przewidujący użycie jednego pojemnika przechowującego wszystkie dane. Przykładem może być alfabetyczny spis pracowników firmy albo książka telefoniczna. Największe wady takiego sposobu modelowania to długi czas dostępu do informacji oraz często występująca redundancja (nadmiarowość) informacji.
- **Model hierarchiczny** – oparty na strukturze odwróconego drzewa, czyli takiego w którym korzeń znajduje się na szczycie hierarchii. W bazie istnieją węzły nadrzędne i podrzędne (czyli właśnie hierarchia). Zasada drzewa głosi, iż jeden ojciec może mieć wiele dzieci, ale każde dziecko ma tylko jednego ojca. Przykładem wykorzystania tego modelu był system IBM IMS, zapoczątkowany w roku 1966 na potrzeby programu Apollo – było to oprogramowanie przechowujące informacje o materiałach wykorzystywanych przy rozbudowie i obsłudze rakiety Saturn V.
- **Model obiektowy** – zrealizowany zgodnie z paradygmatem programowania obiektowego. Rzeczywistość reprezentujemy za pomocą obiektów tworzonych na podstawie klas. Każdy obiekt ma atrybuty opisujące jego stan. Istnieją także metody, czyli specjalne funkcje należące do klas. Metody reprezentują zachowanie obiektu oraz określają zasady dostępu do jego wartości składowych, zapewniając dodatkową

możliwość **hermetyzacji** danych w celu ochrony ich integralności. Interakcja z obiektami jest możliwa tylko przez udostępniony w klasach interfejs. Przykłady użycia tego modelu sięgają lat 70. XX wieku - np. projekt systemu bazodanowego EXODUS, zrealizowany na Uniwersytecie Wisconsin-Madison, system IRIS stworzony w firmie Hewlett-Packard, czy ORION zaprojektowany przez inżynierów w Microelectronics and Computer Technology Corporation.

- **Model sieciowy** – bazował na strukturze hierarchicznej (drzewiastej), lecz wprowadził pewne modyfikacje – używało się w nim tzw. typów rekordów i kolekcji, zaś reprezentacja tych typów oraz powiązań miała postać grafu zorientowanego, nazywanego siecią (stąd nazwa modelu). Ten model został zaadoptowany w 1969 r. przez grupę CODASYL Data Base Task Group, dzięki czemu w latach 70. XX wieku znalazł praktyczne zastosowania. Z czasem jednak model został ostatecznie zastąpiony przez architekturę relacyjną.
- **Model relacyjny** – najpopularniejszy obecnie, opracowany przez Edgara Franka Codda, sposób reprezentacji fragmentu otaczającej nas rzeczywistości. Fundamentem modelu jest matematyczne pojęcie relacji, występujące w teorii mnogości, a związane nierozdzielnie z tzw. iloczynem kartezjańskim. W relacyjnych bazach danych nie używamy klas ani obiektów, tylko encji oraz instancji. Podstawową formą przechowywania danych jest tabela (relacja) zbudowana z kolumn (atrybutów) oraz wierszy (rekordów, krotek).
- **Model relacyjno-obiektowy** – kompromis pomiędzy relacyjnymi, a obiektowymi bazami danych, łatwiejszy w implementacji od czystego modelu obiektowego. Idea modelu polega na tym, że operujemy na danych tak jak na obiektach, czyli z możliwością zaprogramowania własnych metod, używających technik obiektowych (na przykład **dziedziczenia**, **polimorfizmu**). Jednak wewnętrzny mechanizm przechowywania danych jest relacyjny. Można ująć to następująco: mechanizmy obiektowe wykorzystywane są wewnątrz relacyjnej bazy danych.

Więcej informacji na temat wymienionych modeli znajdziesz w filmie towarzyszącym tej lekcji.

Relacyjne bazy danych

Zajmijmy się pojęciami związanymi z najpopularniejszymi obecnie, relacyjnymi bazami danych. Tego typu zbiór danych uporządkowany jest w **tabelach** (relacjach), czyli w klasycznych, znanych każdemu strukturach złożonych z **wierszy** i **kolumn**. Posłużmy się przykładem. Informacje o książkach przechowywanych w bazie danych szkolnej biblioteki, zostały umieszczone w następującej tabeli:

idksiazki	tytul	autor_imie	autor_nazwisko
1	Pan Tadeusz	Adam	Mickiewicz

idksiazki	tytul	autor_imie	autor_nazwisko
2	Krzyżacy	Henryk	Sienkiewicz
3	Lalka	Bolesław	Prus

Atrybut (kolumna) to część tabeli przechowująca **dane określonego typu** (rodzaju) – na przykład liczbę, napis, czas, datę albo wartość logiczną. W pokazanej tabeli jest to każda pojedyncza kolumna (np. autor_nazwisko).

