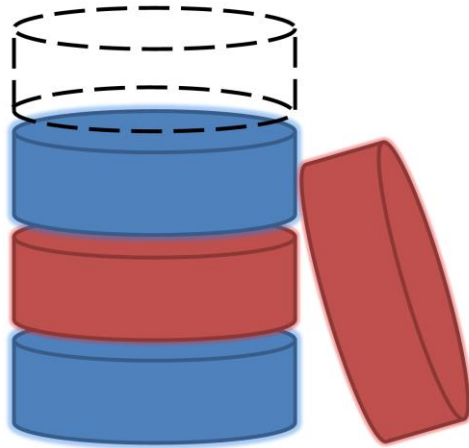


Modele de date



1



Sumar Cursuri

BD –concepte de bază

Modelul relațional:

- BD relaționale
- Forme normale
- Algebră relațională
- SQL

Metode de proiectare:

- Entity-Relationship Model
- Object Oriented Model
- Object – Relational Model

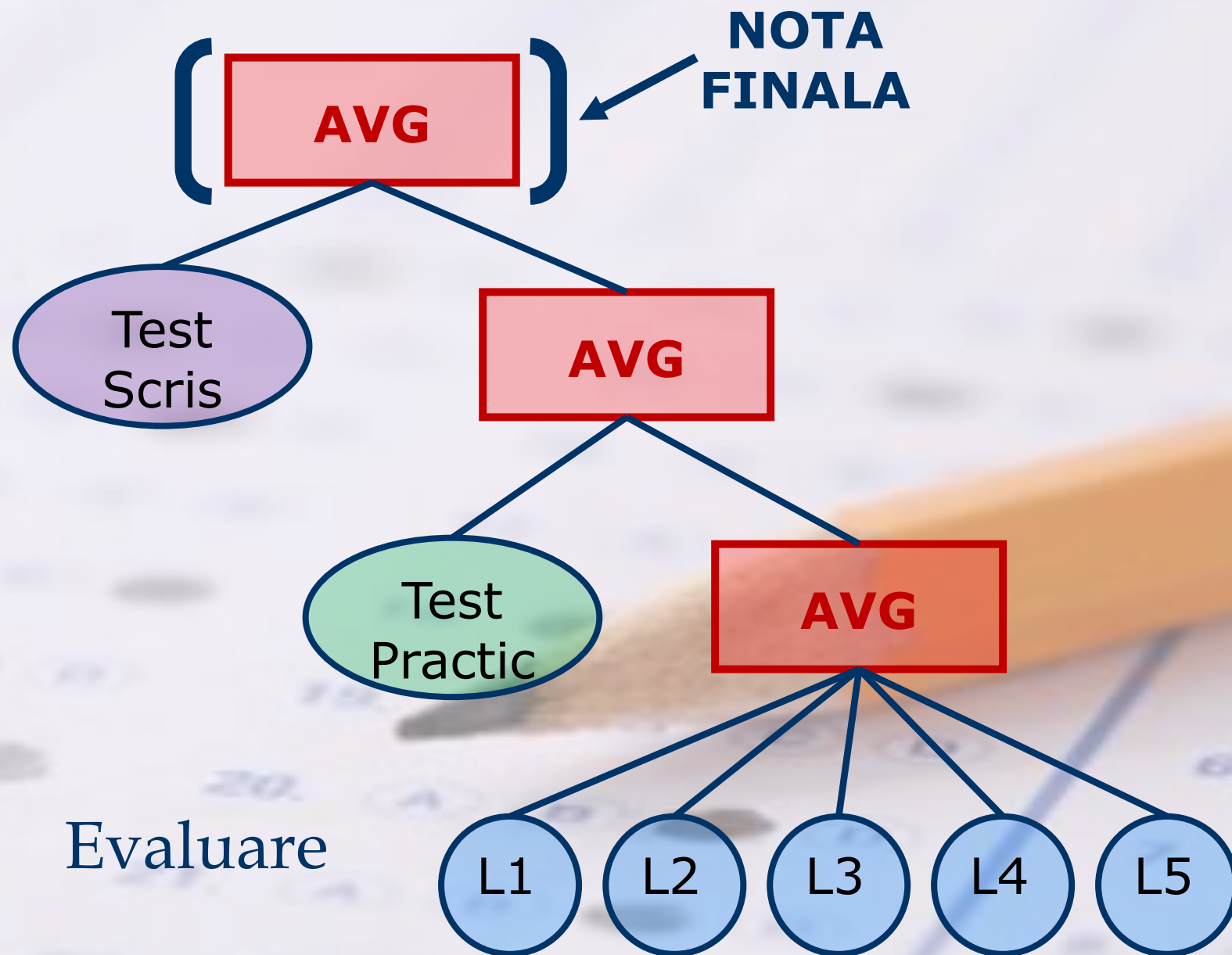
Structura fizică a BD

Indexarea BD

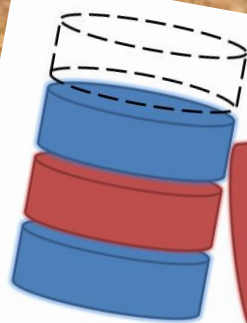
- Arbori B
- Hash Files

Lab: MS SQL Server





www.cs.ubbcluj.ro/~dsuciu
www.danmirceasuciu.ro



BAZE DE DATE
note de curs



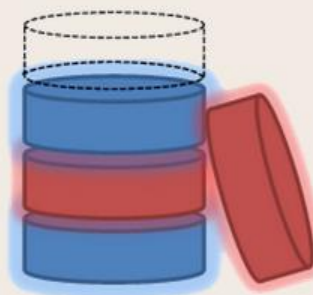
Home

Courses

Papers & Presentations

Hobbies

More



BAZE DE DATE

Ciclu de studii: Licență

Specializare: Informatică / Matematică- Informatică

Semestrul: 3

Ore: 2+1+2 / 2+1+1

Cod conectare curs - Microsoft Teams: 1s9g8le

Seminarii Informatică / Matematică-Informatică: Camelia ANDOR ([email](#), [pagină web](#))

