

## *Baze podataka (usmeni)*

- **Baza podataka** – organizirana kolekcija podataka spremljena elektronički na računalnom sustavu
- **Model podataka** – skup pravila koja određuju kako sve može izgledati logička struktura baze podataka
- **Vrste modela:** mrežni, hijerarhijski, relacijski, objektni, polustrukturirani
- **Relacijski model podataka** - kolekcije tablica koje predstavljaju podatke i veze između podataka
  - Tablice se nazivaju relacije
  - Svaka tablica definira n-torku određenog tipa koja se sastoji od fiksnog broja atributa
- **relacija** – naziv tablice u relacijskom modelu podataka
- **atribut** – jedan „stupac” relacije
- **domena atributa** - definira skup dozvoljenih vrijednosti za atribut
- **n-torka** - je uređeni skup parova naziv-vrijednost, fiksne veličine
- **shema relacije** - svaka relacija ima shemu, koja sadrži: uređeni skup atributa i domene za svaki atribut
- **shema baze podataka** - nacrt strukture baze podataka koja predstavlja logički dizajn baze podataka
- **shema dijagram** – grafički prikaz sheme baze podataka
- **ključevi** - ključevi se koriste za razlikovanje pojedinih n-torki unutar relacije
- **super ključ** - je skup atributa koji zajedno, jednoznačno definiraju svaku n-torku unutar relacije
- **kandidat ključ** – minimalni super ključ
- **primarni ključ** - jedan od kandidata ključeva se odabire kao primarni način za jedinstveno identificiranje n-torki (dvije skupine: prirodni ključ i surogat ključ)
- **strani ključ** – primarni ključ druge relacije
- **dvije vrste jezika za rad s bazama podataka:**
  - DDL – specificiranje sheme baze podataka (ograničenja: domene, referencijalni integritet, opća ograničenja, dozvole)
  - DML – zadavanje upita i izmjena podataka (proceduralni i deklarativni)
- **Relacijska algebra** – proceduralni upitni jezik, sastoji se od 6 relacijskih operacija (selekcija, projekcija, unija, razlika, kartezijev produkt, preimenovanje), predstavlja matematičku relaciju
- Svaka operacija uzima jednu ili dvije relacije, a konačni rezultat je stvaranje nove relacije
- **Dotadne operacije:** presjek, natural join, theta join, outer join, left join, right join, full join, dijeljenje, dodjeljivanje
- **Proširene operacije relacijske algebre:**
  - **Generalizirana projekcija** – koristeći operaciju mogu se specificirati atributi, kao i funkcije na atributima
  - **Funkcijske agregacije** – različite funkcije (korak grupiranja, korak primjene funkcije)
- **Null vrijednost** – nepoznata ili nepostojeća vrijednost
- **Operator dodjeljivanja** – promjena podataka

- SQL relacije su sadržane unutar baza podataka
- **Domene atributa:**
  - *Char* – polje znakova fiksne veličine (niz otp. točno iste dužine),
  - *Varchar* – polje znakova varijabilne veličine (nizovi),
  - *Int* – cijeli broj,
  - *Numeric* – broj s korisnički specificiranom preciznosti,
  - *Float* – vrijednost s najmanje N bitova preciznosti,
  - *Date, Time, Datetime* – pohrana vremenskih podataka,
  - *Blob, Clob, Text* – velika tekstualna polja.
- **Stvaranje i korištenje baze** – *Create database* (kreiranje) *Use* (korištenje)
- **Stvaranje tablica unutar baze** – *Create table*
- **Stvaranje redova** – *Insert into... values* (tablice se stvaraju prazne pa se naknadno dodaju redovi)
- **Delete i drop** – brisanje redova i brisanje tablice
- **Alter table** – izmjena strukture unutar tablice; dodavanje i brisanje redova
- Korištenjem \* u naredbi *Select* kažemo da želimo dohvatiti sve attribute relacije
- **Distinct** – uklanjanje duplikata
- **Operacije sa skupovima:** Unija i except (ne koristi se u SQL-u)
- **Sortiranje** – *Order by (asc i desc)*
- **Grupiranje** – *Group by*
- **Agregiranje** – sum, avg, count, min, max
- **Filtriranje** – *Having*
- **Ograničenja** – skup pravila koja imaju cilj osiguravanja kvalitete podataka pohranjenih u bazu podataka:
  - Primarni ključ – baza ne omogućava unošenje iste vrijednosti primarnog ključa (*Primary key*)
  - Strani ključ – primarni ključ tablice r = strani ključ unutar tablice s (*Foreign key*)
  - *Unique* - jedinstven
  - *Check* – postavljanje uvjeta na vrijednost podataka, sprječava unos neispravnih podataka
  - *Not null* – ne pohranjuje null vrijednost
  - DDL ograničenja
- **DDL ograničenja:**
  - a) na domene (deklariranje atributa (int, char..)),
  - b) na referencijalni integritet (provjera integriteta vrijednosti u stranim ključevima),
  - c) na općim ograničenjima (uvjeti koje baza mora zadovoljavati (npr. 20 kolegija mora imati svaki smjer)),
  - d) na dozvole (prava korisnika na pristup podacima)
- **SQL upiti (relacijska algebra):**
  - *Select* – odgovara projekciji
  - *From* – odgovara Kartezijevom produktu
  - *Where* – odgovara selekciji

