

1

# Baze de date relaționale

## Introducere în SQL

Ce este o Bază de date?



*colecție voluminoasă de elemente similare*



*stocate în vederea analizei*

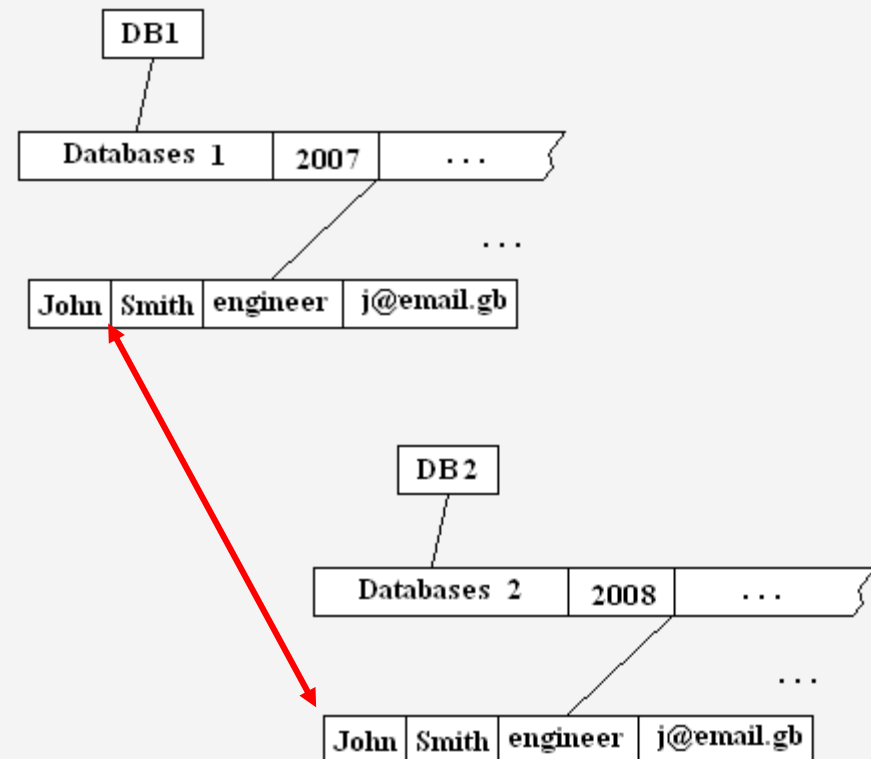
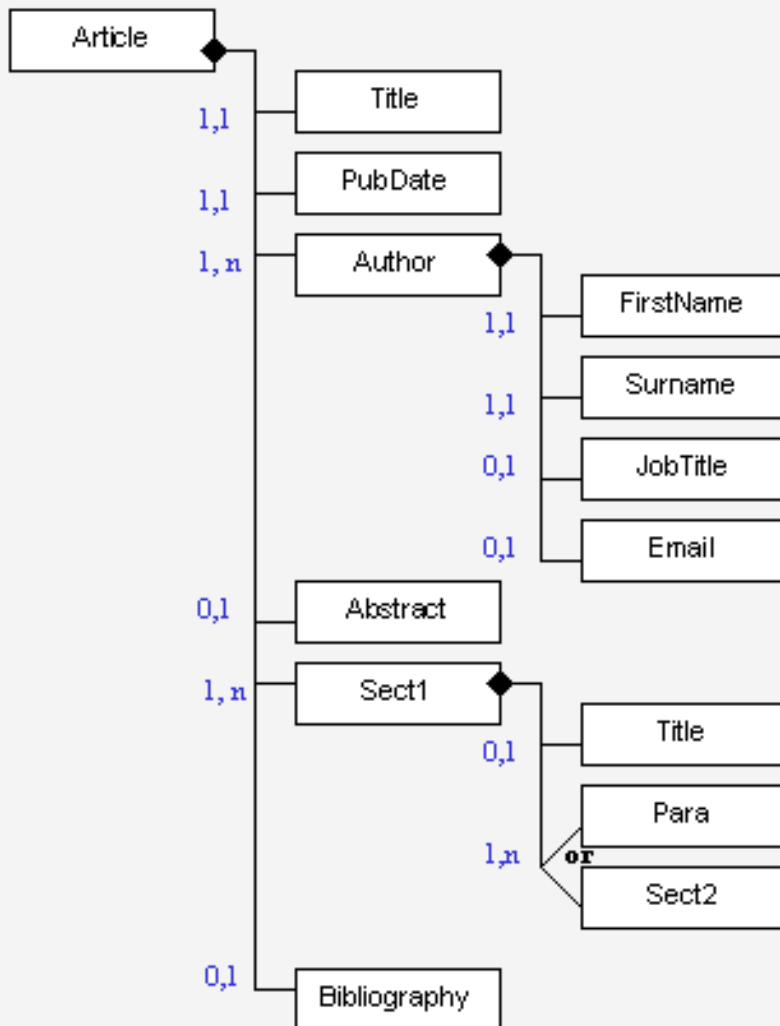


*care persistă o perioadă lungă de timp*

Modelează aspecte ale lumii reale  
prin intermediul unui *model de date*

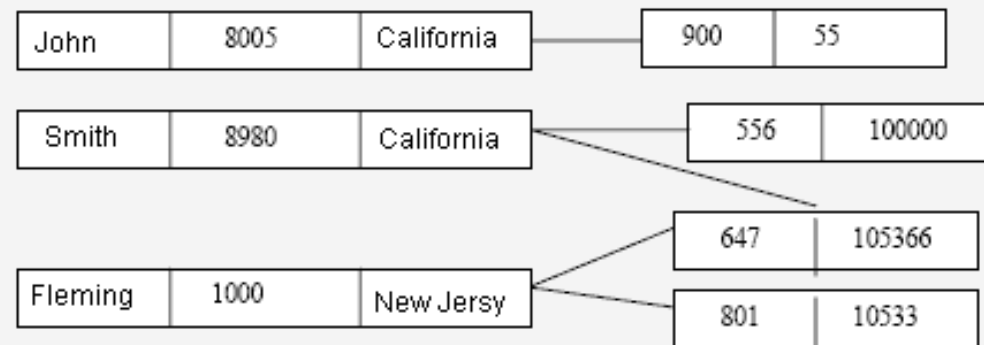
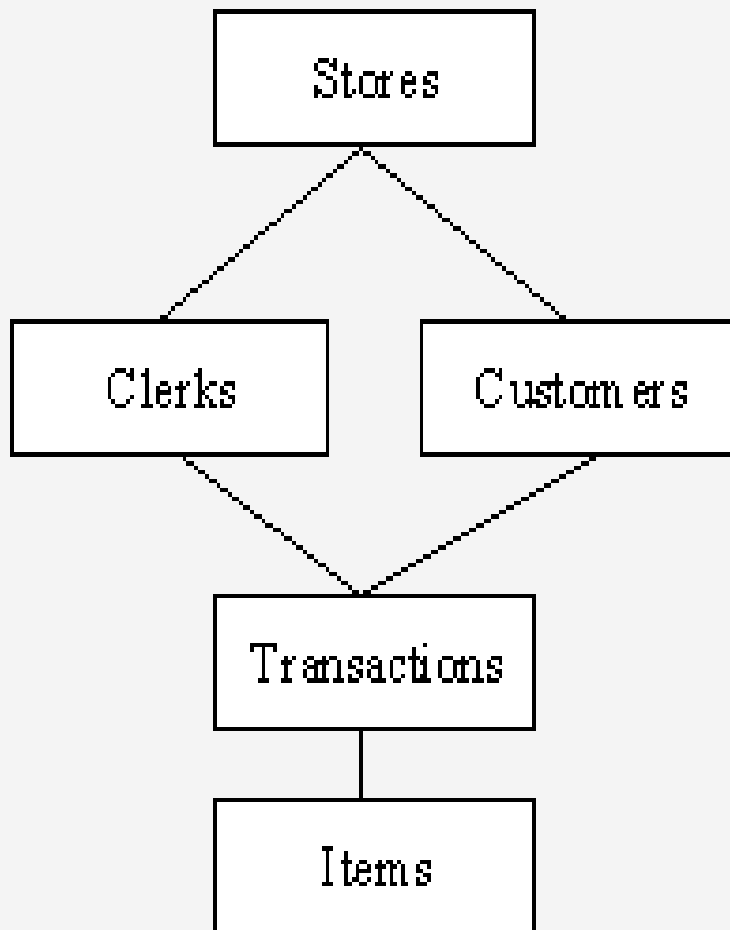
# Modelul ierarhic

- Concepte:
  - structură arborescentă,
  - tip entitate,
  - relații 1-n,...



# Modelul rețea

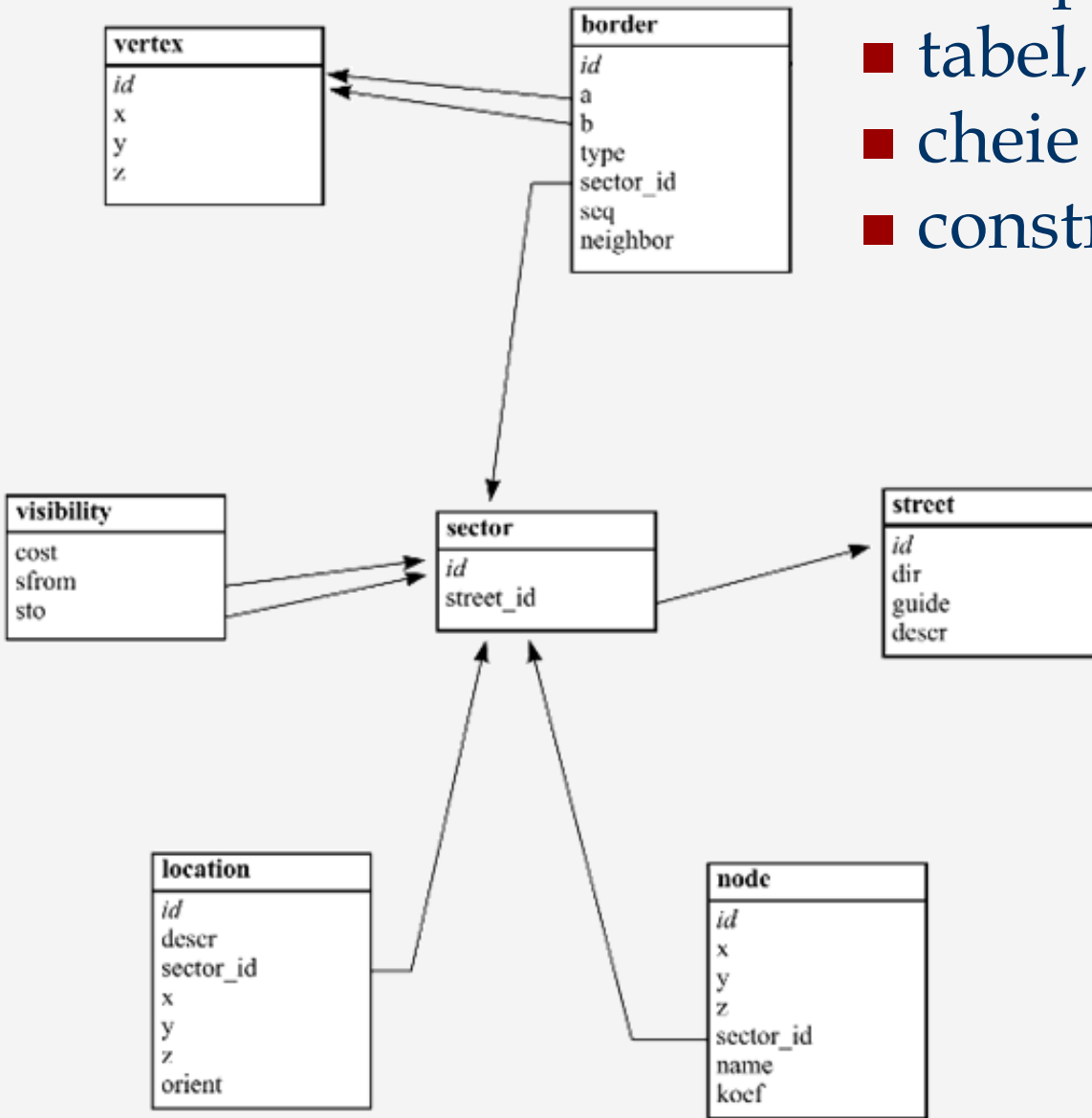
- Concepte:
  - structură de graf,
  - tip articol,
  - tip relație,...