Laboratoare Informatică:

Camelia ANDOR ([email](#), [pagină web](#))

Emilia POP ([email](#), [pagină web](#))

Ioan SIMA ([email](#), [pagină web](#))

Bogdan IOAN ([email](#))

Ce este o Bază de date?



large collection of related data items



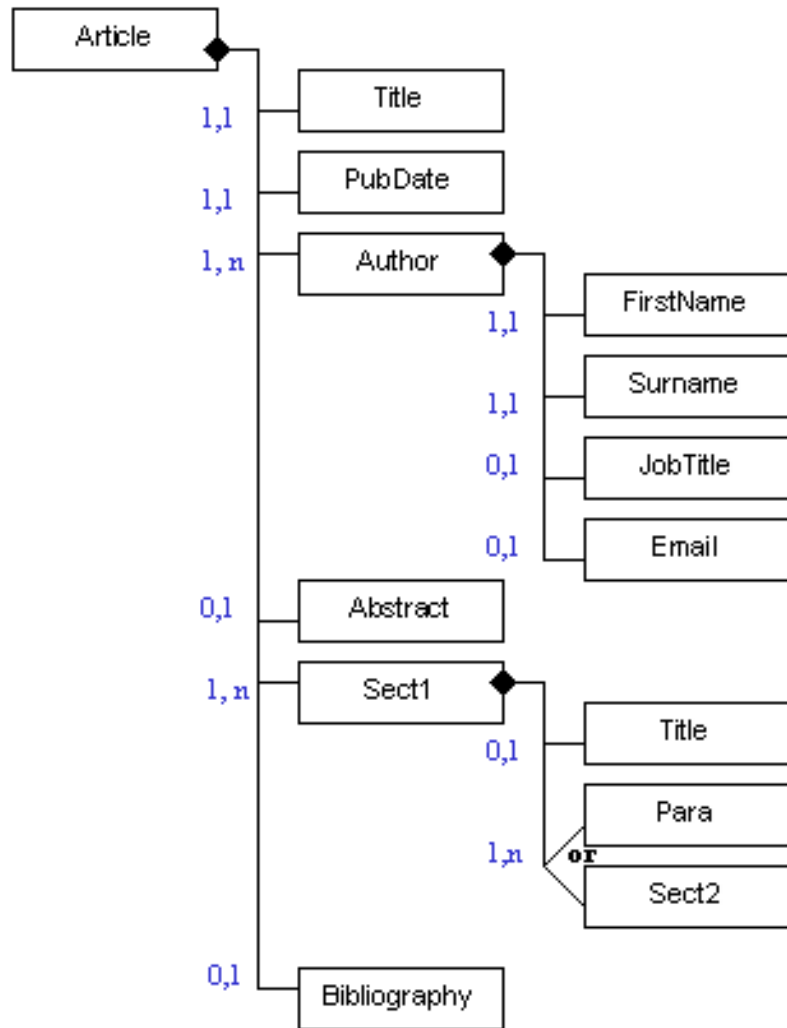
stored for record-keeping and analysis



