

## BAZY DANYCH

Wykład 1  
dr inż. Agnieszka Bołtuć

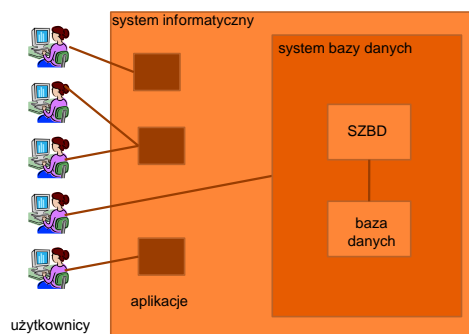
### ORGANIZACJA ZAJĘĆ

- **Wykładowca** – dr inż. Agnieszka Bołtuć, pokój 304,  
e-mail: [aboltuc@ii.uwb.edu.pl](mailto:aboltuc@ii.uwb.edu.pl)
- **Liczba godzin i forma zajęć**: 30 godzin wykładu oraz  
30 godzin laboratorium
- **Konsultacje**: czwartek 10:15-12:00
- **Wykłady w wersji elektronicznej**:  
[ii.uwb.edu.pl/~aboltuc](http://ii.uwb.edu.pl/~aboltuc)
- **Forma zaliczenia wykładu**: egzamin, część  
teoretyczna + zadanie projektowe

### PLAN WYKŁADU

- Podstawowa terminologia i definicje
- Charakterystyka baz danych
- Użytkownicy
- System zarządzania bazą danych
- Model danych
- Architektura systemu baz danych
- Architektura komunikacyjna
- Podział baz danych
- Tworzenie baz danych
- Historia zarządzania danymi

### SCHEMAT SYSTEMU INFORMATYCZNEGO



## SYSTEM BAZ DANYCH

### Składniki systemu baz danych

- Baza danych,
- Model danych,
- System zarządzania bazą danych (SZBD) (*ang. database management system, DBMS*)

## BAZA DANYCH - DEFINICJA

- Baza danych (*ang. database, DB*) to zbiór danych zapisanych w ściśle określony sposób w strukturach odpowiadających założonemu modelowi danych.

źródło: Wikipedia

- Zbiór danych reprezentujący wybrany fragment rzeczywistości.

źródło: Banachowski

- Zbiór danych trwałych wykorzystywanych przez system aplikacji danej organizacji (firma, bank, szpital, etc.)

źródło: Date

## GŁÓWNE CECHY BAZY DANYCH

- **Trwałość** – najczęściej dane są przechowywane przez długi okres czasu, zazwyczaj nieokreślony,
- **Duże wolumeny** – duża liczba danych wymaga innych metod dostępu niż liniowe,
- **Zgodność z rzeczywistością** – w związku z tym, iż baza stanowi odzwierciedlenie fragmentu rzeczywistości, to niezbędne jest zapewnienie zgodności danych z bazy z tymi rzeczywistymi, które ulegają zmianom:
  - aspekt administracyjny – aktualizacja danych w chwili ich zmiany w rzeczywistości,
  - aspekt wewnętrzny – dane winny spełniać określone warunki: reguły rzeczywistości, biznesu (np. wiek <120)

## WŁAŚCIWOŚCI BAZY DANYCH

- **Spójność danych** – poprawność danych z punktu widzenia kryteriów: wiernego odzwierciedlenia danych rzeczywistych, spełnienie ograniczeń nałożonych przez użytkowników, odporności na anomalie dostępu, awarie sprzętowo-programowe oraz błędy użytkowników,
- **Współdzielenie danych** – dane przechowywane w bazie powinny być dostępne dla wielu osób w tym samym czasie, należy zapewnić rozwiązywanie konfliktów typu zapis-zapis,
- **Integracja danych** – baza danych powinna być zbiorem nie posiadającym danych redundancyjnych; **dane redundancyjne** to dane powtarzające się,

## WŁAŚCIWOŚCI BAZY DANYCH

- **Integralność danych** – baza powinna zapewniać aktualizację danych po obu stronach związków będących odzwierciedleniem rzeczywistości, zapobiegnie to sprzeczności pomiędzy dwoma pozycjami reprezentującymi ten sam fakt, powinna istnieć możliwość nakładania ograniczeń integralnościowych,
- **Bezpieczeństwo danych** – aby zapewnić integralność danych należy zabezpieczyć dostęp do bazy, wprowadzenie autoryzacji dostępu: dostęp tylko dla użytkowników identyfikowanych unikalną nazwą i hasłem, a także przydzielanie określonych uprawnień do określonych części bazy danych,