# Modelul relațional

## ■ Concepte:

- tabel, câmp
- cheie primară, cheie străină
- constrângeri de integritate,...

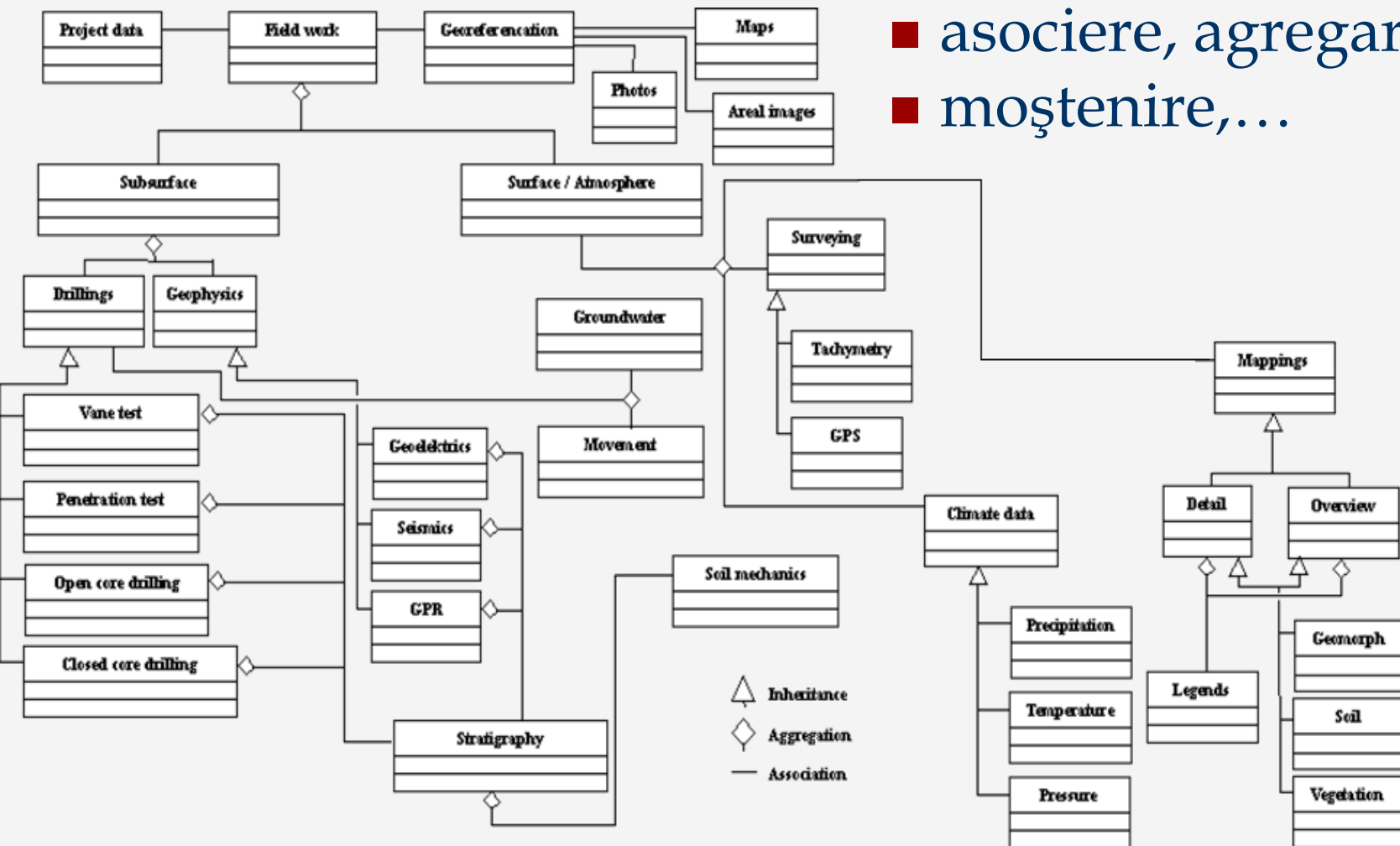


vertex

id	x	y	z
1	30	30	30
2	30	30	0
3	30	0	0
4	0	0	0
5	0	0	30
6	0	30	30

# Modelul orientat obiect

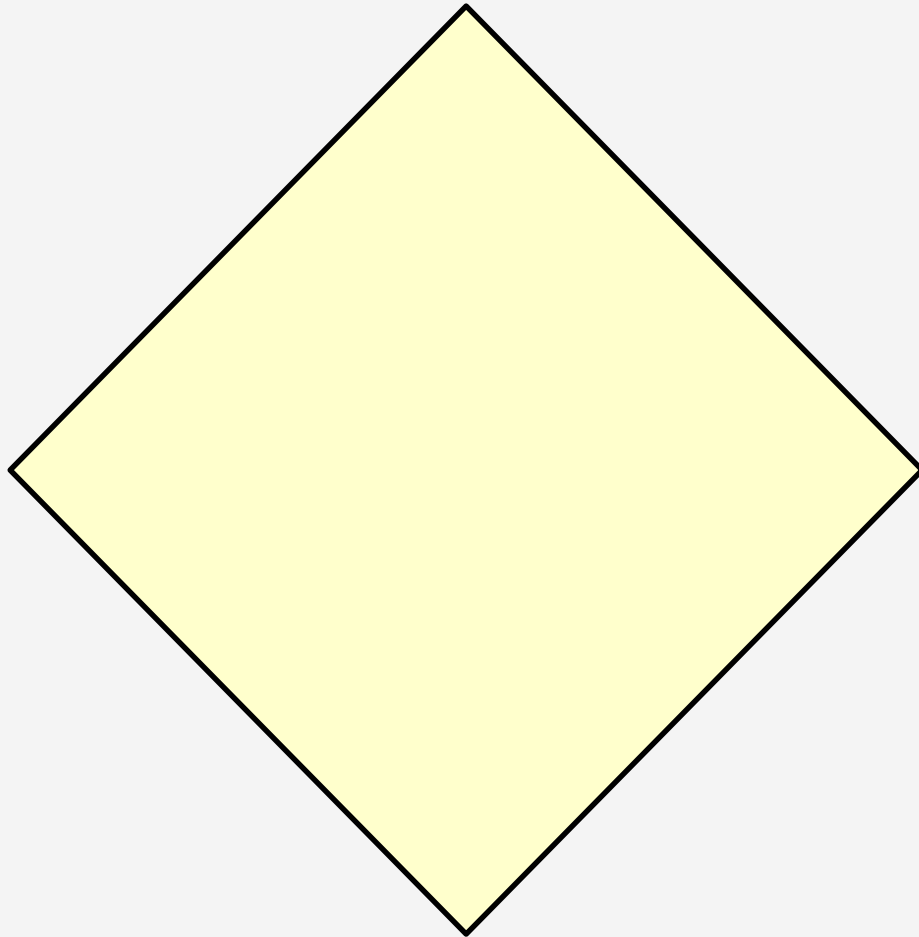
- Concepte:
  - clase, atribut, metode
  - asociere, agregare
  - moștenire,...



# Schemă

VS.

# Date



# Instanță

# Ce este un *SGBD*?

- Colecție integrată de **instrumente** pentru
  - ... crearea unei baze de date și specificare structurii acesteia;
  - ... interogarea și modificarea eficientă a datelor;
  - ... securizarea datelor;
  - ... controlul accesului la date de către *mai mulți* utilizatori la *un moment dat*;
  - ...