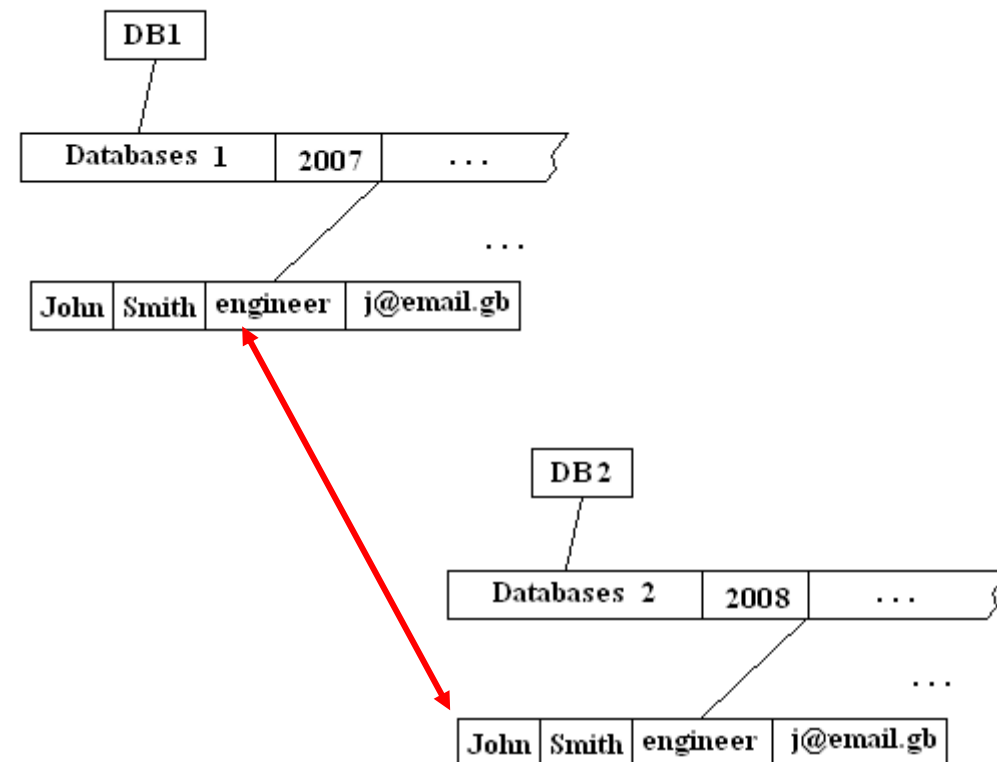
that exists over a long period of time

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

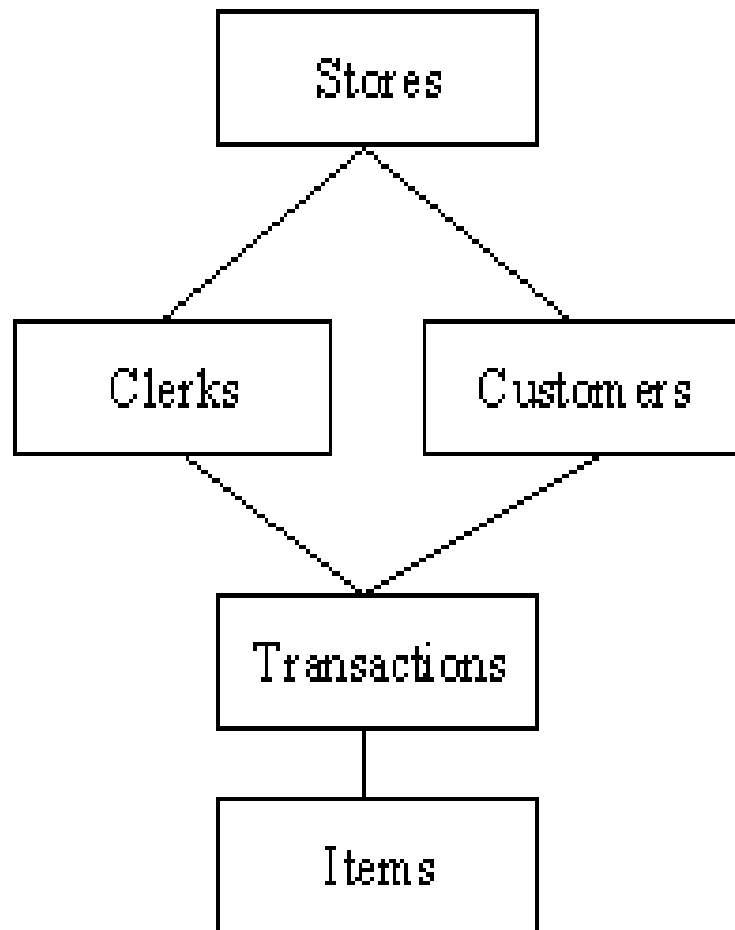
Modelul ierarhic



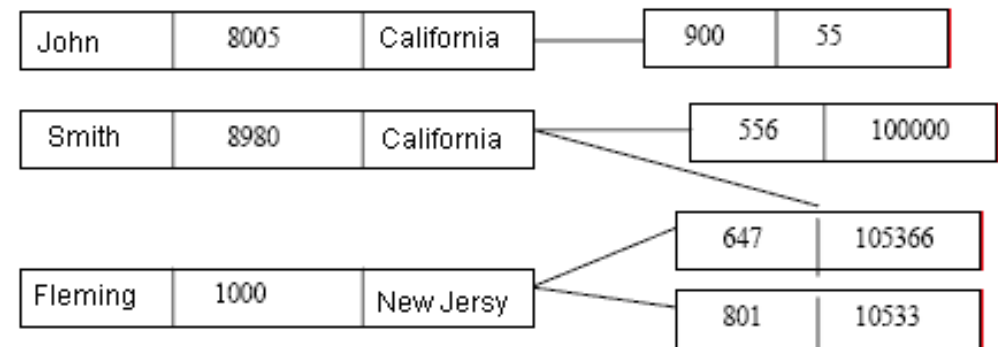
- Concepte:
 - structură arborescentă,
 - tip entitate,
 - relații 1-n,...



