

## DUOMENŲ BAZĖS

### Įvadas

Šiame kurse bus pateiktos žinios apie duomenų bazes (DB), informacines sistemas (IS), duomenų bazių valdymo sistemas. Studentams bus aiškinama kaip sukurti informacinę sistemą. Pagrindinis dėmesys, kaip matosi iš kurso pavadinimo, bus skiriamas duomenų bazių projektavimo klausimams. Taip pat bus pateikti SQL kalbos pagrindai.

Pavyzdžiai bus rodomi su Oracle korporacijos sukurta programinė įranga. Oracle DB buvo pasirinkta, nes tai yra vienas iš lyderių savo srityje. Oracle produktais naudojasi daug stambių įmonių ir organizacijų. Lietuvoje tai yra tokios įmonės kaip SEB bankas, TEO LT, Registrų centras, Lietuvos energija ir kt. Naudoja Oracle ir mūsų universitetas VGTU.

Daugiausiai dėmesio bus skiriama DB projektavimui, nes tai yra vienas iš esminių IS kūrimo etapų. Nuo DB projektavimo kokybės dažnai priklauso ir viso IS kūrimo projekto tolimesnė sėkmė. Studentams bus išaiškinti pagrindiniai reliacinių DB kūrimo principai. Atskiras dėmesys bus skirtas DB normalizacijai.

Kurso pabaigoje studentai bus supažindinti su tokiais temomis, kaip duomenų saugumas, duomenų importas bei eksportas, duomenų saugyklos.

### Pagrindinės DB sąvokos

Gyvenime mums dažnai tenka susidurti su būtinybe saugoti duomenis. Pavyzdžiui, mes turime užrašų knygeles, kur fiksuojame pažįstamų telefonus bei adresus, perkame įvairius žinytus, kur irgi yra pateikta nemažai susistemintos informacijos. Universitete yra duomenų bazė, kur saugoma informacija apie grupes, studentus, jų pažangumą ir pan. Bankai turi duomenų bazes, kur saugoma informacija apie klientų sąskaitas. Pavyzdžius galima būtų vardinti dar ilgai, kas rodo, kad duomenų bazės užima svarbią vietą žmonių veikloje.

Pagrindinė DB sąvoka yra lentelė. Tarkime, norime sukurti lentelę, kurioje būtų atvaizduoti tam tikros grupės studentai. Tokios lentelės pavyzdys parodytas apačioje:

<b>PAVARDE</b>	<b>VARDAS</b>	<b>GRUPE</b>
Pavardė_1	Vardas_1	TM-09/1
Pavardė_2	Vardas_2	TM-09/1
Pavardė_3	Vardas_3	TM-09/1
Pavardė_4	Vardas_4	TM-09/1
Pavardė_5	Vardas_5	TM-09/1

Kaip matome, sukurtoje lentelėje yra trys stulpeliai, kuriuose talpinama informacija apie studentus (šiuo atveju vardas, pavardė ir grupė). Lentelėje yra 5 eilutės, t.y. pateikta informacija apie 5 studentus. Bet vardas ir pavardė nepilnai identifikuoja studentą. Netgi toje pačioje grupėje gali būti bendravardžių ir bendrapavardžių. Todėl ši lentelė turi būti papildyta lauku (arba keliais laukais), kurių reikšmė lentelėje negali

kartotis. Pavyzdžiui, papildykime mūsų lentelę lauku PAZ\_NR, kuriame bus saugomas studento pažymėjimo numeris, kuris yra skirtingas kiekvienam studentui.

PAZ_NR	PAVARDE	VARDAS	GRUPE
529147	Pavardė_1	Vardas_1	TM-09/1
521948	Pavardė_2	Vardas_2	TM-09/1
521949	Pavardė_3	Vardas_3	TM-09/1
521950	Pavardė_4	Vardas_4	TM-09/1
521951	Pavardė_5	Vardas_5	TM-09/1

Laukas PAZ\_NR vadinamas unikaliu lentelės STUDENTAI raktu. Šio lauko reikšmė negali kartotis, t.y. lentelėje STUDENTAI negali būti dviejų įrašų su vienodu PAZ\_NR.

Kartais unikalų raktą tenka kurti ne iš vieno, bet iš dviejų ir daugiau laukų. Sakykime lentelėje fiksuojame informaciją apie kažkokius dokumentus. Kiekvienas dokumentas identifikuojamas pagal numerį, kuris savo ruožtu formuojamas iš dviejų skaičių: pirmas skaičius yra kalendoriniai metai (metai, kuriais buvo sukurtas dokumentas), o antras eilės numeris, priskiriamas pagal dokumentų pasirašymo eiliškumą. Kitais žodžiais tariant, mes galime turėti penktą dokumentą 2011 metais ir penktą dokumentą 2010 metais, bet negalime turėti, sakykime, ketvirto dokumento 2008 metais. Tokios lentelės pavyzdys pateiktas apačioje.