## WŁAŚCIWOŚCI BAZY DANYCH

- **Niezależność danych** – odporność aplikacji na zmiany w strukturze przechowywania danych i sposobie dostępu, cecha jedynie częściowo osiągnięta,
- **Standaryzacja danych** – ułatwia wymianę danych i ich przenoszenie pomiędzy systemami, ważne są także standardy w nazywaniu danych i prowadzeniu dokumentacji,
- **Metadane** – baza powinna dawać wsparcie dla metadanych (danych o danych),

## UŻYTKOWNICY

Grupy zasadnicze:

- Użytkownicy (z prawem odczytu, z prawem odczytu i modyfikacji),
- Programiści aplikacji,
- Projektanci baz danych,
- Analitycy systemowi,
- Administratorzy bazy danych.

Grupy dodatkowe:

- Administratorzy serwerów, sieci komputerowych,
- Projektanci i programiści SZBD,
- Inżynierowie wiedzy,
- Projektanci narzędzi developerskich.

## SYSTEM ZARZĄDZANIA BAZĄ DANYCH (SZBD)

System zarządzania bazą danych jest zorganizowanym zbiorem narzędzi umożliwiającym dostęp i zarządzanie jedną lub więcej bazami danych.

Do podstawowych funkcji realizowanych przez SZBD należą:

- **Definiowanie i obróbka danych** – wsparcie dla języka bazy danych, tworzenie, aktualizacja, usuwanie danych,
- **Struktury danych** – wsparcie dla struktur danych zapewniających składowanie i przetwarzanie dużych wolumenów danych,

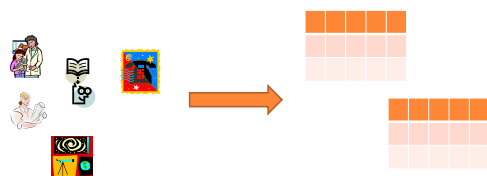
## SYSTEM ZARZĄDZANIA BAZĄ DANYCH (SZBD)

- **Bezpieczeństwo i integralność danych** - zapewnienie bezpieczeństwa danych w przypadku awarii sprzętowo-programowej, zapewnienie autoryzacji dostępu do danych, zapewnienie nienaruszalności integralności danych,
- **Odzyskiwanie danych i współbieżność** – wsparcie dla odzyskiwania danych oraz synchronizacji współbieżnego dostępu do danych,
- **Wydajność** - efektywne przetwarzanie i składowanie danych,
- **Różnorodność interfejsów.**

## SYSTEM ZARZĄDZANIA BAZĄ DANYCH (SZBD)

- Oracle,
- IBM,
- Microsoft,
- Sybase,
- Access,
- MySQL,
- PostgreSQL,
- FireBird.

## MODEL DANYCH - DEFINICJA



Obiekty ze świata rzeczywistego są reprezentowane w bazie danych za pomocą tzw. **modelu danych**. Model danych jest zbiorem ogólnych zasad posługiwania się danymi (definicja danych, operowanie danymi, integralność danych).

## MODEL DANYCH - RODZAJE

- **hierarchiczny** - dane są przechowywane na zasadzie rekordów nadrzędnych-podrzędnych, tzn. rekordy przypominają strukturę drzewa (nie stosowany),
- **sieciowy** – następca hierarchicznego, struktura danych to sieć (nie stosowany),
- **relacyjny** – dane prezentowane jako relacje posiadające atrybuty, zapewnia integralność danych poprzez zastosowanie kluczy,
- **obiektowy** – baza jest zbiorem obiektów posiadających pola i metody, zapewnia dziedziczenie, hermetyzację, itd,
- **obiekto-relacyjny** – połączenie własności modelu relacyjnego i obiektoowego,
- **semistrukturalny** – z wykorzystaniem XML.