# SGBD-uri generale



# SGBD NoSQL



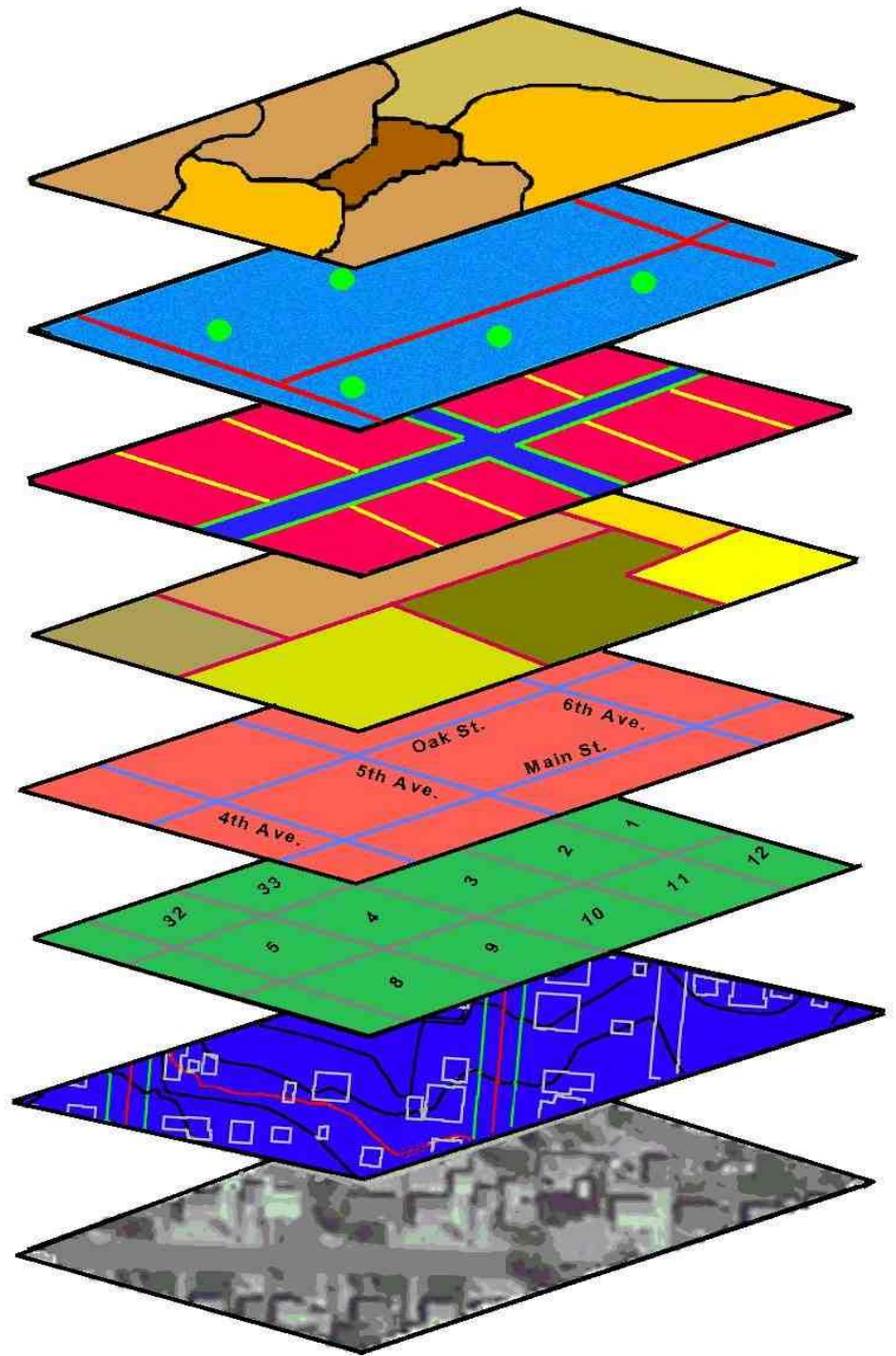




# SGBD Multimedia



# Geographical Information Systems (GIS)





SGBD

Real-Time

# Data Warehouse





SGBD-uri active



# Data Stream Management System





# DBMSs

are

getting

smaller

and

smaller

Databases  
are  
getting  
bigger  
and  
bigger

**Când  
utilizăm  
baze de date?**

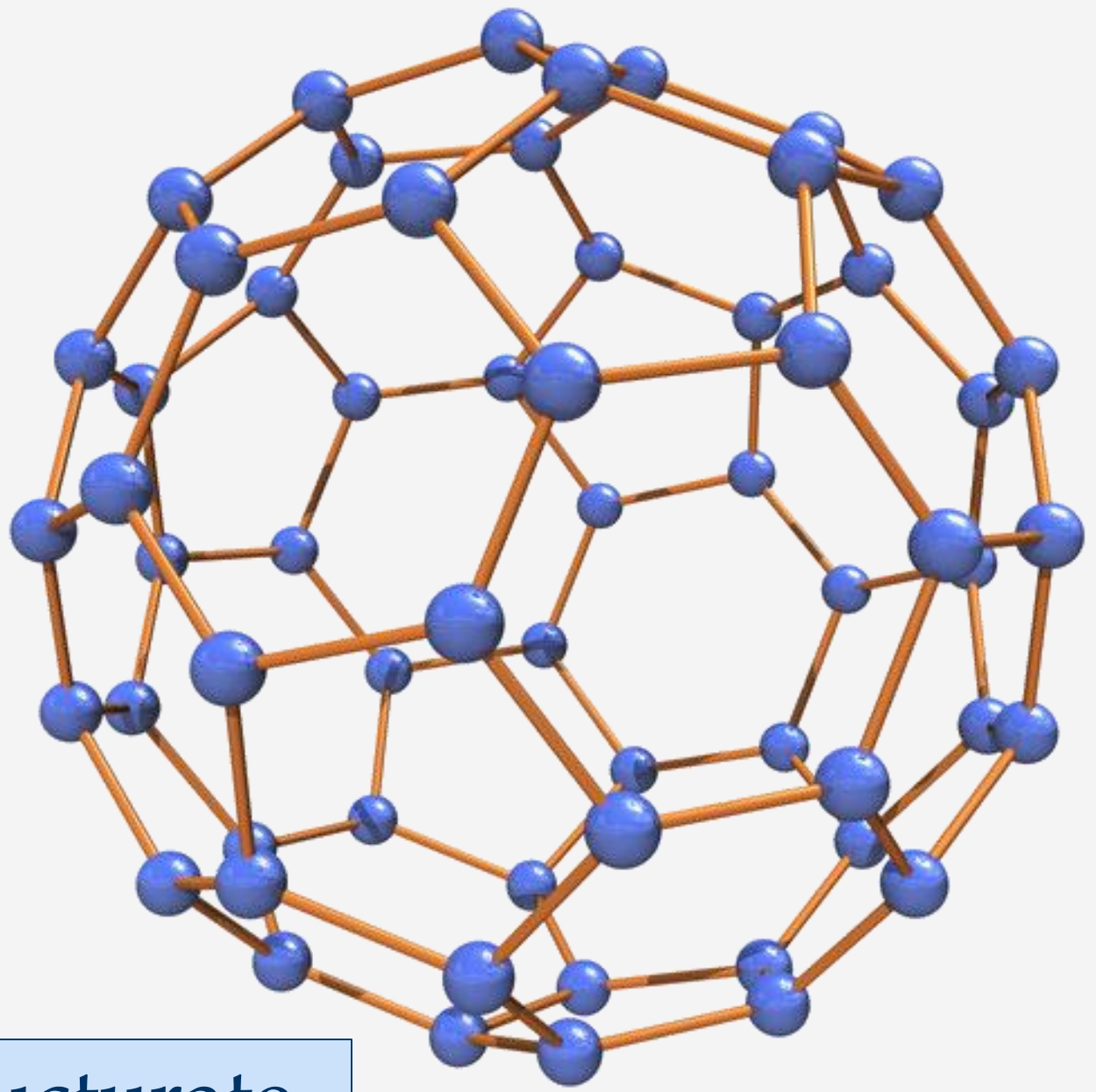




# 1. Persistență

## 2. Cantitate mare de date





### 3. Date structurate



## 4. Integritate

T R U S T

The word 'TRUST' is spelled out using five white, square tiles with rounded corners. Each tile has a single letter in a dark, serif font. The tiles are arranged in a horizontal row on a dark, heavily textured, and cracked surface, possibly stone or old wood. The lighting is dramatic, coming from the upper left, which casts long, dark shadows to the right of each tile, emphasizing their three-dimensional form and the rough texture of the background.





## 5. Accés Concurrent/Distribuit





## 6. Securitate



## 7. Transmiterea datelor altor aplicații



Când  
**NU**  
utilizăm  
baze de date?