- **Logički redosljed:**
  1. From → on → outer
  2. Where → group by → having
  3. Select → distinct → order by → top
- **Korelirani podupit** - kada ugniježđeni podupit referencira attribute iz relacije vanjskog upita
- **Ugniježđeni podupiti** – select unutar select-a; upit unutar upita (može biti u select, where ili from dijelu)
- **Skalarni podupit** - podupiti mogu kao rezultat vratiti jednu vrijednost; to je relacija s jednim atributom i jednom n-torkom (select min (primanja) from nastavnik)
- **Standardni mehanizam** ima tri razine: katalog, shema, tablica
- **E-R dijagram/ E-R model/Shema skupova veza:**
  - definira složenu shemu u grafički oblik; koristi skup osnovnih objekata entitet i veze
  - model za dizajniranje sheme, koristi entitete i veze (e-r dijagram)
- **Entitet** – stvar, objekt u stvarnom svijetu koja se razlikuje od ostalih objekata
- **Skup entiteta** – imenovana kolekcija entiteta istog tipa
- **Veze** – asocijacija između dva ili više entiteta
- **Skup veza** – imenovana kolekcija veza istog tipa
- **E-R dijagram ograničenja:**
  - **kardinalnost mapiranja** – koliko drugih entiteta može biti povezano s entitetom preko određenog skupa- veza:
    - a) one to one (jedan entitet iz A skupa povezan je s jednim entitetom iz B, jedan nastavnik jedan kolegiji),
    - b) one to many (jedan iz A s više iz B ili jedan iz B s jednim iz A, jedan nastavnik – više kolegija),
    - c) many to one (jedan iz A s jednim iz B ili jedan iz B s više iz A, svaki nastavnik može predavati samo jedan kolegij)
    - d) many to many (više iz A s više iz B i obratno, nastavnik može držati više kolegija)
  - **Ograničenje ključeva u skupovima-entiteta** – npr. ključ kandidat ograničava vrijednost ključnih atributa
  - **Ograničenje uključenosti** – ako svaki entitet E sudjeluje u najmanje jednoj vezi R sudjelovanje je potpuno (npr. svaki kredit je vezan uz klijenta), a ako samo neki sudjeluju tada je djelomično (odnos klijenta i kredita tj. klijent ne mora imati kredit)
- **Opisni atribut** – opisuje određenu vezu unutar E-R dijagrama
- **slabi skupovi-entiteta** – ponekad skup – entiteta nema atributa po kojima bi se mogao razlikovati (nije moguće definirati primarni ključ za skup - entiteta), zato slabi skup treba biti povezan s drugim jakim skupom – entiteta
- identifikacijska veza – odnos između dva entiteta (npr. many-to-one bez deskriptivnih atributa)

- **primarni ključ za binarne skupove-veza** (između jakih skupova-entiteta A i B):
  - many-to-many mapiranje – unija svih primarnih ključeva skupova-entiteta postaje primarni ključ skupa-veza
- **konverzija E-R dijagrama u relacijski model:**
  - razlika je što kod E-R dijagrama imamo složene i višeznačne attribute, dok kod relacijskog modela imamo jednostavne
  - **tri komponente procesa konverzije:**
    - a) specifikacije sheme relacije
    - b) specifikacija primarnog ključa
    - c) specifikacija reference stranih ključeva na druge relacije
- **normalne forme** definiraju "dobre" obrasce shema
- **Boyce-Codd normalna forma** - eliminira redundantnost koristeći funkcijske zavisnosti
- **Prva normalna forma (1NF)** - svi atributi moraju biti atomarni - jednostavno u relacijskom modelu, teže ili manje poželjno u SQL-u
- **Trivijalne zavisnosti** – zadovoljavaju sve vrijednosti unutar relacije
- **Selekcija** – ako P rezultira s *null* tada se ista n-torka isključuje iz rezultata
- **Instanca relacije** – označava specifičnu instancu (stanje) neke relacije, tj. relaciju koja sadrži specifični skup n-torki
- **Identificiranje veza** – određene veze identificiraju se isključivo po entitetima koji sudjeluju u njima
- **Referencijski integritet** – konzistentnost između dvije zavisne relacije
- **pravila inferencije za zavisnosti** se nazivaju aksiomi – grupe pravila:
  - a) refleksivnost,
  - b) uvećanje i
  - c) tranzitivnost,
 i dodatna pravila:
  - a) unije
  - b) dekompozicije
  - c) pseudotranzitivnosti
- **natural outer join** – dvije relacije koje se spajaju koriste se svi zajednički atributi, a onim atributima koji nisu upareni dodaje null vrijednost i to se sve uključuje u rezultat
- **dekompozicija s gubitkom** – neke dekompozicije gube informacije
  - djelatnik\_detalji (ime\_prezime, broj mobitela, titula...) pošto nemamo id ili OIB samo ime\_prezime ne može definirati jednostavnost
- **kršenje stranog ključa:**
  - a) ako referencirajuća relacija dobije pogrešnu vrijednost stranog ključa, operacija jednostavno nije dopuštena
  - ako se u tablicu predaje pokušava dodati red koji ima vrijednost nastavnik\_id od nepoznatog nastavnika (nije evidentiran u nastavnik tablici)
  - ako se u tablicu predaje pokušava dodati red koji ima vrijednost kolegij\_sifra od nepoznatog kolegija (nije evidentiran u kolegij tablici)
  - b) kada se rade promjene u referenciranoj relaciji
  - ako se iz tablice nastavnik ukloni red koji je referenciran u tablici predaje