## MODEL DANYCH - ELEMENTY

- o **struktury danych** – niezbędna do reprezentowania w bazie danych obiektów rzeczywistych,



Imię i nazwisko	Nr telefonu
Anna Gosk	600101101
Jerzy Kilt	500345687

- o **operacje na danych** – modele posiadają pewne z góry zdefiniowane operacje możliwe do wykonania na danych,
- o **ograniczenia integralnościowe** – modele umożliwiają nałożenie ograniczeń integralnościowych na reprezentowane w nim dane, np.:

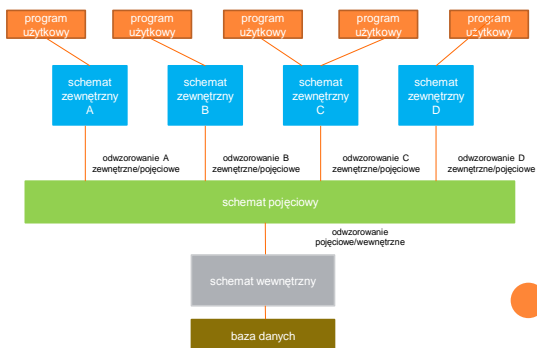
*Data\_zatrudnienia > data\_urodzenia*  
120 > Wiek > 0

## ARCHITEKTURA SBD

Architektura systemu bazy danych została zdefiniowana w standardzie **ANSI/SPARC**. Wyróżniamy w niej trzy poziomy:

- o **wewnętrzny (fizyczny)** – dotyczy fizycznego sposobu przechowywania danych oraz dostępu do danych,
- o **zewewnętrzny (użytkownika)** – dotyczy sposobu w jaki użytkownicy widzą dane,
- o **pojęciowy (konceptyjny)** – pośredni pomiędzy powyższymi, opisuje widok danych postrzegany przez organizację, związki między danymi, etc.

## ARCHITEKTURA ANSI/SPARC



## ARCHITEKTURA KOMUNIKACYJNA - RODZAJE

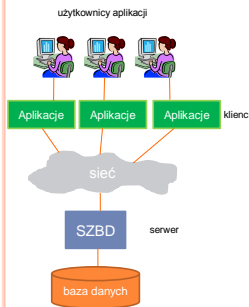
Na architekturę systemu baz danych składają się: opis elementów stanowiących system baz danych, sposób w jaki są one ze sobą powiązane oraz logika tej organizacji.

Rodzaje:

architektura klient-serwer

- o dwuwarstwowa,
- o trójwarstwowa.

## ARCHITEKTURA KOMUNIKACYJNA – DWUWARSTWOWA



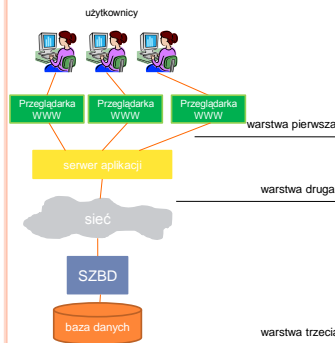
Zalety:

- bezpieczeństwo serwera i danych,
- odciążenie centralnego komputera od obsługi interfejsu użytkownika,
- istnieje wiele technologii wspomagających takie rozwiązania.

Wady:

- trudniejsze administrowanie (wiele komputerów i aplikacji),
- brak kontroli nad użytkownikami i koszty eksploatacji komputerów,
- możliwe problemy z przepustowością,
- całkowity brak dostępu do danych w przypadku awarii serwera.

## ARCHITEKTURA KOMUNIKACYJNA – TRÓJWARSTWOWA



Zalety:

- większość zalet architektury klient-serwer,
- prostsze projektowanie i implementacja.

Wady:

- „silniejszy” sprzęt na serwery aplikacyjne,
- trudniejsze technologie.

## PODZIAŁ BAZ DANYCH

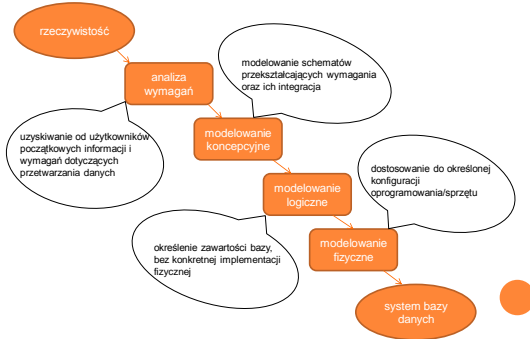
W zależności od przyjętego kryterium bazy można podzielić na:

- o wg modelu danych
  - o hierarchiczne,
  - o sieciowe,
  - o relacyjne,
  - o obiektowe,
  - o obiektowo-relacyjne,
  - o semistrukturalne.
- o wg liczby baz danych
  - o scentralizowane,
  - o rozproszone.