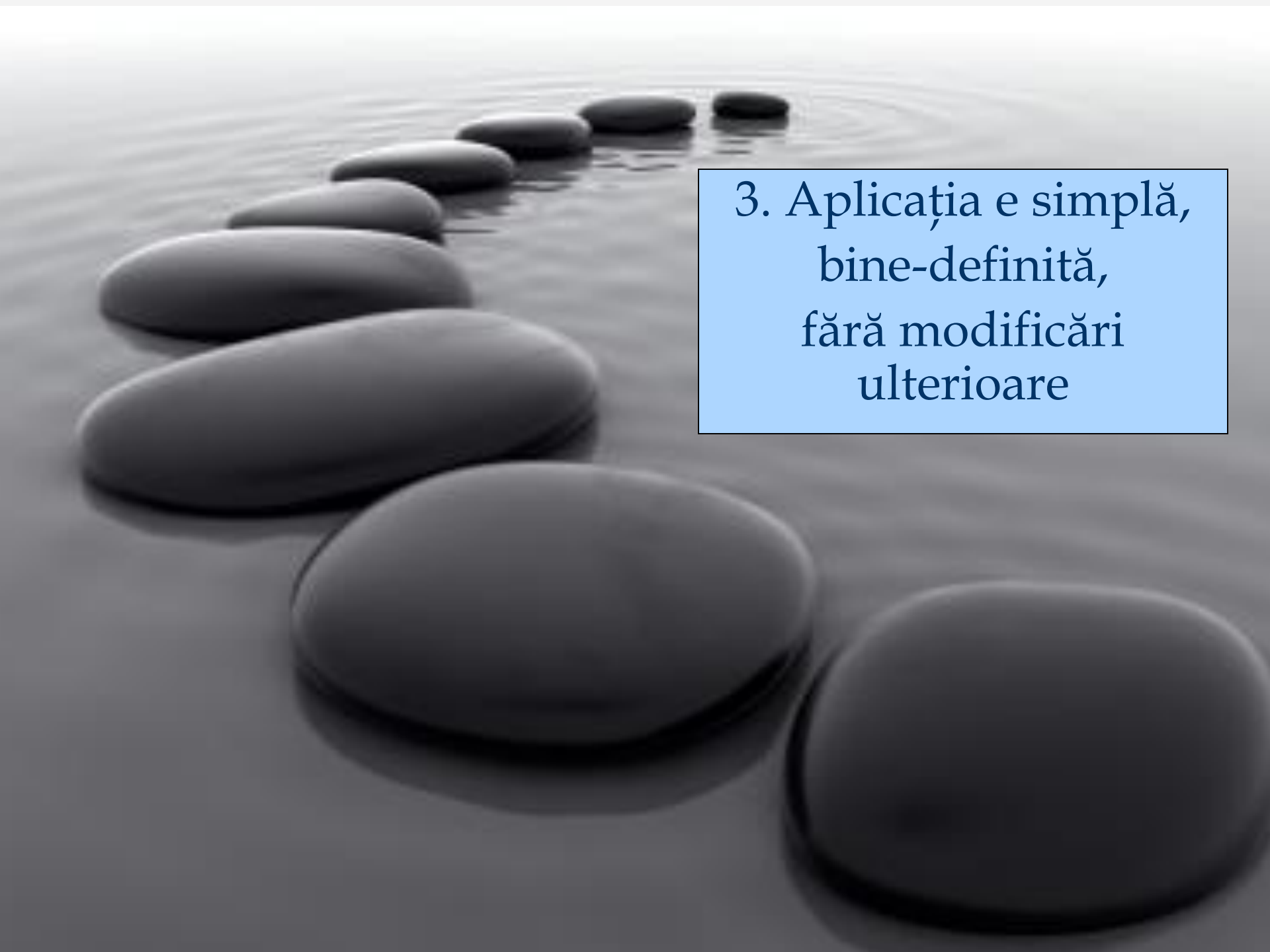
# 1. Investiția inițială e prea mare







## 2. Prea mult efort



3. Aplicația e simplă,  
bine-definită,  
fără modificări  
ulterioare





4. Nu este necesar accesul mai multor utilizatori la date



__NounName	Name	Handles	VM	WS	PM	NPM	Path	Company	CPU	FileVersion	ProductVer
"Process"	"AcDeskBandHlpr"	"127"	"85278720"	"10047488"	"4362240"	"14224"	"C:\Program Files (x86)\Le				
"Process"	"AcPrfMgrSvc"	"167"	"72732672"	"11051008"	"4956160"	"17128"	"C:\Program Files (x86)\Le				
"Process"	"AcSvc"	"421"	"127533056"	"17551360"	"10117120"	"26960"					
"Process"	"armsvc"	"84"	"45129728"	"4509696"	"1601536"	"8528"					
"Process"	"btwdins"	"127"	"60628928"	"7720960"	"3743744"	"1009"					
"Process"	"CamMute"	"111"	"40837248"	"5125328"	"839328"	"1009"					
"Process"	"CcmExec"	"1433"	"32935616"	"39616192"	"584112"	"8432"					
"Process"	"conhost"	"35"	"25186004"	"4130960"	"2248000"	"112"					
"Process"	"conhost"	"45"	"52461568"	"6311936"	"3215360"	"6360"	"C:\windows\system32\conhost.exe"	"M			
"Process"	"csrss"	"1145"	"55078912"	"5623808"	"2977792"	"17768"					
"Process"	"csrss"	"628"	"96055296"	"17383424"	"20799488"	"28704"					
"Process"	"CxAudMsg64"	"98"	"57937920"	"6680576"	"7385088"	"8832"					
"Process"	"daemonu"	"413"	"73506816"	"8957952"	"5935104"	"20132"					
"Process"	"DcaSvc"	"642"	"64344064"	"13996032"	"10178560"	"27904"					
"Process"	"DcaTray"	"644"	"220614656"	"35368960"	"33832960"	"34628"	"C:\Program Files (x86)\DirectA				
"Process"	"dwm"	"108"	"84504576"	"10313728"	"5709824"	"10024"	"C:\windows\system32\Dwm.exe"	"Microso			
"Process"	"EvtEng"	"284"	"105017344"	"20992000"	"12906496"	"22248"					
"Process"	"EXCEL"	"384"	"316968960"	"4223386"	"27746304"	"48584"	"C:\Program Files (x86)\Microsoft				
"Process"	"explorer"	"1153"	"35283344"	"5155984"	"1250640"	"4440"	"C:\windows\Explorer.EXE"	"Mic			
"Process"	"fmapp"	"30"	"53719040"	"6168756"	"51744"	"690"	"C:\Program Files\CONEXANT\ForteConfig				
"Process"	"FwcAgent"	"555"	"59432960"	"11211872"	"9289728"	"42248"					
"Process"	"hkcmd"	"84"	"71577600"	"804048"	"3964928"	"8048"	"C:\windows\system32\hkcmd.exe"	"Intel			
"Process"	"ibmpmsvc"	"61"	"44564480"	"4259840"	"2301952"	"6568"					
"Process"	"Idle"	"0"	"24576"	"0"	"0"	"0"					
"Process"	"iexplore"	"774"	"304881664"	"79278080"	"77529088"	"82152"	"C:\Program Files (x86)\Intern				
"Process"	"iexplore"	"639"	"157421568"	"22876160"	"12095488"	"38648"	"C:\Program Files (x86)\Intern				
"Process"	"igfxpers"	"203"	"83292160"	"10944512"	"5844992"	"10808"	"C:\windows\system32\igfxpers.ex				
"Process"	"LMS"	"112"	"41308160"	"5484544"	"2117632"	"9800"					
"Process"	"lsass"	"1545"	"75075584"	"27377664"	"18477056"	"47272"					
"Process"	"lsm"	"213"	"19779584"	"5664768"	"3756032"	"7600"					
"Process"	"micmute"	"110"	"48328704"	"5672960"	"6172672"	"11120"					
"Process"	"mscorsvw"	"105"	"73936896"	"14417920"	"8912896"	"11184"					
"Process"	"MsitBlSHA"	"118"	"45617152"	"8912896"	"4481024"	"10328"					
"Process"	"msitcertsvc"	"712"	"565702656"	"33861632"	"40435712"	"41284"					
"Process"	"MsitTpmsvc"	"84"	"43319296"	"10129408"	"6975488"	"7568"					
"Process"	"MsMpEng"	"503"	"238940160"	"87953408"	"117288960"	"47112"					
"Process"	"MSOIDSVC"	"612"	"106561536"	"21938176"	"14983168"	"27648"					
"Process"	"MSOIDSVCM"	"72"	"36134912"	"4841472"	"2621440"	"6080"					
"Process"	"msseces"	"341"	"156336128"	"23216128"	"10706944"	"25232"	"C:\Program Files\Microsoft Sec				
"Process"	"NisSrv"	"261"	"80003072"	"4771840"	"9138176"	"18000"					
"Process"	"nusb3mon"	"89"	"75804672"	"5914624"	"2224128"	"10328"	"C:\Program Files (x86)\Renesas El				

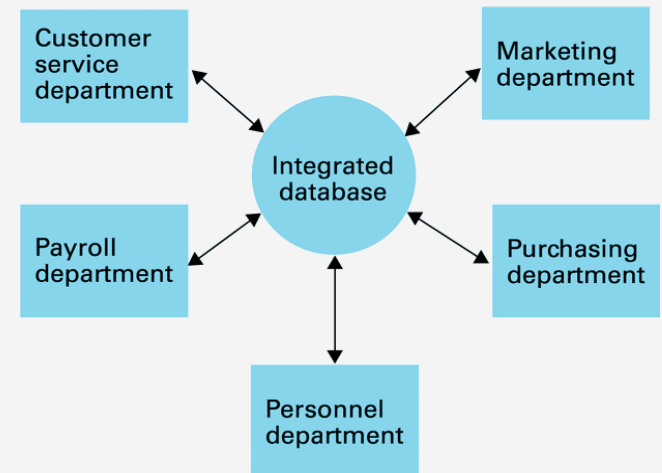
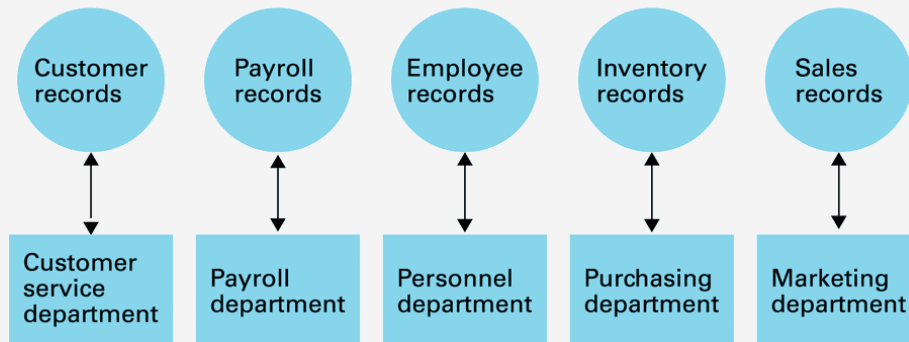


# Fișiere text **vs.** Bază de date

- a. Fișiere text: sistem de stocare unidimensională
- b. Bază de date: stocare multidimensională

b. Database-oriented information system

a. File-oriented information system



# Provocările utilizării fișierelor text

- prezintă mai multe formate de memorare a datelor; codul scris într-o manieră ad-hoc pentru o aplicație nu poate fi exploatat pentru aplicații pe alte fișiere
- există redundanță în memorarea datelor, unele dintre acestea se regăsesc în mai multe fișiere; acest fapt poate duce la inconsistența datelor
- operațiile de citire / scriere sunt descrise în program; se ia în considerare o anumită structură a înregistrărilor, fapt ce conduce la greutate în dezvoltarea unui program (prin schimbarea structurii fișierelor trebuie modificat programul)
- este dificilă obținerea datelor care îndeplinesc anumite condiții
- actualizarea datelor este complexă (e.g., modificarea unor valori din înregistrări, ștergerea unor înregistrări)
- verificarea anumitor condiții de integritate (corectitudine) se face din program
- trebuie gestionată memoria internă (e.g., cum se încarcă o colecție de date de zeci sau sute de GB în memorie pentru procesare?)