Rekord to pojedynczy wiersz w tabeli, czyli **zestaw wartości atrybutów** w niej zapisanych. W tabeli wyżej rekordem jest numer i tytuł książki wraz z informacją o jej autorze: na przykład pierwsza książka, czyli „Pan Tadeusz” autorstwa Adama Mickiewicza (wartości wszystkich czterech atrybutów).

Zestaw atrybutów opisujący w modelu dany rekord (tutaj jest nim książka) w programowaniu oraz nomenklaturze bazodanowej często określamy mianem **encji** lub ewentualnie **klasy**. Konkretnego reprezentanta (przedstawiciela) tej klasy nazywamy zaś **instancją**.

Mnogość słów opisujących zagadnienia związane z bazami danych może na początku nieco deprimować. W celu usystematyzowania użytych wcześniej pojęć, które pochodzą z różnych modeli rozwijanych na przestrzeni lat, posłużymy się następującym zestawieniem:

Współczesne bazy relacyjne	Teoria relacyjna (Edgar Frank Codd)	Teoria związków encji (Peter Chen)	Programowanie obiektowe
tabela	relacja	encja	klasa
rekord, wiersz	krotka (ang. <i>tuple</i>)	instancja	obiekt
atrybut, kolumna	atrybut	atrybut	atrybut, właściwość, cecha

System zarządzania bazą danych

Zajmowanie się bazami danych nie polega jednak wyłącznie na utworzeniu uporządkowanego zbioru informacji, który w określony sposób modeluje rzeczywistość. Wiele pracy włożono również w zaprojektowanie **narzędzi administracyjnych** służących do zarządzania zgromadzonymi danymi.

DBMS (ang. *Database Management System*) lub w wersji polskiej: **SZBD** (System Zarządzania Bazą Danych) to oprogramowanie, które odpowiada za:

- **zarządzanie użytkownikami** systemu bazodanowego, w tym obsługę wbudowanych mechanizmów autoryzacji i walidacji uprawnień,
- **ochronę integralności** (spójności) zbioru danych – zapewnienie poprawności, dokładności i aktualności zbioru informacji,
- możliwości wykonywania na zgromadzonych danych **kwerend** (zapytań w języku SQL) – tzw. standard CRUD (od ang. *Create, Read, Update, Delete*),
- zapewnienie **bezpieczeństwa i trwałości** danych, w tym mechanizmów tworzenia oraz odzyskiwania kopii zbioru informacji.

Kilka współczesnych przykładów oprogramowania DBMS: MySQL, MariaDB, FireBird, PostgreSQL, Microsoft SQL Server, SQLite, Oracle Database, Microsoft Access.

Język SQL

SQL (akronim pojęcia *Structured Query Language*) to strukturalny język zapytań umożliwiający programiście komunikację z bazą danych.

Fundamenty języka zostały opracowane przez firmę IBM. W e-materiałach wiele uwagi poświęcimy właśnie wypracowaniu sprawności w posługiwaniu się językiem SQL.

Zapytanie kierowane do bazy, nazywane też **kwerendą** (ang. *query*), to instrukcja zapisana w języku SQL, w wyniku której dokonywana jest pewna operacja na zbiorze danych. Zapytania mogą np. wyszukiwać informacje, aktualizować lub usuwać zgromadzone wartości, tworzyć tabele lub określać konfigurację struktury tabeli.

Słownik

hermetyzacja

(inaczej: enkapsulacja lub kapsułkowanie danych) mechanizm określania dostępu do atrybutów oraz metod obiektów danej klasy w celu ochrony przechowywanych wartości przed ich niepożądaną modyfikacją

dziedziczenie

w programowaniu obiektowym, mechanizm współdzielenia atrybutów oraz zachowań pomiędzy klasą bazową, a jej pochodną; dzięki dziedziczeniu niektórych rozwiązań (atrybutów i metod) napisany kod można używać w wielu aplikacjach

