



MODELE DE BAZE DE DATE: IERARHIC, REȚEA ȘI RELAȚIONAL

CHIȚU RALUCA-OANA
INFORMATICĂ ID, ANUL 2

CUPRINS

CUPRINS	2
1.BAZE DE DATE	3
1.1 CE ESTE O BAZĂ DE DATE?	3
1.2. SCURT ISTORIC AL BAZELOR DE DATE	4
2. MODELE DE BAZE DE DATE	6
2.1. MODELUL IERARHIC	7
2.2 MODELUL REȚEA	9
2.3 MODELUL RELAȚIONAL	11
3. CONCLUZII	14
TERMENI CHEIE	17
REFERINȚE	18
BIBLIOGRAFIE	19

1.BAZE DE DATE

1.1 Ce este o bază de date?

Acum circa 70.000 de ani a început ceea ce se cheamă Revoluția Cognitivă, perioadă care s-a încheiat cu aproximativ 30.000 de ani în urmă. De ce este importantă? Pentru că atunci, conform teoriilor, pe baza unor mutații pur întâmplătoare, modul de gândire și comunicare a *Omului (Homo sapiens)* a dus la nașterea limbajului nostru. Practic, este perioada în care informația începe să se dezvolte.

Informația reprezintă în fapt o veste, un mesaj, o știre etc. despre diverse stări, evenimente, obiecte sau forme de manifestare a realității din mediul înconjurător. Ea se transmite, fie verbal, prin intermediul comunicării de la o persoană la alta, fie prin purtători de informație precum ziare, TV, cărți, calculator, device-uri electronice (tableta, smartphone).

Conceptual, informația reprezintă o reflectare în planul gândirii umane a legăturilor de cauzalitate, privind aspecte din realitatea ce ne înconjoară¹.

Progresul uman, de la fabricarea uneltelor la scrierea de documente (*foi, registre*) a condus la necesitatea dezvoltării calculatoarelor. Rolul acestora a fost și este stocarea de informații. Însă, în momentul în care informațiile sunt gestionate prin intermediul calculatorului, acestea devin date.

Data este în fapt o reprezentare a informației prelucrate. Deosebirea dintre cele două concepte este destul de clară: informația este ceea ce este înțeles de o persoană, iar data este ceea ce se stochează pe un mediu pasiv.

Din punct de vedere organizatoric, într-un sistem de calcul, datele sunt organizate, în memoria externă sub forma:

- **fișierelor** – reprezintă o colecție, un ansamblu de înregistrări fizice;
- **bazelor de date** – o mulțime de date accesibile unui grup de utilizatori.

O bază de date reprezintă o colecție partajată de date, care conține datele propriuzise, relațiile logice dintre acestea, precum și descrierea datelor (*structura datelor*). O bază de date este proiectată pentru a satisface necesitățile informaționale ale unei organizații².

În vederea unei bune gestiuni a bazelor de date, s-a dezvoltat ceea ce se numește SGBD (*Sistemul de gestionare a bazelor de date*) sau, în engleză DBMS (*Database Management System*). Acesta reprezintă de fapt software-ul de creare și manipulare a unei baze de date.

1.2. Scurt istoric al bazelor de date

Înainte de inventarea computerelor, datele erau stocate fizic prin intermediul hârtiei, însă acest mod de stocare presupunea spațiu alocat mare, timp crescut pentru găsirea unei informații și minimizarea posibilității de backup sau restaurare a acestora.

O dată cu apariția calculatoarelor, lumea bazelor de date s-a schimbat rapid și a evoluat într-un ritm accelerat.

Astfel, în anii '60 s-au dezvoltat două modele principale de baze de date: CODASYL (*Conference on Data System Language*) și modelul de tip ierarhic IMS (*Information Management System*). Din această perioadă, data de bază SABRE a companiei IBM a fost comercializată și utilizată în mod frecvent, mai ales pentru compania America Airlines.