# Provocările utilizării fișierelor text (cont)

- nu există proceduri de securitate adecvate, i.e., e nevoie de politici de securitate în care diferiților utilizatori li se acordă sau nu permisiunea de a accesa anumite porțiuni din date
- nu se poate controla ușor accesul concurent la date
- datele trebuie readuse la o formă corectă dacă sistemul întâmpină probleme în timp ce se operează modificări asupra lor, e.g., o operațiune bancară care transferă bani din contul A în contul B este întreruptă de o pană de curent după ce a scos bani din contul A, dar înainte de a-i fi depus în B; datele trebuie readuse la forma corectă, banii trebuie puși înapoi în contul A
- fișierele sunt utile pentru programe care necesită puține date și sunt folosite de un singur utilizator, însă astăzi crește atât cantitatea de date, cât și numărul utilizatorilor care folosesc o anumită aplicație

# Modelul Relational

# Modele de date

- Modelul ierarhic (1965)
- Modelul rețea (1965)
- **Modelul relațional (1NF) (1970s)**
- Model relațional imbricat (1970s)
- Obiecte complexe (1980s)
- Model obiectual (1980)
- Model relațional-obiectual (1990s)
- XML (DTD), XML Schema (1990s)

# Model relațional - idei

- Utilizează o structură de date simplă: *Tabela*
  - simplu de înțeles
  - utilă în modelarea multor situații/entități din lumea reală
  - conduc la interogări de o complexitate redusă
- Utilizează matematica în descrierea/reprezentarea înregistrărilor și a colecțiilor de înregistrări: *Relația*
  - pot fi modelate formal
  - permit utilizarea de limbaje de interogare formale
  - au proprietăți ce pot fi modelate și demonstrate matematic

# Relația – definiție formală

- O **relație** sau **structura unei relații R** este o listă de **nume de attribute**  $[A_1, A_2, \dots, A_n]$ .
- **Domeniu** = mulțime de valori scalare (tipuri atomice - întreg, text, dată, etc)
- $D_i = \text{Dom}(A_i)$  - domeniul lui  $A_i$ ,  $i=1..n$
- **Instanța unei relații** ( $[R]$ ) e o submulțime a  $D_1 \times D_2 \times \dots \times D_n$

# Relația – definiție formală

- **Grad (aritate)** = numărul tuturor atributelor din structura unei relații
- **Tuplu** = un element al instanței unei relații, o înregistrare. Toate tuplurile unei relații sunt distincte!
- **Cardinalitate** = numărul tupluri unei relații



# Exemplu de relație

- Students(**sid**:integer; **name**:string;**email**:string; **age**:integer; **gr**:integer)

*nume câmp*

*tip câmp  
(domeniu)*

<b>sid</b>	<b>name</b>	<b>email</b>	<b>age</b>	<b>gr</b>
2833	Jones	<a href="mailto:jones@scs.ubbcluj.ro">jones@scs.ubbcluj.ro</a>	19	231
2877	Smith	<a href="mailto:smith@scs.ubbcluj.ro">smith@scs.ubbcluj.ro</a>	20	232
2976	Jones	<a href="mailto:jones@math.ubbcluj.ro">jones@math.ubbcluj.ro</a>	21	233
2765	Mary	<a href="mailto:mary@math.ubbcluj.ro">mary@math.ubbcluj.ro</a>	22	233

*structură  
relație*

*instanță  
relație*

*tuplu*

- cardinalitate = 4, grad = 5, toate tuplurile distincte !

# Baze de date relaționale

- O **bază de date** este o mulțime de relații
- **Structura** unei baze de date este mulțimea structurilor relațiilor acesteia
- **Instanța (starea)** unei baze de date este mulțimea instanțelor relațiilor acesteia

# Reprezentarea grafică a relațiilor






Students(*sid*:string, *name*:string, *email*:string, *age*:integer, *gr*:integer)




Courses(*cid*: string, *cname*: string, *credits*:integer)




Enrolled(*sid*:string, *cid*:string, *grade*:double)



Teachers(*tid*:integer; *name*: string; *sal* : integer)




Teaches(*tid*:integer; *cid*:string)

Students	
	sid
	name
	email
	age
	gr

Courses	
	cid
	cname
	credits

Teachers	
	tid
	name
	sal

Teaches	
	tid
	cid

Enrolled	
	sid
	cid
	grade

# Constrângeri de integritate (CI)

- **CI**: sunt condiții ce trebuie să fie îndeplinite de către *orice* instanță a unei baze de date
  - specificate la momentul definirii structurii relației
  - verificate la modificarea conținutului relației
- O instanță a unei relații că este *legală* dacă satisface toate CI specificate
  - SGBD nu va permite instanțe *ilegale*

# Constrângeri de integritate - exemple

- Students(*sid:string*, *name:string*, *email:string*, *age:integer*, *gr:integer*)
  - Constrângere de domeniu: *gr:integer*
  - Constrângere de interval:  $18 \leq \textit{age} \leq 70$
- TestResults(*sid:string*, *TotalQuestions:integer*, *NotAnswered:integer*, *CorrectAnswers:integer*, *WrongAnswers:integer*)
  - $\textit{TotalQuestions} = \textit{NotAnswered} + \textit{CorrectAnswers} + \textit{WrongAnswers}$  – **nu e o CI!**

# Chei Primare

- O mulțime de attribute reprezintă o **cheie** a unei relații dacă:
  1. Nu există două tuple care au aceleași valori pentru toate attributele

**ȘI**

  2. Acest lucru nu este adevărat pentru nici o submulțime a cheii
- Dacă a 2-a afirmație este falsă → **super cheie**
- Dacă există >1 cheie pentru o relație → **chei candidat**
- Una dintre cheile candidat este selectată ca **cheie primară**



## Chei străine (externe)

- O **cheie străină (externă)** este o mulțime de câmpuri a unei relații utilizate pentru a `referi` un tuplu al unei alte relații (un fel de `pointer logic`).
  - Aceasta trebuie să corespundă cheii primare din a doua relație.

De exemplu pentru

*Enrolled* (*sid*: string, *cid*: string, *grade*: double)

*sid* este cheie externă referind *Students*

# Integritate referențială

■ **Integritate referențială** = nu sunt permise valori pentru cheia străină care nu se regăsesc în tabela referită.

Exemplu de model de date fără integritate referențială:

*Link-uri HTML*



# Integritate referențială

- Fie *Students* și *Enrolled*; *sid* in *Enrolled* este o cheie străină ce referă o înregistrări din *Students*.
- Adaugarea in *Enrolled* a unui tuplu cu un id de student inexistent, acesta va fi respins de SGBD.

*Enrolled*

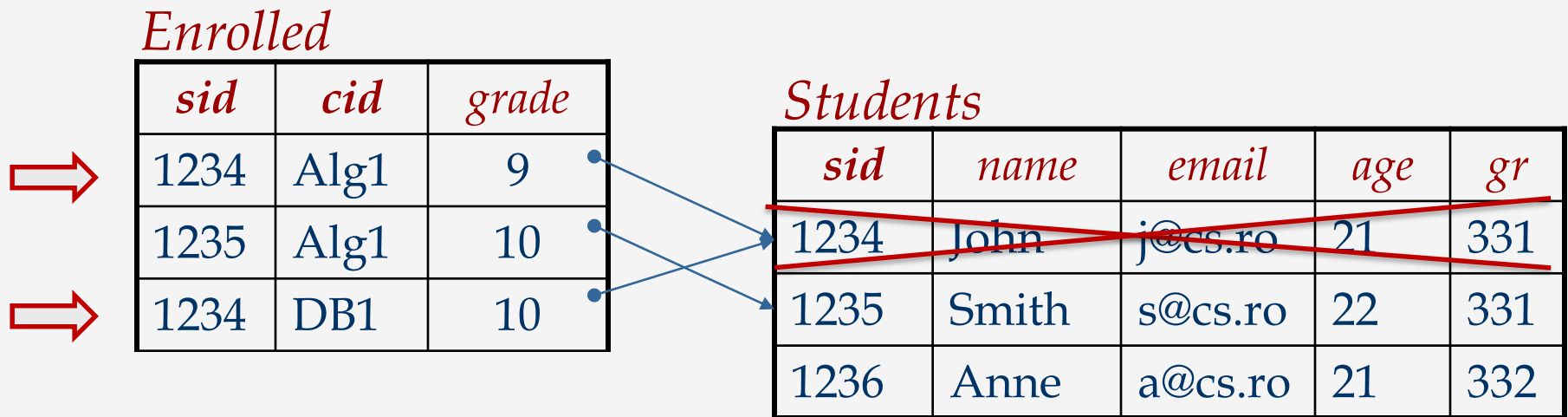
<i>sid</i>	<i>cid</i>	<i>grade</i>
1234	Alg1	9
1235	Alg1	10
1234	DB1	10
<del>1237</del>	<del>DB2</del>	<del>9</del>

*Students*

<i>sid</i>	<i>name</i>	<i>email</i>	<i>age</i>	<i>gr</i>
1234	John	j@cs.ro	21	331
1235	Smith	s@cs.ro	22	331
1236	Anne	a@cs.ro	21	332

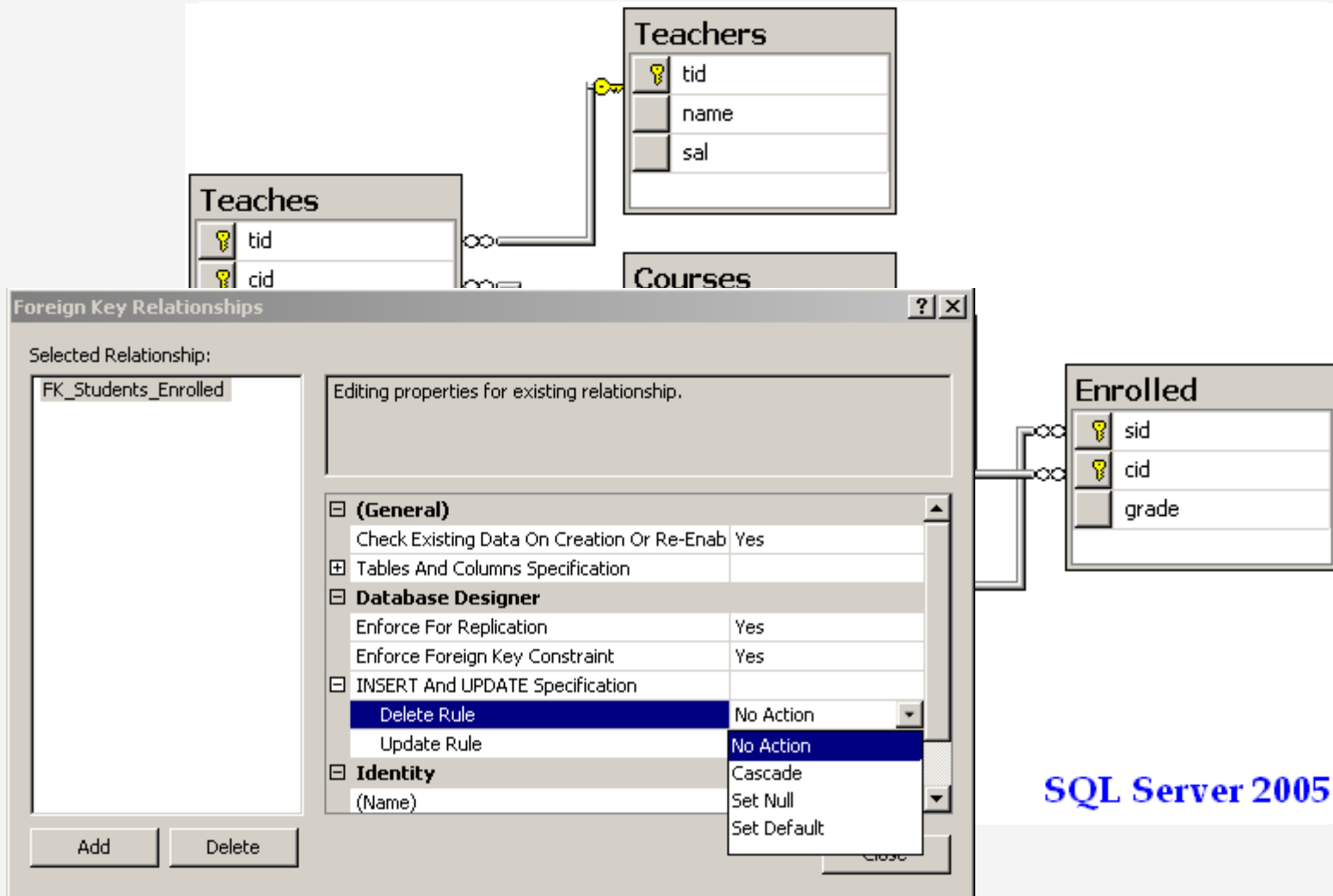
# Integritate referențială

- Dacă o înregistrare din *Students* este ștearsă dar ea este referită din *Enrolled*:
  - se șterg toate înregistrările ce o refera din *Enrolled*.
  - nu se permite ștergerea înregistrării din *Students*
  - sid din *Enrolled* va avea asignată o valoare implicită.
  - sid din *Enrolled* va avea asignată valoarea *null*.