Modelul rețea

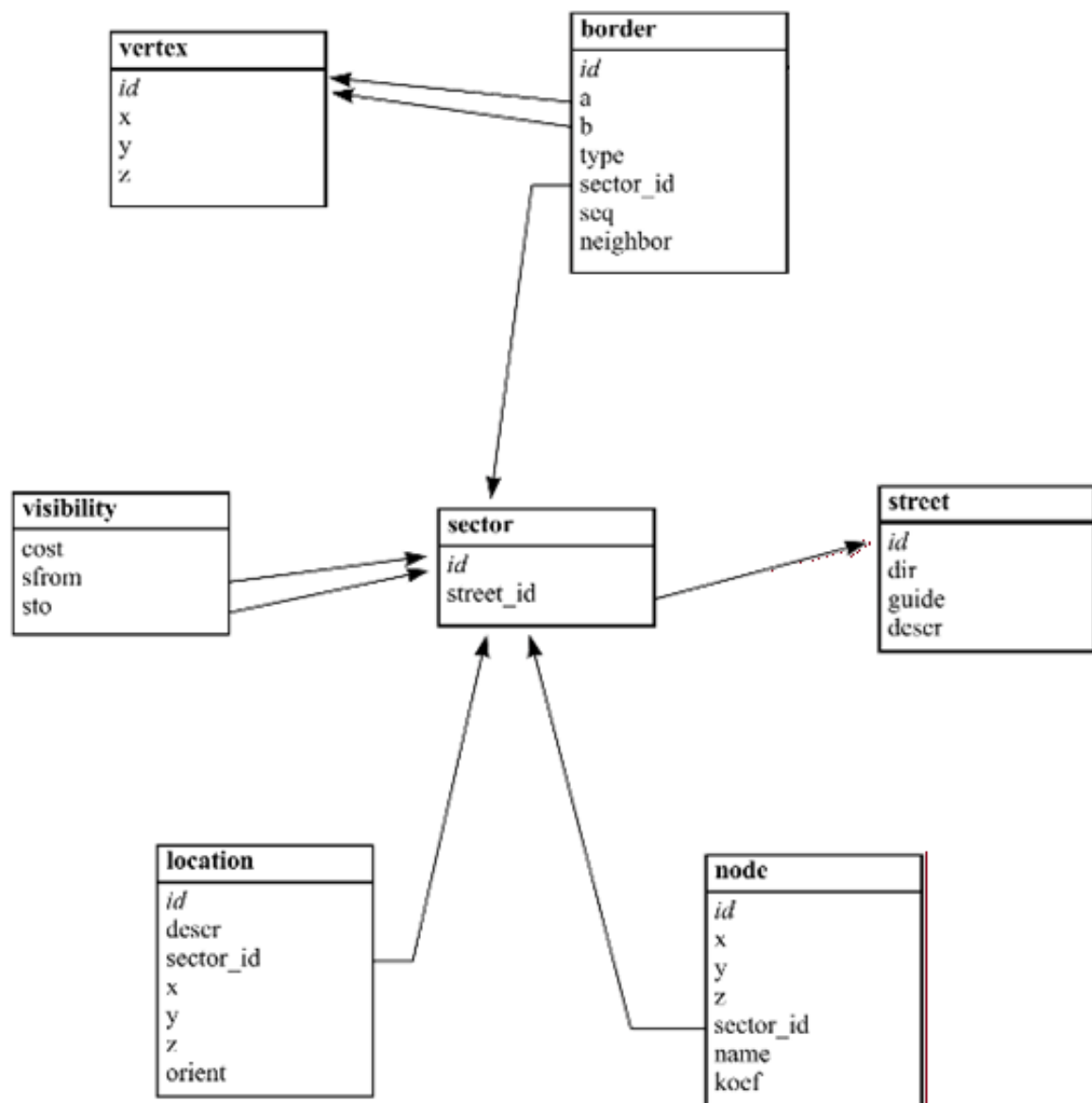


- Concepte:
 - structură de graf,
 - tip articol,
 - tip relație,...