Anii '70 au oferit următoare etapă în evoluția modelelor de baze de date. În această perioadă a luat naștere unul dintre cele mai utilizate modele: **cel relațional**. E.F.Codd este cel care prezintă pentru prima dată acest model, susținând că organizarea logică este necesar a fi deconectată de la stocarea fizică a informațiilor. Acesta a devenit principiul de bază pentru sistemele de baze de date.

Un alt moment important al istoriei, îl reprezintă apariția în anii '80 a limbajului standard al bazelor de date: **SQL** (*Structured Query Language*), acesta fiind selectat de

către Institutul National de Standarde American în 1986 și de către Organizația Internațională pentru Standarde în 1987. Acest limbaj este folosit și în prezent.

Tot în anii '80 bazele de date orientate de obiecte și-au început dezvoltarea.

O altă dezvoltare majoră a avut loc în anii '90, mai punctual în anul 1997, când este introdus limbajul XML (*Extensible Markup Language*). Acesta reprezintă un limbaj de marcare care definește un set de reguli pentru codificarea documentelor într-un format citibil atât de om, cât și de către un computer. Obiectivele sale sunt generalitatea, simplitatea și capacitatea de utilizare a Internetului. Tot mai mulți furnizori de sisteme de baze de date dezvoltă în produsele lor XML-ul.

Ultima etapă majoră a istoriei modelelor de baze de date a avut loc în 2009, când NoSQL Database (*baze de date non-relaționale*) re-intră în discuții. Deși ele au fost dezvoltate încă din 1998 de către Carlo Strozzi, cel care organizează evenimentul de relansare a acestora este Johan Oskarsson.

În prezent, modelele bazelor de date sunt într-o continuă dezvoltare, giganții industriei fiind: IBM (*DB2*), Oracle (*MySQL*) și Microsoft (*Access*). Acestora li se adaugă companiile MongoDB Inc. (*MongoDB*) și MariaDB Corporation Ab (*MariaDB*)

2. MODELE DE BAZE DE DATE

Modelul unei baza de date reprezintă structura logică a acesteia. Cu ale cuvinte, este factorul determinant care ne indică modul în care datele sunt stocate, organizate și manipulate.

Definirea unui astfel de model are în vedere trei elemente:

- **structura modelului:** definirea obiectelor (entităților) și a caracteristicilor lor, precum și stabilirea relațiilor între obiecte
- **operatorii** (de citire, memorare, modificare, joncțiune etc.) care acționează asupra structurilor de date
- **regulile (restricțiile)** impuse pentru ca datele să rămână corecte.

Relațiile dintre înregistrări într-un model de date are menirea să reducă redundanța datelor.

De-a lungul timpului s-au dezvoltat diverse tipuri de modele pentru baze de date. Astfel, cele mai întâlnite sunt:

- modelul de date ierarhic
- modelul de date relațional
- modelul de date rețea
- modelul de date orientat pe obiect
- modelul de date dimensional.

De ce este necesar să utilizăm un astfel de model? Pentru că beneficiile modelării bazelor de date ușurează practic munca programatorilor, a analiștilor de date, a arhitecților de baze de date. Pe lângă acestea, putem adăuga pe listă:

- îmbunătățește comunicare dintre programatori și echipele de business intelligence
- reduce erorile de sistem și a dezvoltării baze de date

- crește performanța bazei de date
- ușurează maparea datelor în întreaga organizație.

În continuare, vom vorbi despre cele mai cunoscute modele și anume cel ierarhic, cel de tip rețea și cel de tip relațional.

2.1. Modelul ierarhic

Dezvoltat în jurul anului 1968 de către IBM sub denumirea de *Information Management System (IMS)*, modelul ierarhic reprezintă primul model de date care a stat la baza realizării unor SGBD-uri.

Acest tip de model este o colecție de structuri de tip arbore (*arborescent*) și este format din două elemente:

- **nodul master sau nodul rădăcină (key)** – permite identificare tipului, locației și a ordinii înregistrărilor
- **nodurile subordonate** – restul datelor dintr-o înregistrare.