# Reprezentarea grafică a CI

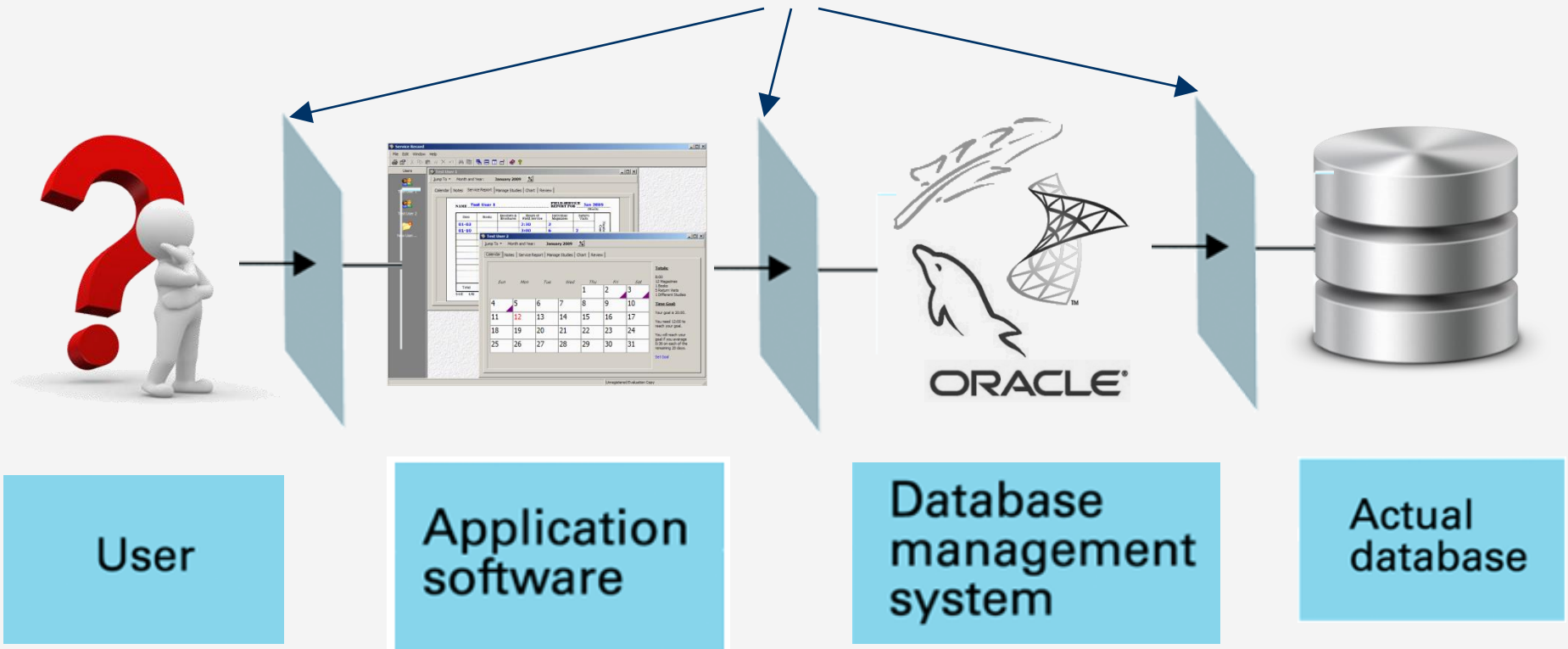


# Cum apar CI?

- CI se bazează pe semantica entităților din lumea reală / conceptuală modelate.
- Putem verifica dacă o CI este încălcată de instanța unei tabele, însă **NU** vom putea deduce dacă o CI este adevărată doar consultând o singură instanță.
  - O CI se referă la *toate instanțele* posibile ale unei tabele
- Cheile primare și externe sunt cele mai comune CI;

# Nivelele de abstractizare

Nivele diferite  
de abstractizare



**STUDENT**

Name

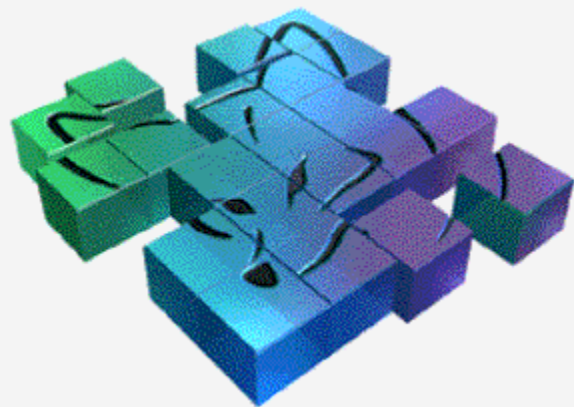
Date of birth

Sex

CNP

Group





# Structura fizică

Faculty.dbc

42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

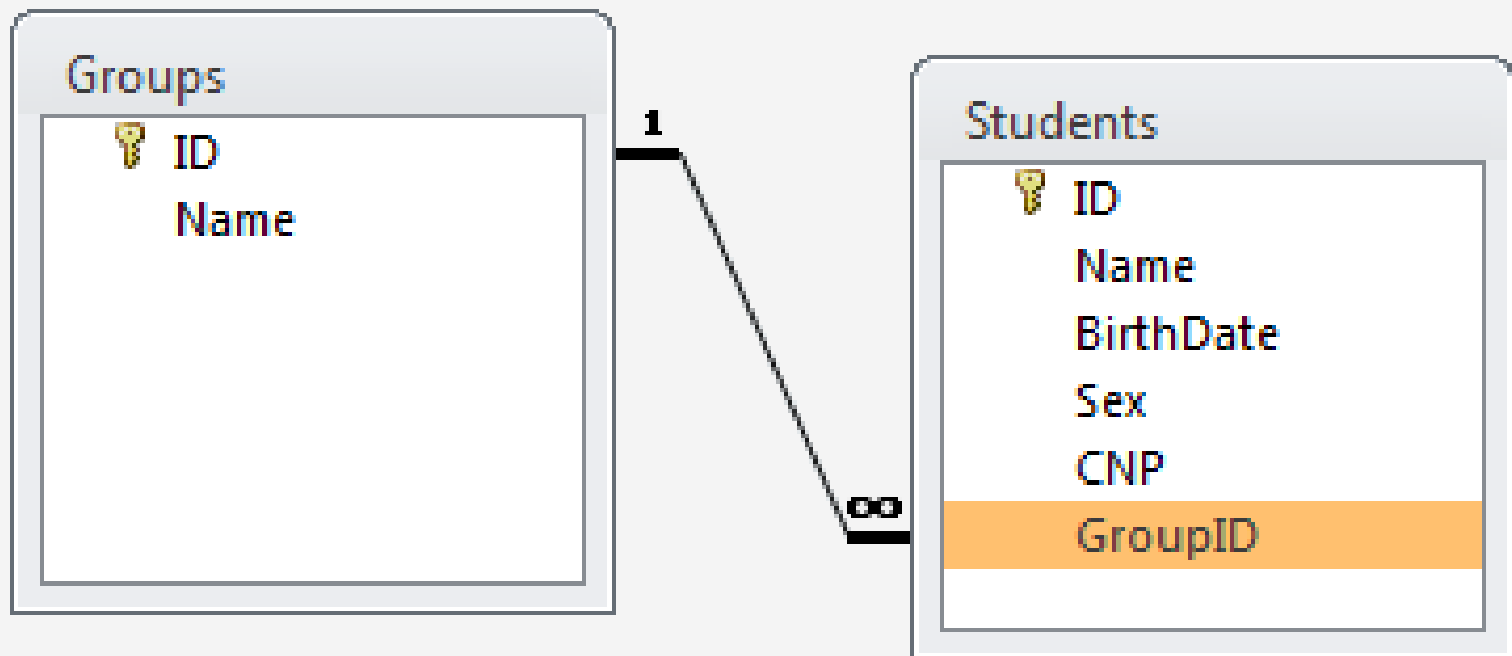
Students.dbf

20	20	20	31	56	31	2e	30
38	31	39	32	44	65	66	61
61	67	65	20	53	65	74	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	00	ff	01	00	7c	80	00
45	41	44	45	52	34	0f	53
4e	55	4d	42	45	52	14	34
34	21	0a	20	20	20	20	20
42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