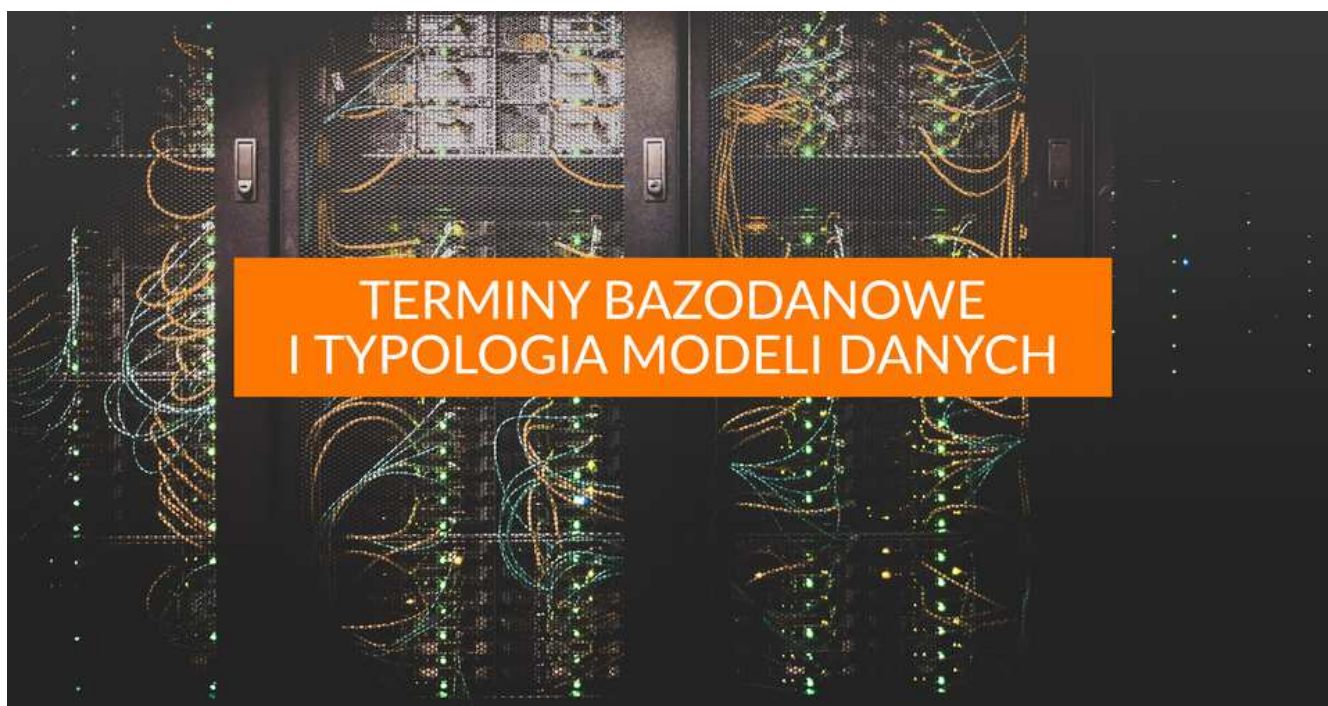
polimorfizm

(inaczej: wielopostaciowość) mechanizm umożliwiający posługiwanie się tą samą metodą w kilku różnych kontekstach, zależnie od zaistniałych okoliczności; działanie metody zależy na przykład od typu dostarczonych jej danych wejściowych

Film edukacyjny

Polecenie 1

Zapoznaj się z filmem, dzięki któremu lepiej zrozumiesz terminologię związaną z bazami danych oraz poznasz typologię modeli danych.



Film dostępny pod adresem </preview/resource/RjLBfJLTUwypQ>

Źródło: Contentplus.pl Sp. z o.o., licencja: CC BY-SA 3.0.

Film przedstawia terminy bazodanowe i typologia modeli danych.

Sprawdź się

Pokaż ćwiczenia:   

Ćwiczenie 1



Ćwiczenie 2



Ćwiczenie 3



Ćwiczenie 4



Ćwiczenie 5



Ćwiczenie 6



Ćwiczenie 7



Ćwiczenie 8



Dla nauczyciela

Autor: Mirosław Zelent

Przedmiot: Informatyka

Temat: Wprowadzenie do baz danych

Grupa docelowa:

Szkoła ponadpodstawowa, liceum ogólnokształcące, technikum, zakres podstawowy

Podstawa programowa:

Cele kształcenia – wymagania ogólne

II. Programowanie i rozwiązywanie problemów z wykorzystaniem komputera oraz innych urządzeń cyfrowych: układanie i programowanie algorytmów, organizowanie, wyszukiwanie i udostępnianie informacji, posługiwanie się aplikacjami komputerowymi.

