

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta informačních technologií



Databázové systémy

2019/2020

Dokumentace projektu

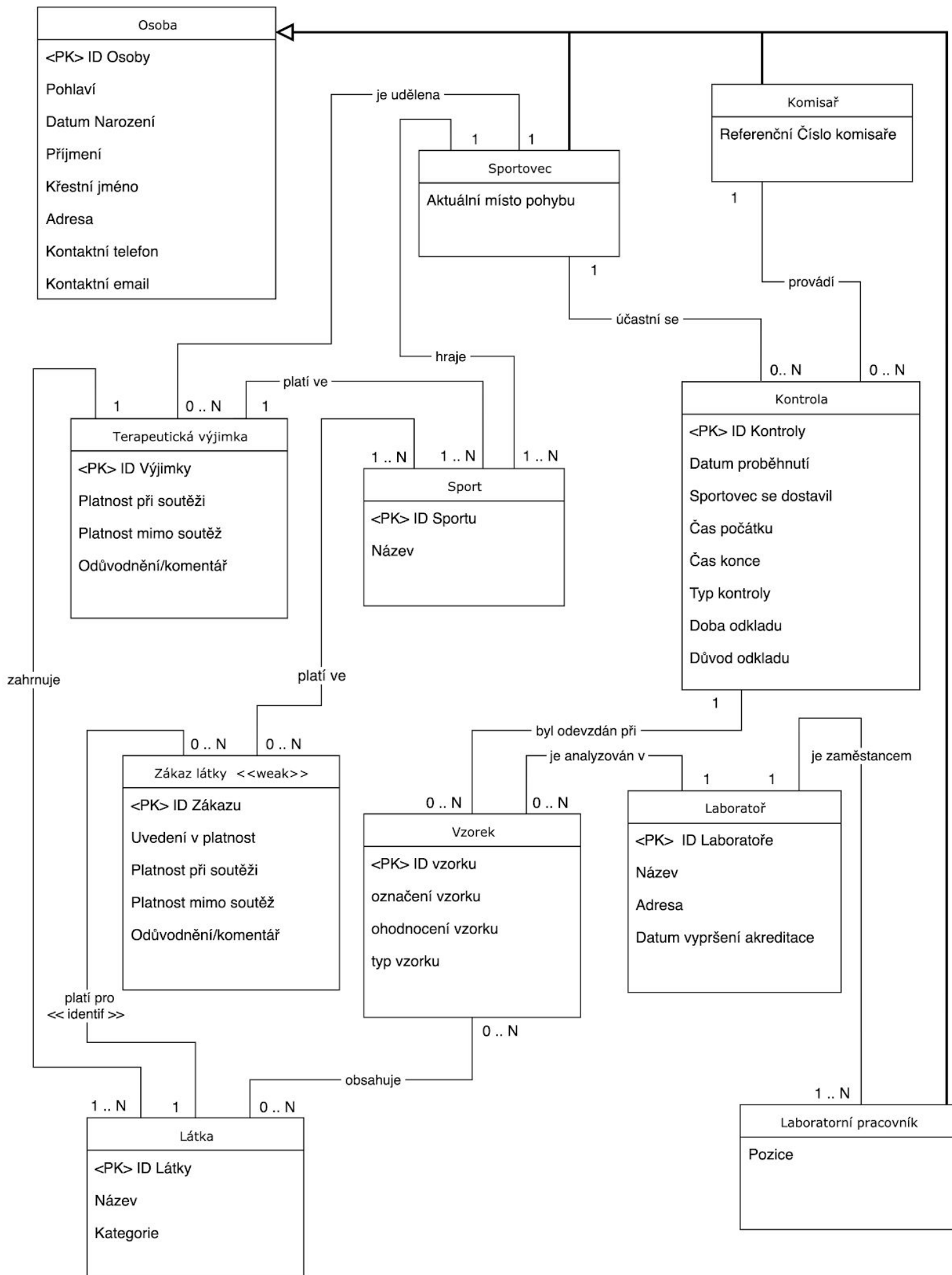
Zadání č. 62 – Antidopingová agentura

Petr Mičulek (xmicul08)
Daniel Jacko (xjacko04)