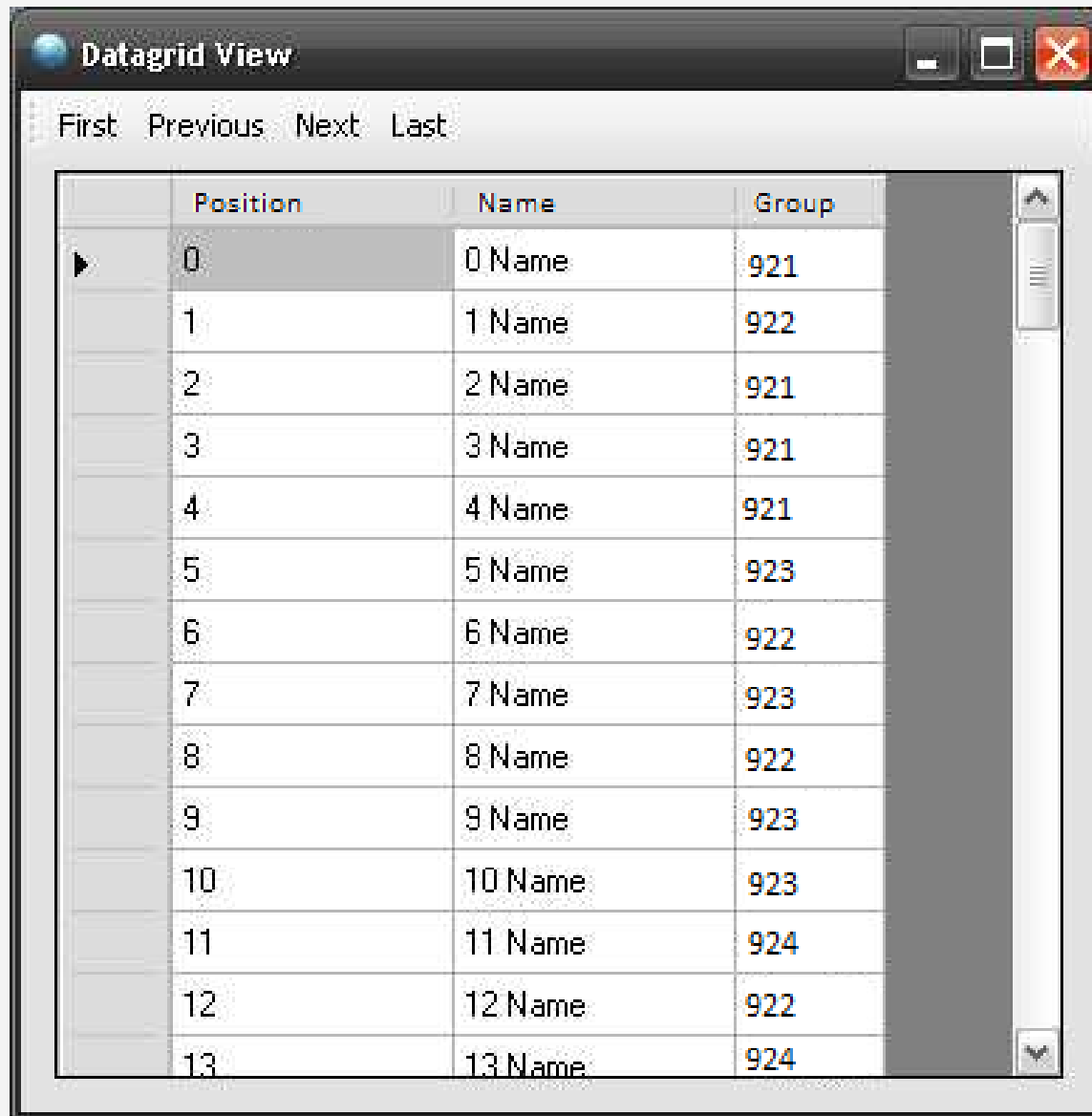
Groups.dbf

20	20	20	31	56	31	2e	30
38	31	39	32	44	65	66	61
61	67	65	20	53	65	74	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	00	ff	01	00	7c	80	00
45	41	44	45	52	34	0f	53
4e	55	4d	42	45	52	14	34
34	21	0a	20	20	20	20	20
42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

# Structura conceptuală



# Vizualizare pentru utilizator



The image shows a software window titled "Datagrid View". Inside the window, there is a table with the following data:

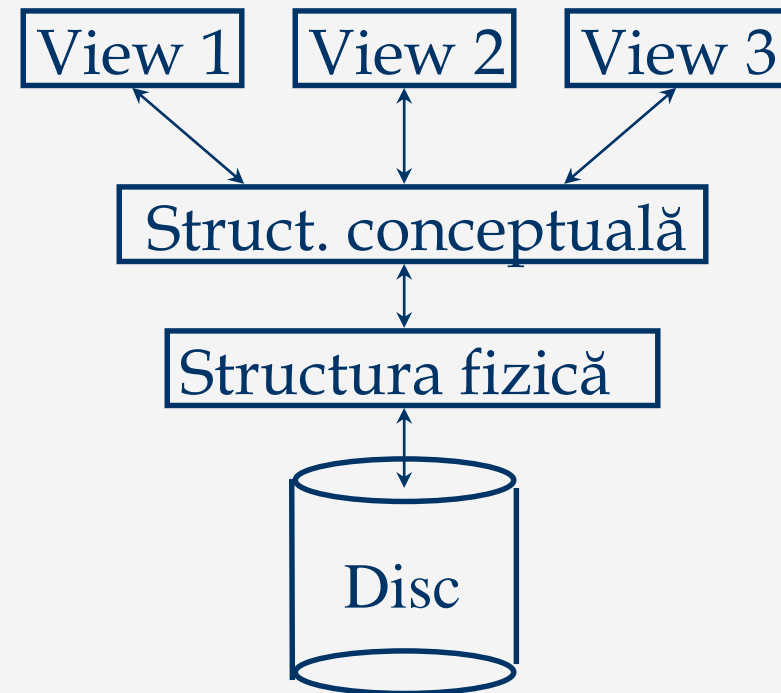
	Position	Name	Group
▶	0	0 Name	921
	1	1 Name	922
	2	2 Name	921
	3	3 Name	921
	4	4 Name	921
	5	5 Name	923
	6	6 Name	922
	7	7 Name	923
	8	8 Name	922
	9	9 Name	923
	10	10 Name	923
	11	11 Name	924
	12	12 Name	922
	13	13 Name	924

Navigation buttons: First, Previous, Next, Last.

# Nivelele de abstractizare

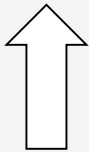
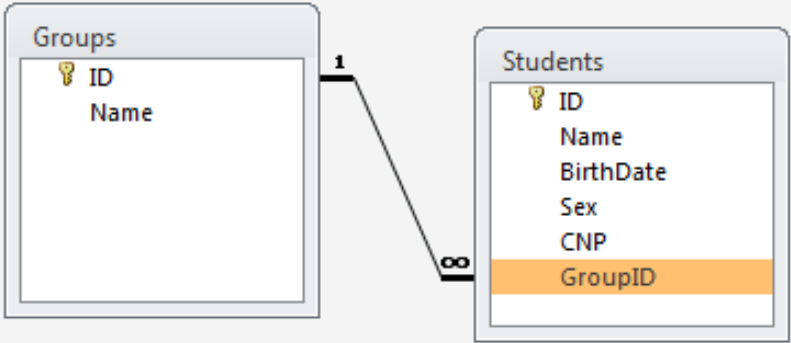
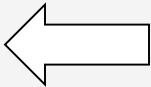
- Mai multe structuri externe (views), câte o singură structură conceptuală (logică) și o structură fizică (internă).

- *Views* – cum văd utilizatorii datele.
- *Conceptual* - modelul logic compus din relații, attribute, etc
- *Fizic* - fișierele de date și indecși





Datagrid View			
First Previous Next Last			
	Position	Name	Group
▶	0	0 Name	921
	1	1 Name	922
	2	2 Name	921
	3	3 Name	921
	4	4 Name	921
	5	5 Name	923
	6	6 Name	922
	7	7 Name	923
	8	8 Name	922
	9	9 Name	923
	10	10 Name	923
	11	11 Name	924
	12	12 Name	922
	13	13 Name	924



Faculty.dbc

42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

Students.dbf

20	20	20	31	56	31	2e	30
38	31	39	32	44	65	66	61
61	67	65	20	53	65	74	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	00	ff	01	00	7c	80	00
45	41	44	45	52	34	0f	53
4e	55	4d	42	45	52	14	34
34	21	0a	20	20	20	20	20
42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

Groups.dbf

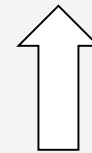
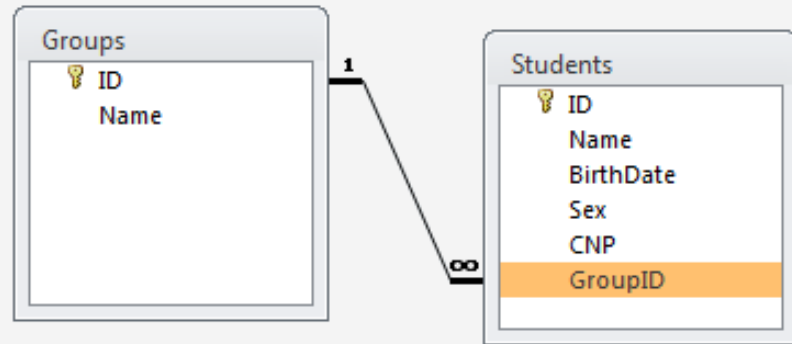
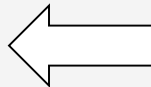
20	20	20	31	56	31	2e	30
38	31	39	32	44	65	66	61
61	67	65	20	53	65	74	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	00	ff	01	00	7c	80	00
45	41	44	45	52	34	0f	53
4e	55	4d	42	45	52	14	34
34	21	0a	20	20	20	20	20
42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

# Independența fizică a datelor

Datagrid View

First Previous Next Last

	Position	Name	Group
0		0 Name	921
1		1 Name	922
2		2 Name	921
3		3 Name	921
4		4 Name	921
5		5 Name	923
6		6 Name	922
7		7 Name	923
8		8 Name	922
9		9 Name	923
10		10 Name	923
11		11 Name	924
12		12 Name	922
13		13 Name	924



Faculty.mdf

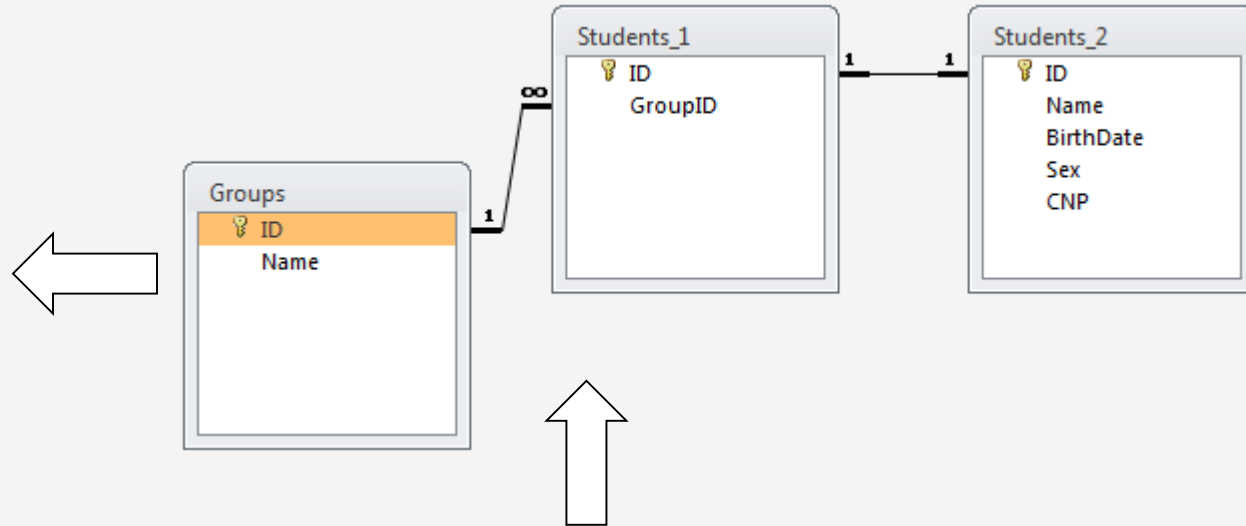
20	20	20	31	56	31	2e	30
38	31	39	32	44	65	66	61
61	67	65	20	53	65	74	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	00	ff	01	00	7c	80	00
45	41	44	45	52	34	0f	53
4e	55	4d	42	45	52	14	34
34	21	0a	20	20	20	20	20
42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