Fișierele în cadrul acestui model de baze de date sunt organizate într-o manieră de tipul părinte-copil, fiecare nod părinte fiind capabil de a avea mai multe noduri-copii, însă fără ca un nod-copil să poată avea mai multe noduri de tip tată. Acest mod de organizare este specific sistemelor de fișiere. În general, avem un director rădăcină care va conține subdirectoare, respectiv fișiere. În figura 1 avem o reprezentare grafică a unui model de tip ierarhic, A1 reprezentând nodul-rădăcină.

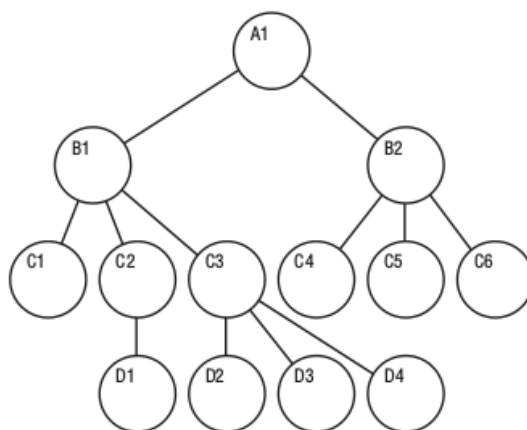


Figura 1. *Hierarchical Database Model*³

Modelul ierarhic, deși a reprezentat o îmbunătățire semnificativă în ceea ce privește gestionarea fișierelor care nu au legătură unele cu celelalte, are, totuși unele dezavantaje care nu pot fi ignorate. Astfel, avem:

a. avantajele modelului ierarhic

- oferă o integritate a datelor
- avantajos pentru un sistem de tip unul-mai multi
- viteză și eficiență pentru date care au o structură arborescentă

b. dezavantajele modelului ierarhic

- operațiunile de inserare sau ștergere pot deveni complexe
- fiecare legătură (asociere) trebuie definită în mod clar în momentul creării bazei de date
- flexibilitate redusă

Un exemplu al unui astfel de model ar putea arăta așa:

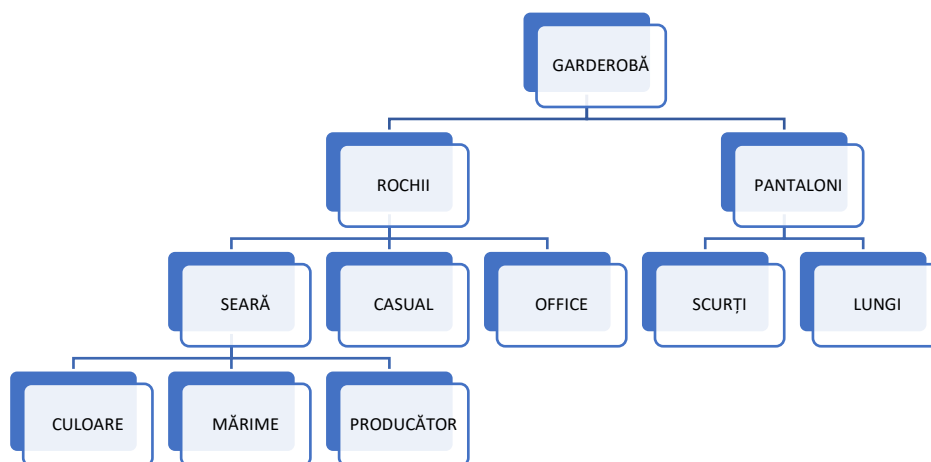


Fig 2. Modelul ierarhic

2.2 Modelul rețea

Apărut în anul 1969, modelul de tip rețea reprezintă o extindere a celui de tip ierarhic, diferența între cele două fiind că în cazul modelului rețea un nod-copil poate avea mai multe noduri-părinte.

Acest tip de model este cel mai apropiat de reprezentarea bazelor de date sub forma diagramelor entitate-asociere. În fapt, putem considera că este un graf orientat în care arcele sunt fie simple (*dacă relația este de tip 1-1*), fie duble (*dacă relația este de tip 1-n*).