METAI	NUMERIS	...	...
2010	1	...	...
2010	2	...	...
2010	3	...	...
2011	1	...	...
2011	2	...	...
2011	3	...	...

Šiuo atveju unikalų raktą sudaro iš karto du laukai: METAI ir NUMERIS. Kiekvieno iš šių laukų reikšmės gali kartotis. Pavyzdžiui mes turime 3 įrašus su lauko METAI reikšme lygia 2010, taip pat 2 įrašus su lauko NUMERIS reikšme 2. Bet dviejų laukų METAI ir NUMERIS kombinacija yra unikali visai lentelei. T.y. mes negalime įrašyti antro įrašo, kad lauko METAI reikšmė būtų 2010, o lauko NUMERIS reikšmė 3.

Be unikalios raktų lentelėje dažniausiai sukuriamas ir pirminis raktas. Susitarkime, kad kiekvienai lentelei visada sukursime pirminį raktą ir pavadinsime jį ID. Pirminis raktas skirtas identifikuoti įrašą ir jo reikšmė (kaip ir unikalios raktų) negali kartotis.

Papildykime lentelę STUDENTAI pirminiu raktu:

ID	PAZ_NR	PAVARDE	VARDAS	GRUPE
101	529147	Pavardė_1	Vardas_1	TM-09/1
102	521948	Pavardė_2	Vardas_2	TM-09/1

103	521949	Pavardė_3	Vardas_3	TM-09/1
104	521950	Pavardė_4	Vardas_4	TM-09/1
105	521951	Pavardė_5	Vardas_5	TM-09/1

Dažniausias pirminis raktas yra tiesiog skaičius, kuris automatiškai nuosekliai suteikiamas kiekvienam naujam įrašui. Jis neturi jokios semantinės prasmės. Dažniausiai DB vartotojas net nežino (ir nesidomi) šio skaičiaus reikšmės. Bet duomenų bazėje pirminio rakto vaidmuo yra svarbus, nes per jį atliekamas lentelių sąryšis. Pailiustruokime tai paprastu pavyzdžiu.

Sakykime, kad norime turėti ne tik informaciją apie studentus, bet ir apie tai, kiek užsienio kalbų jie moka. Tuomet prie jau esamos lentelės STUDENTAI galime sukurti lentelę STUDENTU\_KALBOS

ID	STUDENTO_ID	KALBA
1025	101	anglų
1026	101	vokiečių
1027	101	rusų
1028	105	anglų
1029	105	lenkų
1030	102	prancūzų

Lentelėje STUDENTU\_KALBOS yra laukas STUDENTO\_ID, kuris vadinamas išoriniu raktu. Šitas laukas turi ryšį su lentelės STUDENTAI pirminiu raktu ID. Lentelėje STUDENTU\_KALBOS nėra tiesioginės informacijos apie studentą. T.y. mes neturime čia nei studento pažymėjimo numerio, nei pavardės, nei vardo, nei grupės. Bet visą šią informaciją galima lengvai gauti iš lentelės STUDENTAI. Čia galioja tokia taisyklė:

**STUDENTAI.ID = STUDENTU\_KALBOS.STUDENTO\_ID**

Sakykime, kad mus domina, kiek užsienio kalbų moka studentas, kurio pažymėjimo numeris yra 529147. Iš lentelės STUDENTAI, matome, kad įrašo, kur saugoma informacija apie šį studentą pirminio rakto reikšmė ID = 101. Tuomet lentelėje STUDENTU\_KALBOS surandame visus įrašus su išorinio rakto reikšme STUDENTO\_ID = 101. Gauname, kad studentas moka 3 užsienio kalbas: anglų, vokiečių ir rusų.

Reikia atkreipti dėmesį, kad išorinio rakto reikšmės gali kartotis, bet jos negali būti kitokios, negu atitinkamo pirminio rakto reikšmės. T.y. pagal įvestus duomenis, lentelėje STUDENTU\_KALBOS gali būti kelios eilutės su STUDENTO\_ID = 102, nes studentas su ID = 102 gali mokėti kelias kalbas. Bet šioje lentelėje negali būti nei vieno įrašo su STUDENTO\_ID = 109, nes lentelėje STUDENTAI nėra įrašo, kurio ID = 109.