## PODZIAŁ BAZ DANYCH

- o wg zastosowań
  - o bazy danych przetwarzania transakcyjnego (ang. *On-Line Transaction Processing, OLTP*),
  - o bazy danych przetwarzania analitycznego (ang. *On-Line Analytical Processing, OLAP*),
  - o bazy danych wspomagające projektowanie (ang. *Computer Aided Design, CAD*),
  - o bazy danych do przechowywania informacji geograficznej (ang. *Geographical Information Systems, GIS*),
  - o bazy danych do wspomaganie wytwarzania oprogramowania (ang. *Computer Aided Software Engineering, CASE*).

## TWORZENIE BAZ DANYCH



## TWORZENIE BAZ DANYCH

Podstawowe scenariusze powstawania bazy danych:

- **Przyrostowy** („kawałek-po-kawałku”) – dla każdego działu lub odcinka działalności firmy tworzy się oddzielna bazę i system informacyjny,
  - zalety: sprawdza się przy braku doświadczenia, szybkie powstawanie,
  - wady: redundancja, brak pełnego obrazu stanu firmy, brak możliwości wykonywania zadań globalnych,
- **Zintegrowanego systemu** – trudny, ale brak wad poprzedniego.

## KOMUNIKACJA Z BAZĄ

- Język SQL,
- Aplikacje: formularze, raporty.

## HISTORIA ZARZĄDZANIA DANYMI (WG GRAYA)

- **Zarządzanie ręczne (4000 p.n.e- 1900 n.e)**  
zapis opisujący majątek królewski i podatki w Sumerze notowany jest na ok. 4000 p.n.e,
- **Zapisy na kartach perforowanych (1900-1955)**  
rok 1800 – wynalezienie mechanizmu Jacquarda dotyczącego wytwarzania wyrobów na podstawie wzorów zakodowanych na kartach perforowanych, rok 1890 – Herman Hollerith zastosował karty przy opracowaniu spisu ludności USA i założył przedsiębiorstwo produkujące taki sprzęt (później IBM),

## HISTORIA ZARZĄDZANIA DANYMI (WG GRAYA)

### o Zarządzanie programowane (1955-1970)

początek programów obliczeniowych, wzrost wydajności poprzez wynalezienie taśmy magnetycznej (1 taśma=10000 kart perforowanych), przetwarzanie danych zorganizowanych w rekordy i zbiory, później powstały wsadowe systemy przetwarzania transakcji – dane zorganizowane we wsady były sortowane raz dziennie, wady: opóźnione wykrywanie błędów, brak możliwości poznania bieżących danych

## HISTORIA ZARZĄDZANIA DANYMI (WG GRAYA)

### o Sieciowe zarządzanie on-line (1965-1980)

nowe urządzenie do przechowywania danych (bębny, dyski magnetyczne), rozwój przetwarzania zdalnego, ulepszenie struktur plików, pozwalało to na przeprowadzanie bieżących operacji, model DBTG (Charles Bachman) zdefiniował standard języka definiowania i operowania danymi,

## HISTORIA ZARZĄDZANIA DANYMI (WG GRAYA)

### o Relacyjne zarządzanie danymi (1980-1995)

opracowany przez Codda w 1970 roku, pierwszy prototyp o nazwie system R zbudowany przez IBM, rok 1983 – IBM prezentuje bazę DB2, obecnie wiele innych popularnych,

### o Multimedialne bazy danych (1995-)

przechowywanie danych typu głos, wideo, obrazy, duże zapotrzebowanie na bazy oparte na pojęciu obiektowości.

## WYKŁAD PRZYGOTOWANO NA PODSTAWIE

- o C. J. Date, Wprowadzenie do systemów baz danych, WNT, Warszawa, 2000,
- o [http://wazniak.mimuw.edu.pl/index.php?title=Bazy\\_danych](http://wazniak.mimuw.edu.pl/index.php?title=Bazy_danych)
- o L. Banachowski, Bazy danych tworzenie aplikacji. PLJ, Warszawa, 1998,
- o P. Beynon-Davies, Systemy baz danych, WNT, 2003.