Modelul relațional

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

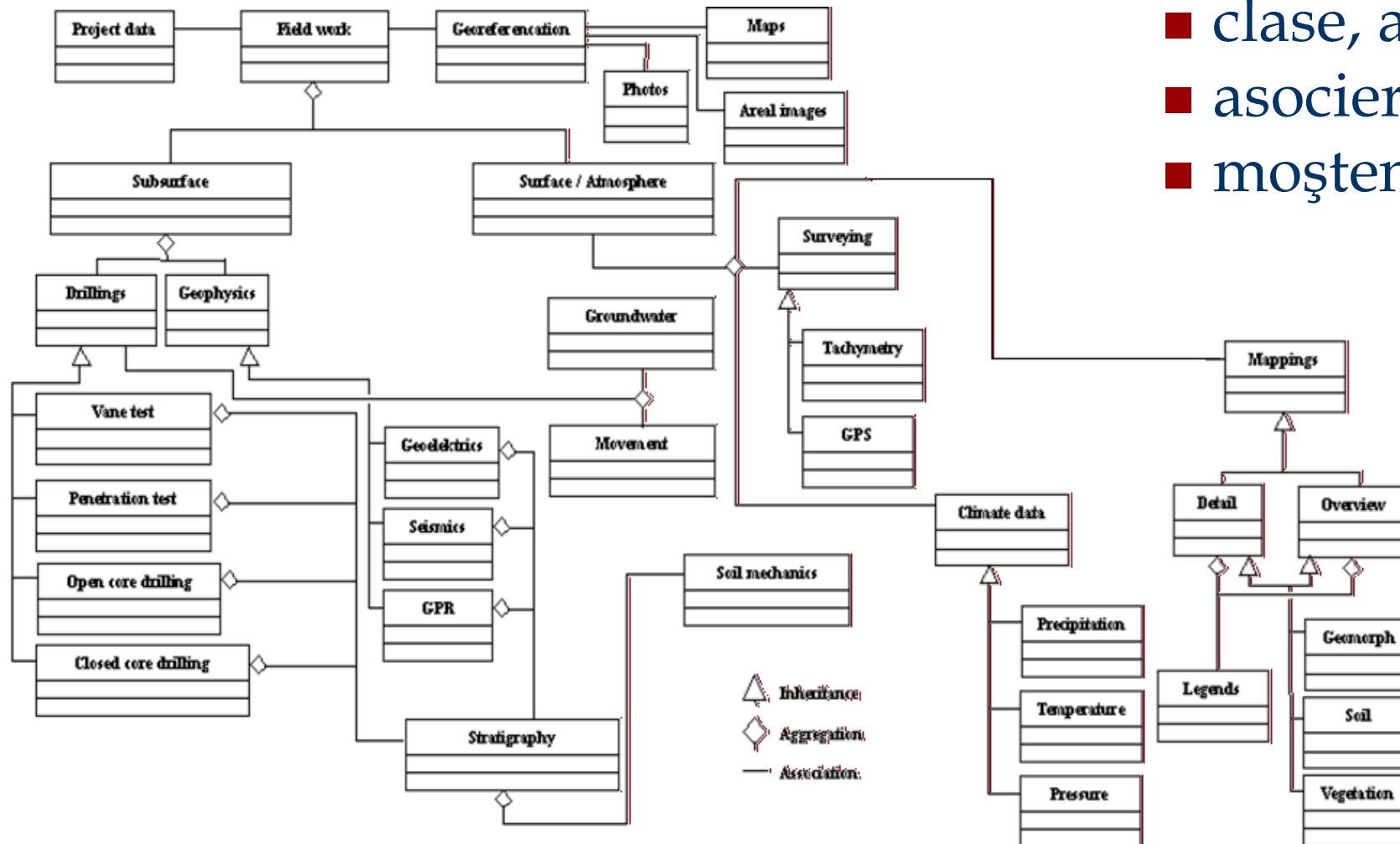


vertex

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

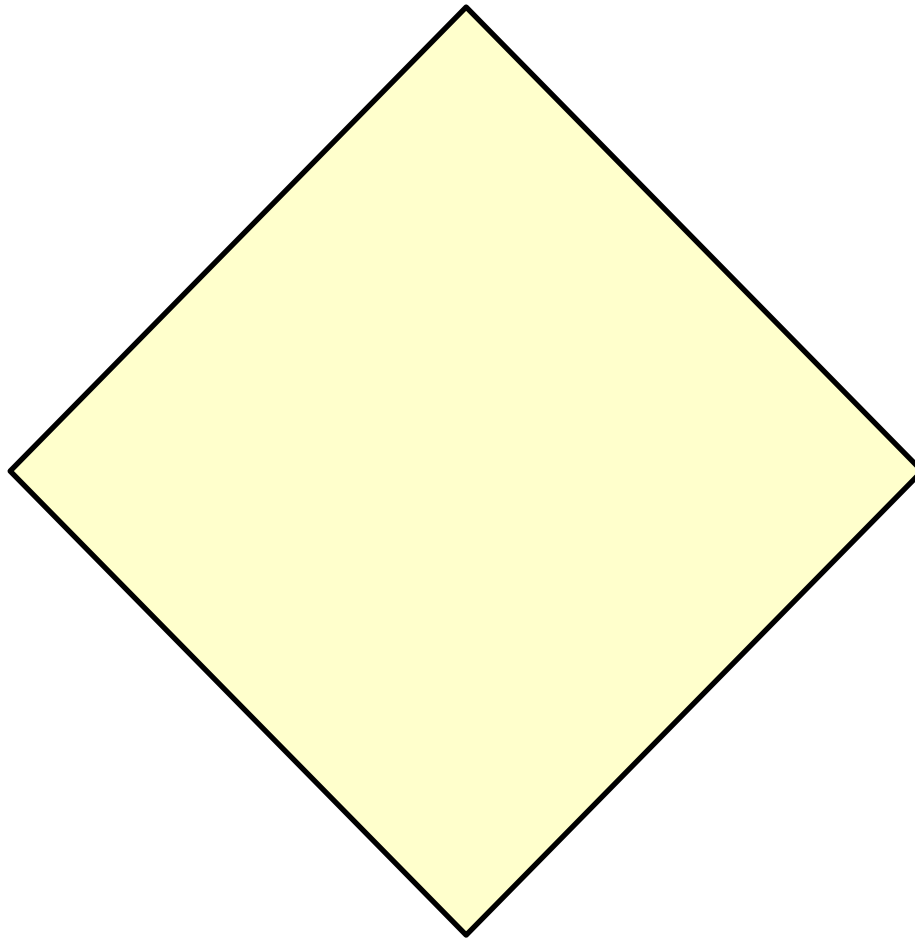
- Concepte:
 - clase, atribut, metode
 - asociere, agregare
 - moștenire,...



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



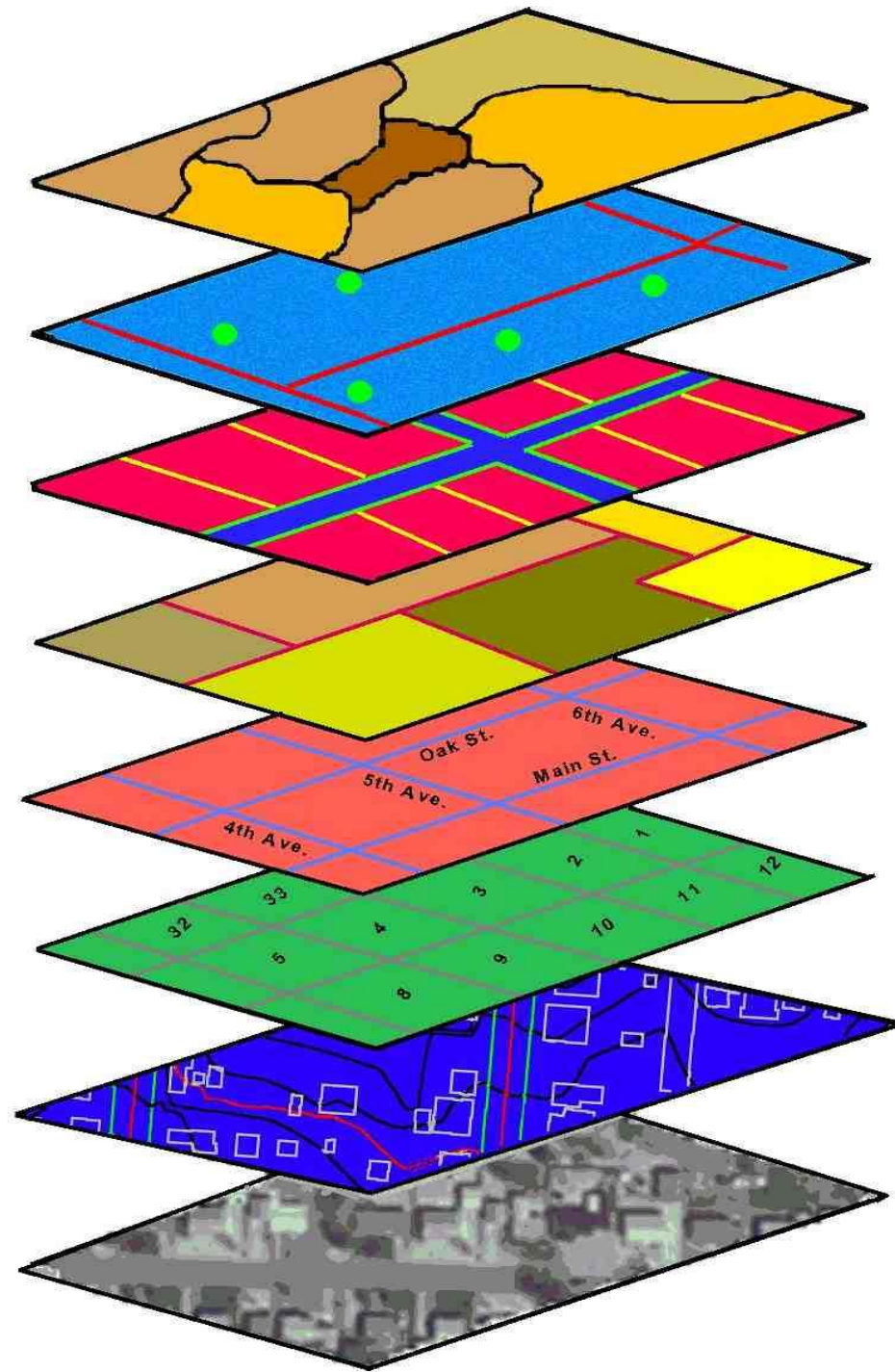
SGBD NoSQL





SGBD Multimedia

Geographical Information Systems (GIS)