Printre caracteristicile modelului rețea se numără complexitatea redusă, crearea de asocieri prin intermediul unei liste de legături și viteza de lucru mare pentru sistemele care implementează astfel de modele.

O reprezentare grafică a acestui model putem vedea în figura 3:

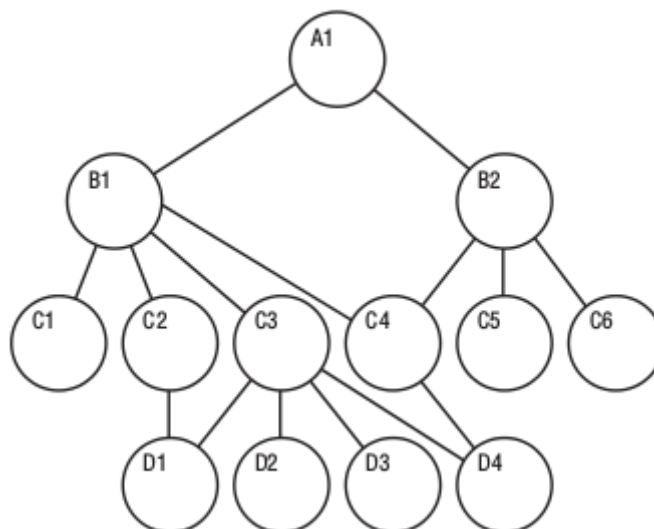


Figura 3. *Hierarchical Database Model*⁴

Avantajele și dezavantajele modelului rețea sunt:

a. avantaje

- este simplu și ușor de implementat
- capabil de a gestiona mai multe tipuri de relații ($1:1$, $1:M$, $M:N$)
- nu permite ca un membru să existe fără ca acesta să aibă un membru superior (*posesor*)

b. dezavantaje

- schema unei astfel de structuri este foarte complexă
- design-ul nu este ușor de utilizat
- nu oferă independență structurală

Exemplificarea unui astfel de model o regăsim în figura următoare:

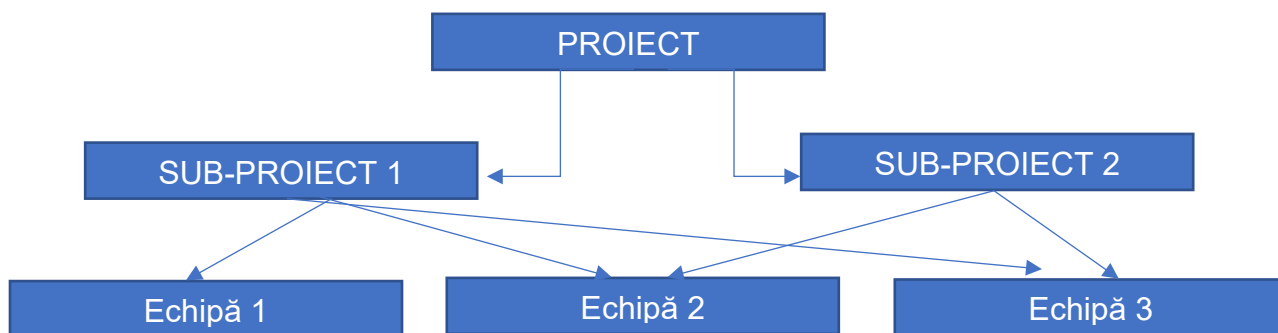


Fig 4. Modelul rețea

2.3 Modelul relațional

Propus de cercetătorul american **E.F.Codd** în anul 1970, modelul relațional este, în prezent, cel mai utilizat model de baze de date.

Reprezintă un model matematic care descrie datele ca o colecție de relații, iar valorile acestor date sunt definite prin așa-numitele domenii. Datele sunt manipulate prin intermediul algebrei relaționale și a calculului relațional.

Acest tip de model este implementat de majoritatea Sistemelor de gestiune a bazelor de date (**SGBD**), fiind numite baze de date relaționale.