Treści nauczania – wymagania szczegółowe

II. Programowanie i rozwiązywanie problemów z wykorzystaniem komputera i innych urządzeń cyfrowych.

Zakres podstawowy. Uczeń:

3) przygotowuje opracowania rozwiązań problemów, posługując się wybranymi aplikacjami:

d) wyszukuje informacje, korzystając z bazy danych opartej na co najmniej dwóch tabelach, definiuje relacje, stosuje filtrowanie, formułuje kwerendy, tworzy i modyfikuje formularze, drukuje raporty,

Kształtowane kompetencje kluczowe:

- kompetencje cyfrowe;
- kompetencje osobiste, społeczne i w zakresie umiejętności uczenia się;
- kompetencje matematyczne oraz kompetencje w zakresie nauk przyrodniczych, technologii i inżynierii.

Cele operacyjne (językiem ucznia):

- Zdefiniujesz pojęcia: informacja, baza danych, system zarządzania bazą danych, model danych, język SQL.
- Sklasyfikujesz rozwijane przez lata rodzaje modeli danych.

- Opisziesz podstawowe założenia i definicje związane z najczęściej stosowanym obecnie relacyjnym modelem danych.

Strategie nauczania:

- konstruktywizm;
- konektywizm.

Metody i techniki nauczania:

- dyskusja;
- rozmowa nauczająca z wykorzystaniem multimediu i ćwiczeń interaktywnych;
- ćwiczenia praktyczne.

Formy pracy:

- praca indywidualna;
- praca w parach;
- praca w grupach;
- praca całego zespołu klasowego.

Środki dydaktyczne:

- komputery z głośnikami, słuchawkami i dostępem do internetu;
- zasoby multimedialne zawarte w e-materiale;
- tablica interaktywna/tablica, pisak/kreda.

Przebieg lekcji

Przed lekcją:

1. **Przygotowanie do zajęć.** Nauczyciel loguje się na platformie i udostępnia e-materiał: „Wprowadzenie do baz danych”. Nauczyciel prosi uczniów o zapoznanie się z treściami w sekcji „Przeczytaj”.

Faza wstępna:

1. Ustalenie celu lekcji i kryteriów sukcesu.
2. **Rozpoznanie wiedzy uczniów.** Nauczyciel zadaje uczniom pytania dotyczące ich aktualnego stanu wiedzy w obszarze poruszanego tematu, np.
 - czym jest informacja?
 - jakie znasz bazy danych?
 - czym jest hermetyzacja?
 - czym jest polimorfizm?Chętni lub wybrani uczniowie udzielają na nie odpowiedzi.

Faza realizacyjna:

1. **Praca z multimedium.** Nauczyciel wyświetla zawartość sekcji „Film samouczek”. Uczniowie wspólnie zapoznają się z treścią filmu, dzięki któremu przyswoją sobie podstawową nomenklaturę bazodanową oraz poznają typologię modeli danych. Zapisują ewentualne problemy i pytania. Po czym następuje dyskusja, w trakcie której nauczyciel wyjaśnia niezrozumiałe treści.
2. **Ćwiczenie umiejętności.** Prowadzący zapowiada uczniom, że w kolejnym kroku będą rozwiązywać ćwiczenia nr 1-5 z sekcji „Sprawdź się”. Każdy z uczniów robi to samodzielnie. Po ustalonym czasie wybrani uczniowie przedstawiają rozwiązania. Nauczyciel w razie potrzeby koryguje odpowiedzi, dopowiada istotne informacje, udziela uczniom informacji zwrotnej.

Faza podsumowująca:

1. Nauczyciel ponownie wyświetla na tablicy temat lekcji zawarty w sekcji „Wprowadzenie” i inicjuje krótką rozmowę na temat zrealizowanych celów (czego uczniowie się nauczyli).
2. Na koniec zajęć nauczyciel prosi uczniów o rozwinięcie zdania: „Na dzisiejszych zajęciach nauczyłam/łem się jak...”.

Praca domowa:

1. Uczniowie wykonują ćwiczenia 6-8 z sekcji „Sprawdź się”.

Wskazówki metodyczne:

- Nauczyciel może wykorzystać multimedium w sekcji „Przeczytaj” do pracy przed lekcją. Uczniowie zapoznają się z jego treścią i przygotowują do pracy na zajęciach w ten sposób, żeby móc samodzielnie rozwiązać zadania dołączone do e-materiału „Wprowadzenie do baz danych”.