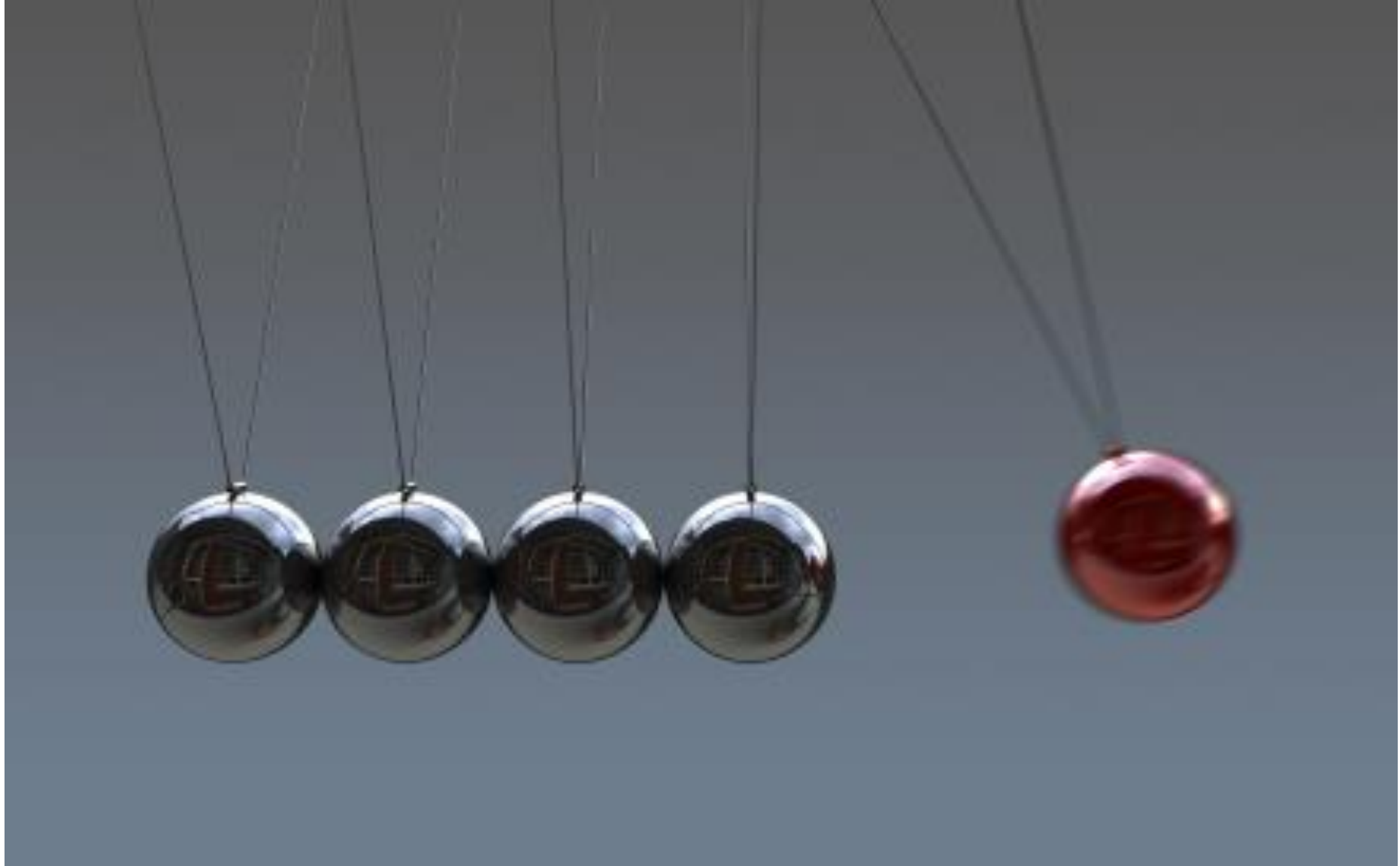


SGBD

Real-Time

Data Warehouse





SGBD-uri active

Data Stream Management System



DBMSs

are

getting

smaller

and

smaller

Databases

are

getting

bigger

and

bigger

P
A
R
A
L
L
E
L

Supporting

C
O
M
P
U
T
I
N
G

**Când
utilizăm
baze de date?**

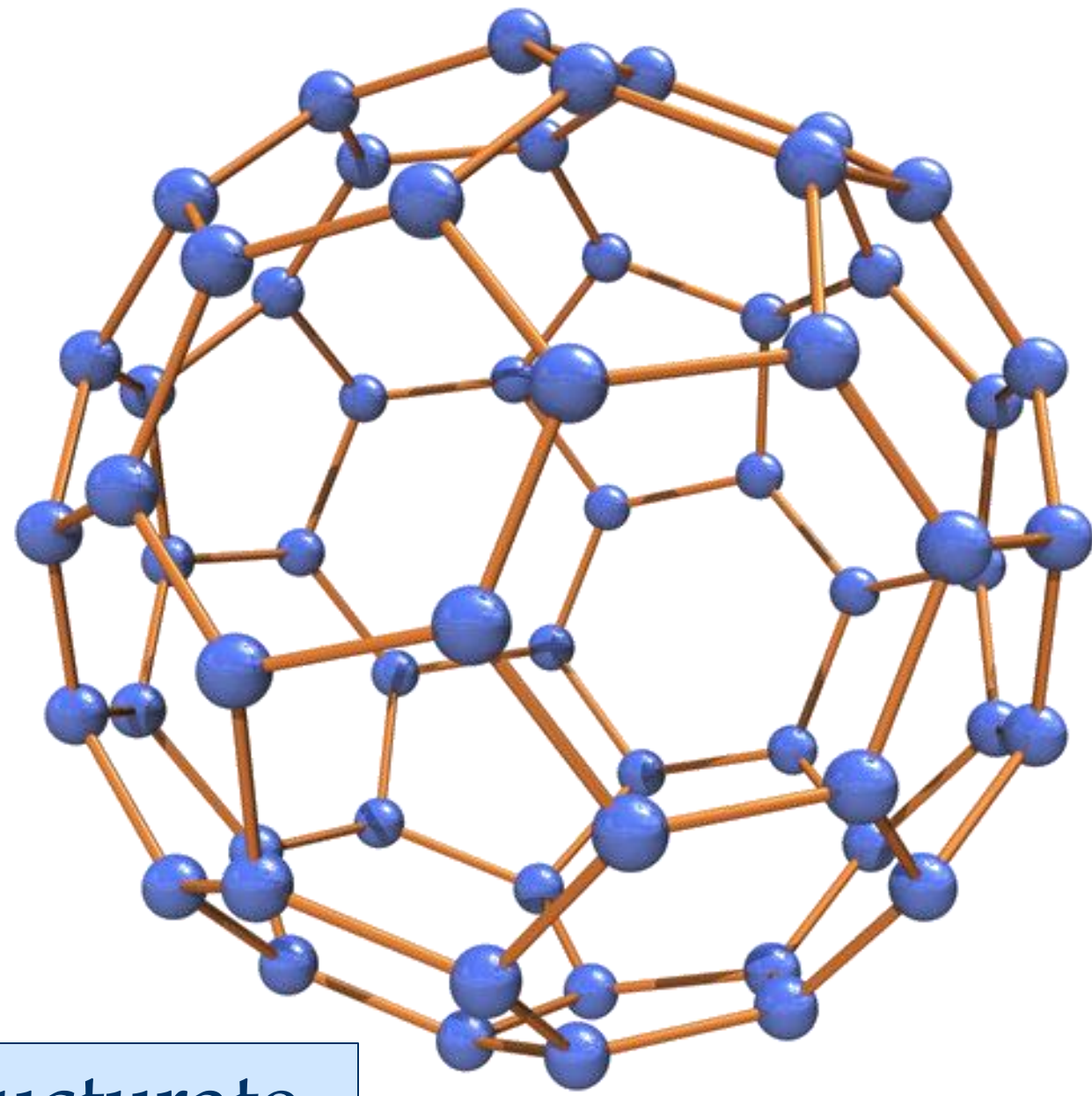


can't erase you

1. Persistență

2. Cantitate mare de date





3. Date structurate

4. Integritate

T R U S T



5. Acces Concurrent/Distribuit

6. Securitate



7. Transmiterea datelor altor aplicații



Când
NU
utilizăm
baze de date?