O bază relațională de date se prezintă sub forma unuia sau mai multor tabele între care se stabilesc o serie de relații. O astfel de tabelă este, în fapt, un tabel clasic format din coloane (*câmpuri*) și linii (*înregistrări*).

Modelul relațional are terminologia sa unică, astfel că acesta utilizează următoarele concepte :

- **domeniul (tipul datei)** : definește setul de valori posibile pe care o dată le poate folosi

- **relație** : se prezintă sub forma unui tabel bidimensional, fiind o colecție de *tupluri* (succesiune de n elemente). Fiecărui tuplu conține valori ale unui număr fix de *atribute*.
- **atribut** : compus dintr-un nume și un domeniu, reprezintă o coloană a tabelului.
- **tuplu** : este setul de valori de atribuit, fiind o linie din tabel.

Specific tabelor utilizate în modelul de tip relaționale sunt *cheile primare* și *cheile străine*. Cheile primare identifică în mod unic fiecare tuplu al relației. Deși recomandat, nu este o cerință obligatorie ca un tabel să conțină o cheie primară. Pe de altă parte, cheile străine definesc relația dintre două tabele, valorile lor fiind identice cu cele ale cheilor primare din alte tabele față de cel în care sunt definite.

Un alt specific întâlnit în acest tip de model este **normalizarea**. Ce este? O colecție de scheme relaționale care se supun unui set de reguli. Aceasta elimină anomaliile și redundanța, pierderea informațiilor fiind eliminată. Ca și exemplu, gândiți-vă că aveți două tabele: *Angajați* și *Activități*. Dacă modificăm o intrare în tabela *Angajați*, este necesară modificarea datelor și în cea de-a doua. Însă, prin normalizare acest pas se poate realiza automat prin intermediul cheilor : *angajatID* – *activitateID*.

În figura 5 putem observa vizual cum arată o bază de dată folosind modelul relațional:

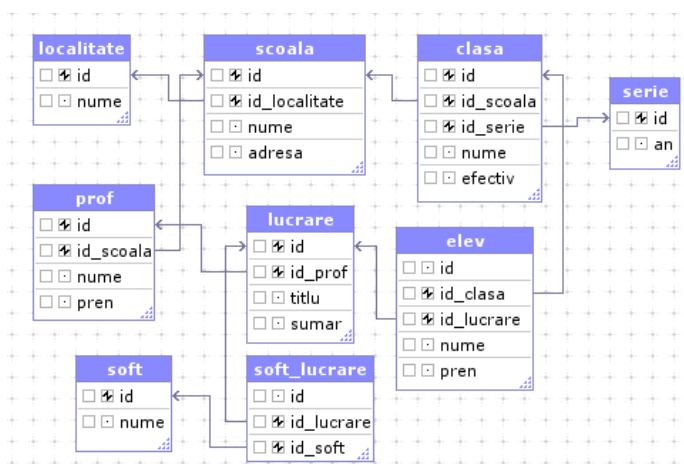


Fig 5. Modelul relațional

Referindu-ne la avantajele și dezavantajele modelului relațional, pe scurt, acestea sunt :

a. avantaje :

- se focusează pe oferirea unei independențe crescute a datelor
- oferă vizualizarea logică a datelor
- opțiunile de import-export oferă facilitarea operațiunilor de backup și restaurare a datelor

b. dezavantaje:

- necesită o stocare fizică mare a datelor
- cost crescut
- timp crescut pentru mentenanță

Spre exemplificare unui astfel de model, avem următorul exemplu :