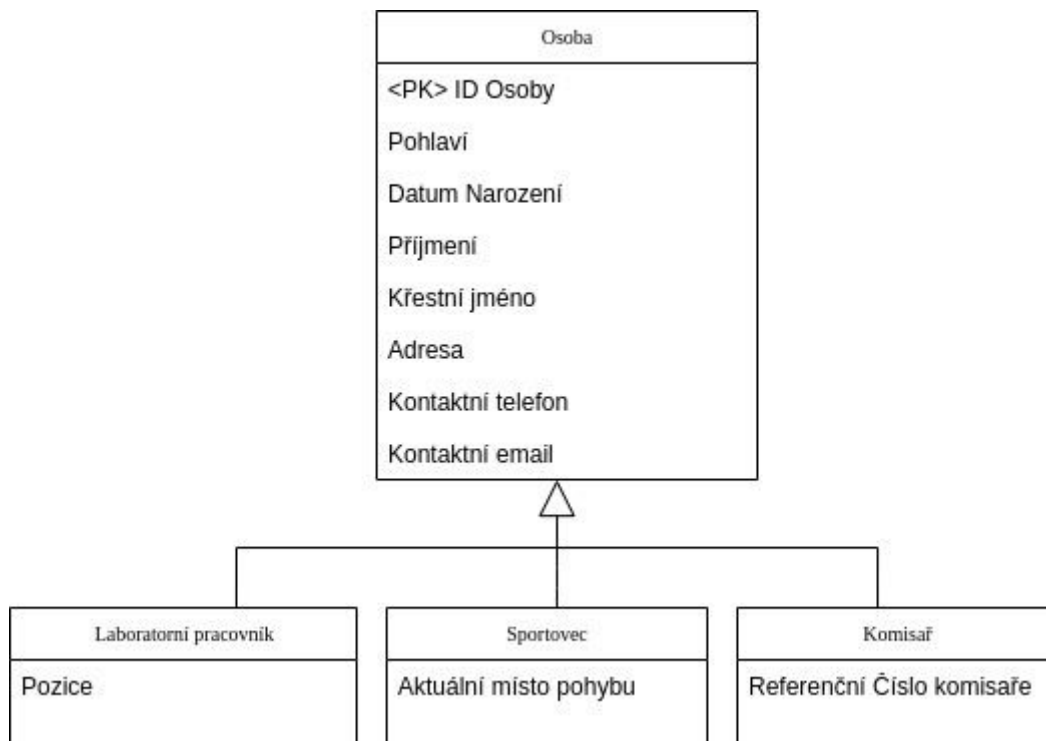
Zadání

Po nedávném rozruchu ve světě sportu se rozhodla antidopingová agentura změnit svůj interní informační systém a vás na to najala. Pro agenturu pracují dopingoví komisaři, kteří budou mít možnost zadávat neohlášené kontroly sportovců na vybrané látky v určený čas a místo konání. IS těsně před konáním kontroly odešle sportovci upozornění. U sportovce se evidují všechny jeho osobní informace včetně pohlaví a aktuálního místa pobytu a sport, který vykonává (může jich být více). V záznamu z dopingové kontroly se uvádí, jestli byl sportovec zastižen a jestli byla kontrola vykonána při soutěži (uvede se druh sportu) a nebo mimo soutěž. Také se uvádí, jestli byl povolen odklad (při postupu v soutěži, delším tréninku). Sportovec může mít povoleno užívat některé léky po určitou dobu z důvodu zranění. Je povinen uvést dopingovému komisaři látky, které jsou součástí těchto léků. Sportovci se odeberou vždy dva vzorky krve nebo moči (hlavní a kontrolní) označené číselným kódem. Ty se pak zapečetěné posílají do laboratoří s platnou akreditací. U laboratoří si mimo akreditace evidujeme jejich adresu a zaměstnance. Laboratorní pracovník pak doplní, jestli byl vzorek negativní, pozitivní nebo neprůkazný. Ve druhém případě zadají i o které zakázané látky se jednalo. U zakázaných látek se uvádí jejich název, popis, od kterého roku byly zakázány a jsou rozděleny do kategorií. Mimo soutěž se některé zakázané látky z kategorií stimulanty, narkotika, kanabinoidy... nekontrolují. U kontroly se bere v potaz i sport, který vykonává sportovec, tzn. např. automobilový jezdec při soutěži nesmí mít v krvi alkohol. Výsledek kontroly je neprodleně oznámen sportovci a ten následně může požádat o ověření kontrolního vzorku.