# Independența logică a datelor

Datagrid View

First Previous Next Last

	Position	Name	Group
▶	0	0 Name	921
	1	1 Name	922
	2	2 Name	921
	3	3 Name	921
	4	4 Name	921
	5	5 Name	923
	6	6 Name	922
	7	7 Name	923
	8	8 Name	922
	9	9 Name	923
	10	10 Name	923
	11	11 Name	924
	12	12 Name	922
	13	13 Name	924



Faculty.mdf

20	20	20	31	56	31	2e	30
38	31	39	32	44	65	66	61
61	67	65	20	53	65	74	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	00	ff	01	00	7c	80	00
45	41	44	45	52	34	0f	53
4e	55	4d	42	45	52	14	34
34	21	0a	20	20	20	20	20
42	53	54	42	20	2e	30	37
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
20	20	20	20	20	20	20	20
03	50	81	01	f0	06	4f	52
4c	45	2d	49	44	01	14	3c
54	2d	4e	41	4d	45	01	13
45	54	2d	4e	55	4d	42	45

# Interogări

- Posibile informații pe care dorim să le obținem din baza de date anterioară (*Faculty Database*) :
  - Care este numele studentului cu *sid* = 2833?
  - Care este salariul profesorilor care predau cursul *Alg100*?
  - Câți studenți sunt înscriși la cursul *Alg100*?
- Astfel de întrebări referitoare la datele stocate într-un SGBD se numesc *interogări*.
- → *limbaj de interogare*

# Limbaje SGBD

- *Data Definition Language (DDL)*
  - Definesc structura **conceptuală**
  - Descriu **constrângerile de integritate**
  - Influențează **structura fizică** (în anumite SGBD-uri)
- *Data Manipulation Language (DML)*
  - Operații aplicate instanțelor unei baze de date
  - DML procedural (**cum?**) vs. DML declarative (**ce?**)
- *Limbaaj gazdă*
  - Limbaj de programare obișnuit ce permite utilizatorilor să includă comenzi DML în propriul cod



# Limbaje de interogare pentru BD relaționale

SQL (Structured Query Language)

**SELECT** *name* **FROM** *Students* **WHERE** *age* > 20

Algebra

$\pi_{name}(\sigma_{age > 20}(Students))$

Domain Calculus

$\{ \langle X \rangle \mid \exists V \exists Y \exists Z \exists T : Students(V, X, Y, Z, T) \wedge Z > 20 \}$

T-uple Calculus

$\{ X \mid \exists Y : Y \in Students \wedge Y.age > 20 \wedge X.name = Y.name \}$

# Structured Query Language (SQL)

- Dezvoltat de IBM (*system R*) în anii 1970
- Ulterior a apărut nevoia de standardizare
- Standarde (ANSI):
  - SQL-86
  - SQL-89 (minor revision)
  - SQL-92 (major revision) - *1,120 pagini*
  - SQL-99 (major extensions) - *2,084 pagini*
  - SQL-2003 (secțiuni SQL/XML) - *3,606 pagini*
  - SQL-2006
  - SQL-2008
  - SQL-2011






# Nivele SQL




- Data-definition language (DDL):
  - Creare / stergere / modificare *tabele* și *views*.
  - Definire *constrangeri de integritate* (CI's).
- Data-manipulation language (DML)
  - Permit formularea de interogari
  - Inserare / ștergere / modificare înregistrări.
- *Controlul accesului:*
  - Asignează sau elimină drepturi de acces si modificare a *tabelelor* și a *view-urilor*.

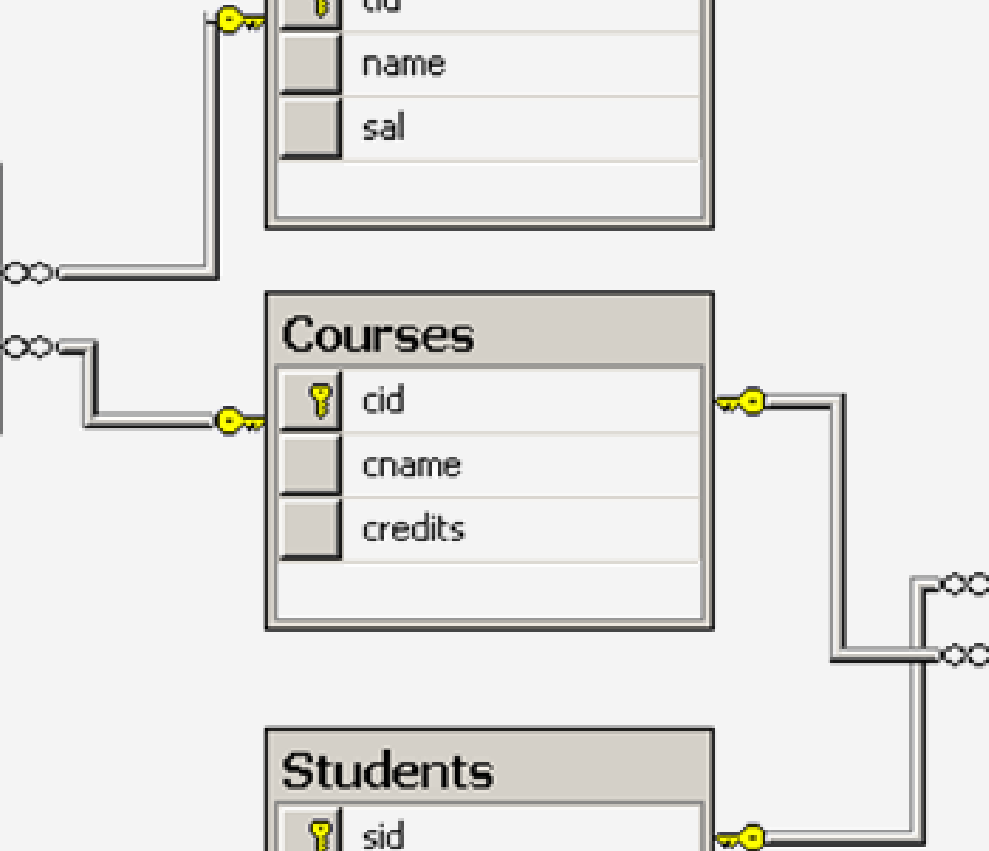
Teaches	
 tid	
 cid	

Teachers	
 tid	
 name	
 sal	

Courses	
 cid	
 cname	
 credits	

Students	
 sid	
 name	
 email	
 age	
 gr	

Enrolled	
 sid	
 cid	
 grade	



## *Students*

<i>sid</i>	<i>name</i>	<i>email</i>	<i>age</i>	<i>gr</i>
1234	John	j@cs.ro	21	331
1235	Smith	s@cs.ro	22	331
1236	Anne	a@cs.ro	21	332

## *Enrolled*

<i>sid</i>	<i>cid</i>	<i>grade</i>
1234	Alg1	9
1235	Alg1	10
1237	DB2	9

## *Courses*

<i>cid</i>	<i>cname</i>	<i>credits</i>
Alg1	Algorithms1	7
DB1	Databases1	6
DB2	Databases2	6



# SELECT

Studentii cu vârsta de 21 de ani:

```
SELECT *  
FROM Students S  
WHERE S.age = 21
```

<i>sid</i>	<i>name</i>	<i>email</i>	<i>age</i>	<i>gr</i>
1234	John	j@cs.ro	21	331
1236	Anne	a@cs.ro	21	332

Returnează doar numele și adresele de e-mail:

```
SELECT S.name, S.email  
FROM Students S  
WHERE S.age = 21
```

<i>name</i>	<i>email</i>
John	j@cs.ro
Anne	a@cs.ro

# Interogare SQL simplă

```
SELECT [DISTINCT] target-list  
FROM relation-list  
WHERE qualification
```

- *relation-list* - lista de nume de relații/tabele.
- *target-list* - listă de attribute ale relațiilor din  
*relation-list*
- *qualification* - comparații logice (*Attr op const* sau *Attr1 op Attr2*, unde *op* is one of  $<$ ,  $>$ ,  $=$ ,  $\leq$ ,  $\geq$ ,  $\neq$ ) combinate cu AND, OR sau NOT.
- *DISTINCT* (optional) - indică faptul că rezultatul final nu conține duplicate.

# Evaluare conceptuală

```
SELECT [DISTINCT] target-list  
FROM relation-list  
WHERE qualification
```

- Calcul produs cartezian al tabelelor din *relation-list*.
- Filtrare înregistrări ce nu verifică *qualifications*.
- Ștergere attribute ce nu aparțin *target-list*.
- Dacă **DISTINCT** e prezent, se elimină înregistrările duplicate.

1. PRODUS CARTEZIAN

2. ELIMINA LINII

3. ELIMINA COLOANE

4. ELIMINA DUPLICATE

Această strategie e **doar**  
la nivel *conceptual*!

Modul actual de evaluare  
a unei interogări  
e **mult** optimizat

# Range Variables

```
SELECT S.name, E.cid  
FROM Students S, Enrolled E  
WHERE S.sid=E.sid AND E.grade=10
```



```
SELECT name, cid  
FROM Students, Enrolled  
WHERE Students.sid=Enrolled.sid  
AND grade=10
```



Interogare: *Studentii care au cel puțin o notă*

```
SELECT S.sid  
FROM Students S, Enrolled E  
WHERE S.sid=E.sid
```

- Rezultatul e diferit cu **DISTINCT**?
- Ce efect are înlocuirea *S.sid* cu *S.sname* în clauza **SELECT**?  
Rezultatul e diferit cu **DISTINCT** în acest caz?