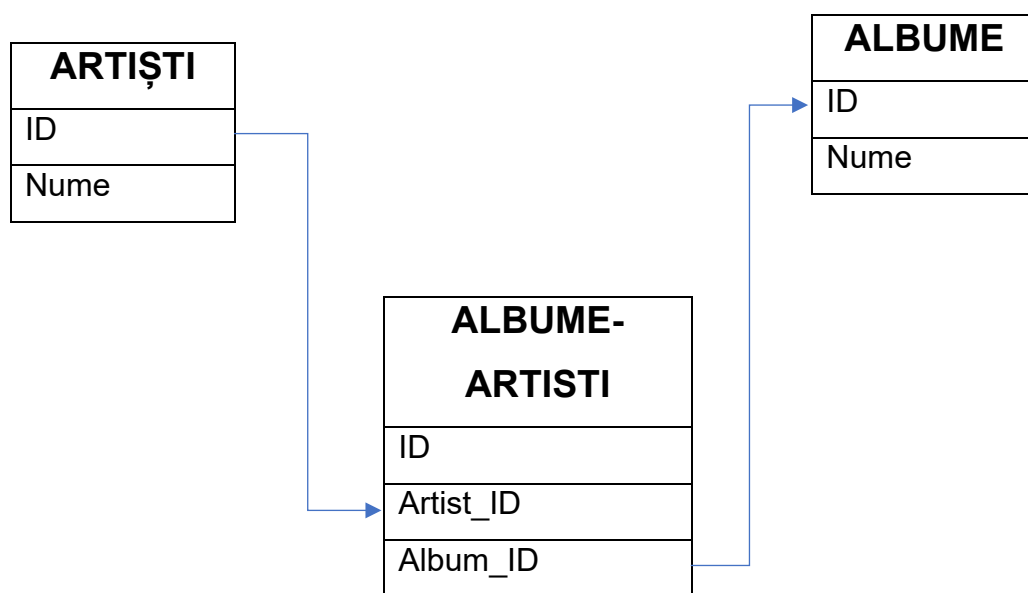


Fig 6. Modelul relațional

3. CONCLUZII

Analizând diferitele tipuri de modelare a bazelor de date, observăm ușor că dezvoltarea acestora s-a datorat necesității de a economisi timpul dedicat înregistrării și evidenței informațiilor.

Privind istoria bazelor de date, concluzionăm că evoluția lor a fost direct proporțională cu evoluția tehnologiei și implicit, a nevoii specialiștilor de a putea manipula datele într-un mod securizat și facil.

În urma prezentării celor 3 tipuri de date (*ierarhic*, *rețea*, *relațional*), putem sublinia următoarele concluzii bazate pe diverse criterii și anume:

a. înregistrare date

- modelul ierarhic: folosește metoda ierarhică, este cel mai vechi model și nu mai este foarte utilizat în prezent
- modelul rețea: înregistrează datele pe bază de legături sau pointeri
- modelul relațional: utilizează tabele și relațiile sunt stabilite pe baza câmpurilor comune

b. organizare date

- modelul ierarhic: utilizează structura de tip arbore
- modelul rețea: folosește grafice direcționate
- modelul relațional: datele sunt organizate sub formă de tabele

c. implementare relații

- modelul ierarhic: relații de tipul 1 : 1 sau 1 : n
- modelul rețea: relații de tipul 1 : 1, 1 : n sau n : m
- modelul relațional: relații de tipul 1 : 1, 1 : n sau n : m

d. structuri de date utilizate

- modelul ierarhic: elementul de bază este pointer-ul
- modelul rețea: baza este lista înlănțuită
- modelul relațional: reprezentare logică

e. inserare elemente noi

- modelul ierarhic: existența anomaliilor datorită faptului că accesul se face prin nodul părinte
- modelul rețea : nu prezintă anomalii
- modelul relațional : nu prezintă anomalii

f. ștergere elemente

- modelul ierarhic: existența anomaliilor datorită faptului că este dificil să ștergi nodul-părinte
- modelul rețea : nu prezintă anomalii
- modelul relațional : nu prezintă anomalii

g. modificare date

- modelul ierarhic: nu este facilă
- modelul rețea : nu există probleme
- modelul relațional : datorită normalizării se elimină redundanța, modificare devenind astfel un proces facil

h. independența datelor

- modelul ierarhic: nu oferă independența datelor
- modelul rețea : oferă independență parțială a datelor
- modelul relațional : oferă independență totală a datelor

i. interogare baze de date

- modelul ierarhic: nu suportă
- modelul rețea : nu suportă
- modelul relațional : suportă prin intermediul limbajului SQL

j. access date

- modelul ierarhic: folosit pentru accesare date complexe și asimetrice
- modelul rețea: folosit pentru accesare date complexe și simetrice
- modelul relațional : folosit pentru accesare date complexe și simetrice

k. design-ul bazei de date

- modelul ierarhic: greu de creat
- modelul rețea : greu de creat și manipulat
- modelul relațional : cel mai ușor model de creat din cele trei

l. flexibilitate

- modelul ierarhic: puțin flexibil
- modelul rețea : mai flexibil decât modelul ierarhic
- modelul relațional : mai flexibil decât modelul ierarhic