Schéma databáze – ER diagram



Generalizace/Specializace



Při tvorbě ER diagramu jsme považovali za vhodné využít vztah generalizace u modelování entit, které představují povolání/roli osob. O všech osobách je uchovávána množina základních informací. Každá osoba má zároveň nějakou roli, kterou může být laboratorní pracovník, sportovec nebo komisař. Každá osoba v systému vždy vystupuje ve své jedné konkrétní roli, ačkoli návrh nebrání tomu, aby jedna osoba měla rolí více a mohla být zároveň komisařem i sportovcem.

(Situace, kdy by daná osoba - v roli komisaře - prováděla dopingovou kontrolu sama se sebou - v roli sportovce - jsme v implementaci neřešili. Považujeme to spíše za problém organizace práce na mezilidské úrovni.)

Implementace 4. úlohy

Triggery

Prvním implementovaným triggerem je doplnění primárního klíče pro řádek v tabulce address, pro případ, že by nebyl zadáný ručně.

Druhý a třetí trigger také slouží k validaci vstupních dat pro tabulku person. Druhý umožňuje zajistit, aby pro osobu vždy existoval alespoň nějaký způsob kontaktování – tedy telefonní číslo nebo email.

Třetí trigger validuje formát zadané emailové adresy pro lepší varování v případě nesprávně zadané adresy.

Procedury

Implementované procedury demonstrují použití PL/SQL k získání bohatších informací z databáze.

```
currently_in_country_percentage(country IN varchar2)
```

Tato procedura slouží k vypsání procent sportovců pobývajících aktuálně v zadané zemi (parametr country). Také demonstruje použití výjimky, kterou odchytí v případě, kdy celkový počet sportovců byl 0 a při výpočtu procent by došlo k dělení nulou.

```
not_at_home_list()
```

Tato procedura slouží k vypsání seznamu sportovců, kteří aktuálně pobývají na jiné než domácí adrese.

Explain Plan

Pro zjištění náročnosti operace select jsme si zvolili jednodušší dotaz, který spojuje informace ze dvou tabulek a jeho výsledkem je seznam dvojic (sport, počet sportovců).

Id	Operation	Name	Rows	Bytes	Cost (%CPU)	Time
0	SELECT STATEMENT		10	650	3 (34)	00:00:01
1	HASH GROUP BY		10	650	3 (34)	00:00:01
2	INDEX FAST FULL SCAN	SYS_C001560275	10	650	2 (0)	00:00:01

Index

Výše zmíněný dotaz není náročný, ale dá se znatelně optimalizovat použitím indexu nad tabulkou reprezentující vztah M:N mezi sportovcem a sportem. Došli jsme k závěru, že “sort group by nosort” je výhodnější, jelikož je využito “pipeliningu”, zatímco u “hash group by” k tomu nedochází. Nedochází sice k využití “index fast full scan”, ale ve výsledku jsme dosáhli nižší ceny.

Id	Operation	Name	Rows	Bytes	Cost (%CPU)	Time
0	SELECT STATEMENT		10	650	1 (0)	00:00:01
1	SORT GROUP BY NOSORT		10	650	1 (0)	00:00:01
2	INDEX FULL SCAN	INDEX_ATHLETE_SPORT	10	650	1 (0)	00:00:01

Materialized views

Byl vytvořen jeden materializovaný pohled pro propojení čísla vzorku a jména laboratoře, ve které je analyzován.

Options použité u tohoto pohledu zajišťují, aby se dal používat pro urychlení často používaného dotazu. Proto je nutné, aby byl automaticky aktualizován v případě změn v původních tabulkách.

Přístupová práva

Skript obsahuje příkazy pro přidělení všech práv kolegovi pro vývoj.

Závěr

K vytváření skriptu jsme použili nástroj DataGrip od společnosti JetBrains, který je pro studenty k dispozici zdarma. Skript jsme testovali na školním serveru Oracle12c. Informace potřebné k úspěšnému vypracování projektu jsme získali z učební látky předmětu IDS, demonstračních cvičení, oficiální dokumentace Oracle Database a neoficiálních zdrojů jako StackOverflow a GitHub.