TERMENI CHEIE

- **Informația** - veste, mesaj, știre etc. despre diverse stări, evenimente, obiecte sau forme de manifestare a realității din mediul înconjurător.
- **Data** - reprezentare a informației prelucrate
- **Fișiere** – o colecție, un ansamblu de înregistrări fizice
- **Bază de date** – o mulțime de date accesibile unui grup de utilizator
- **SGBD** - software-ul de creare și manipularea unei baze de date.
- **Noduri master** – permite identificare tipului, locației și a ordinii înregistrărilor
- **Noduri subordonate** – restul datelor dintr-o înregistrare.
- **Domeniu** - definește setul de valori posibile pe care o dată le poate folosi
- **Relație** - tabel bidimensional, fiind o succesiune de n elemente
- **Atribut** - compus dintr-un nume și un domeniu, reprezintă o coloană a tabelului.
- **Tuplu** - este setul de valori de atribuit, fiind o linie din tabel.
- **Cheie primară** - identifică în mod unic fiecare tuplu al relației
- **Cheie străină** - definește relația dintre două tabele
- **Redundanță** – eliminare datelor duble
- **Normalizare** - colecție de scheme relaționale care se supun unui set de reguli
- **SQL** - Structured Query Language
- **XML** - Extensible Markup Language
- **NoSQL Database** - baze de date non-relaționale

REFERINȚE

- ¹ Nechita Elena, Baze de date: note de curs, Ed. Alma Mater Bacău, 2012, pg.8
- ² Ovidiu Marcu, Cristina Zotic, Giorgie Daniel Vlad , Informatică manual pentru clasa a XII-a, Ed. Gil, 2007, pg. 6
- ³ MariaDB official site - <https://mariadb.com/kb/en/understanding-the-hierarchical-database-model/>
- ⁴ MariaDB official site - <https://mariadb.com/kb/en/understanding-the-network-database-model/>
- ⁵ Site web personal prezentat ca și referat - <https://docerp.ro/tabelul-unic-versus-baza-de-date-intre-practica-si-teorie/> (*utilizat doar pentru imagine*)

BIBLIOGRAFIE

1. Gyrodi Cornelia, Baze de date relaționale: teorie și aplicații, Ed. Treira, 2000
2. Borza Sorin, Brindașu Paul Dan, Marinela Ință, Baze de date în sisteme informatice de gestiune, Ed. universității Lucian Blaga, 2003
3. Nechita Elena, Baze de date: note de curs, Ed. Alma Mater Bacău, 2012
4. Ovidiu Marcu, Cristina Zotic, Giorgie Daniel Vlad , Informatică manual pentru clasa a XII-a, Ed. Gil, 2007
5. Berg L. Kristi, Seymour Tom, Goel Richa, History of Databases - https://www.researchgate.net/publication/298332910_History_Of_Databases
6. Quickbase.com site oficial (soft de vizualizare date) - <https://www.quickbase.com/articles/timeline-of-database-history>
7. Lucidchart software website - <https://www.lucidchart.com/pages/product>
8. MariaDB official site - <https://mariadb.com/kb/en/understanding-the-network-database-model/>
9. Upwork site oficial (platformă de lansare bussiness) - <https://www.upwork.com/resources/nosql-vs-sql>
10. IBM site oficial - <https://www.ibm.com/cloud/learn/relational-databases>
<https://docerp.ro/tabelul-unic-versus-baza-de-date-intre-practica-si-teorie/>