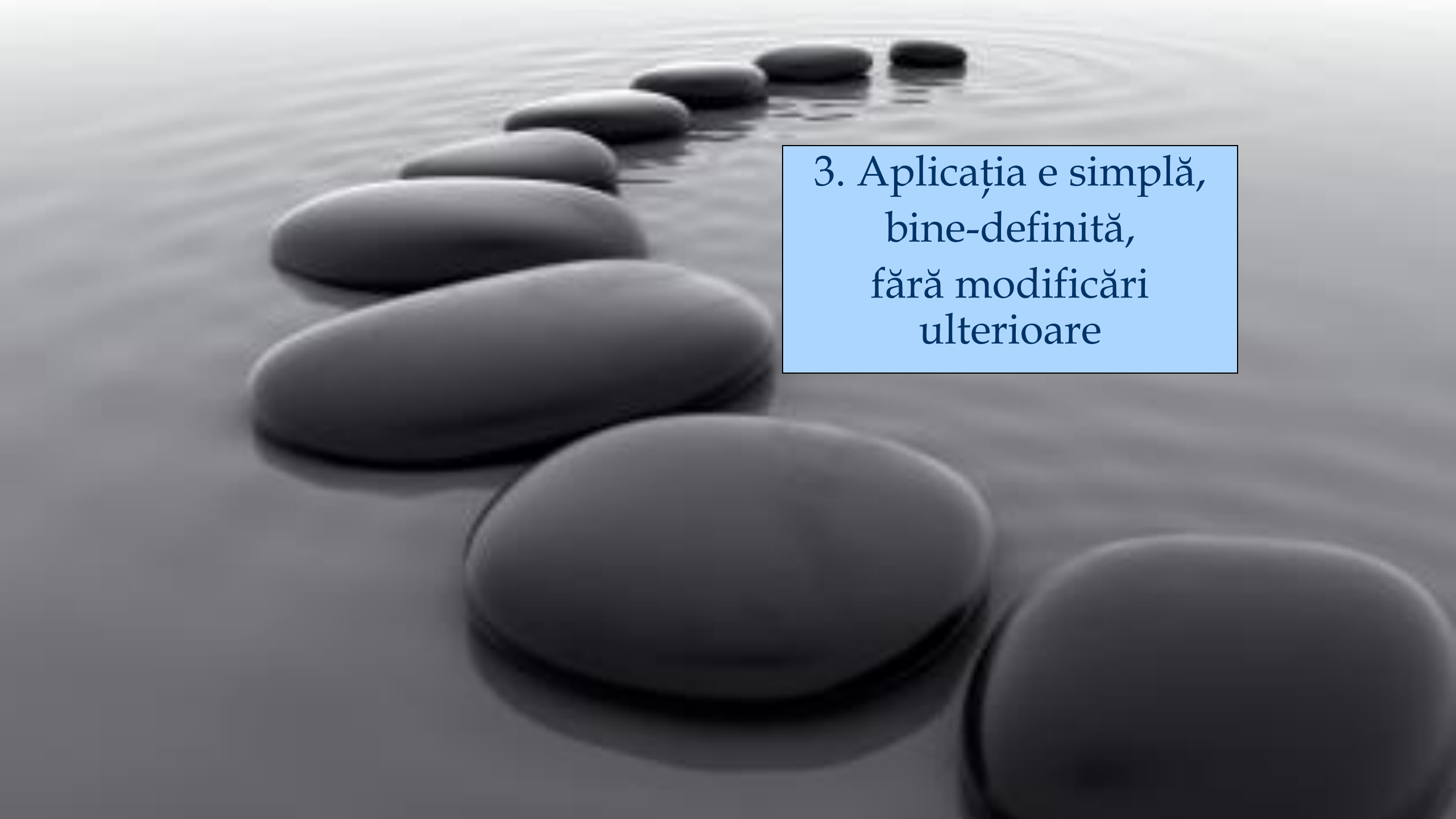


1. Investiția inițială e prea mare





2. Prea mult efort



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



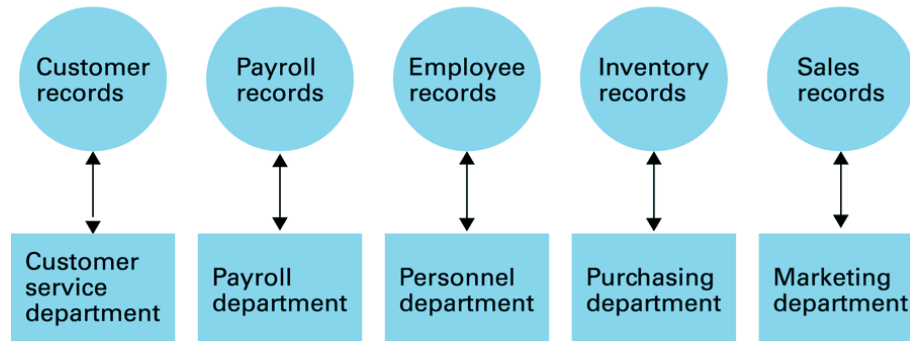
4. Nu este necesar accesul mai multor utilizatori la date

Alternativa:
fişier text

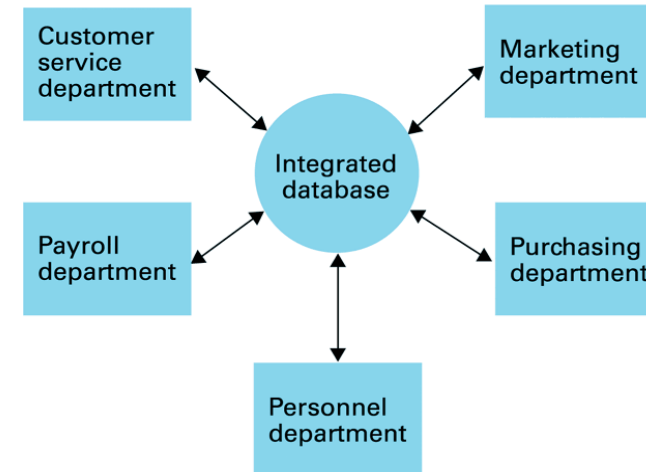
Fișiere text **vs.** Bază de date

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

a. File-oriented information system



b. Database-oriented information system



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 atribut** $[A_1, A_2, \dots, A_n]$.
- **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)

field name

*field type
(domain)*

sid	name	email	age	gr	} <i>relation schema</i>
2833	Jones	<u>jones@scs.ubbcluj.ro</u>	19	231	
2877	Smith	<u>smith@scs.ubbcluj.ro</u>	20	232	} <i>relation instance</i>
2976	Jones	<u>jones@math.ubbcluj.ro</u>	21	233	
2765	Mary	<u>mary@math.ubbcluj.ro</u>	22	233	← <i>relation tuple</i>

- 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

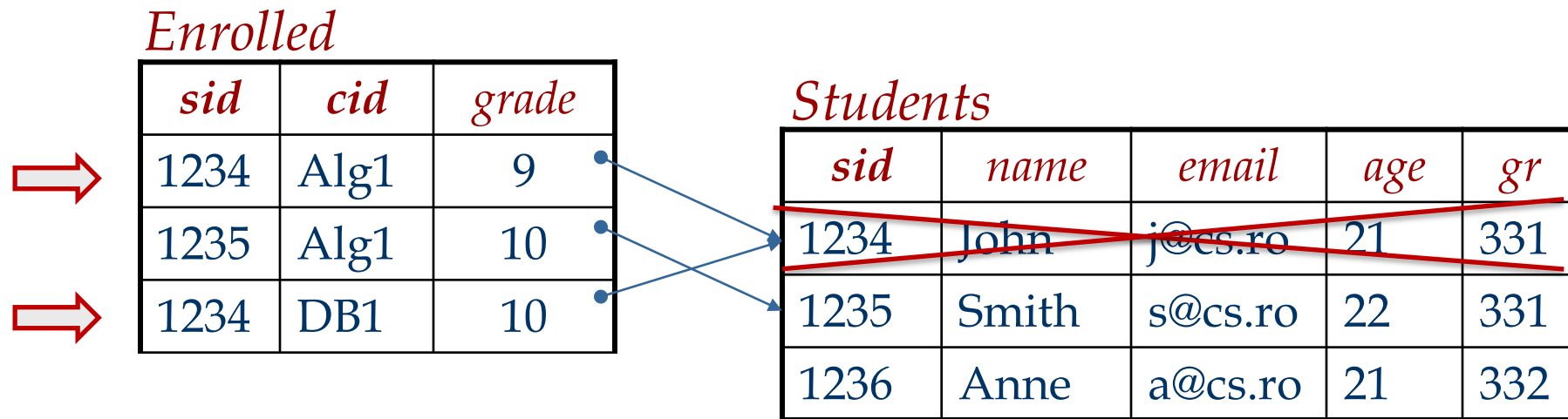
<i>sid</i>	<i>cid</i>	<i>grade</i>
1234	Alg1	9
1235	Alg1	10
1234	DB1	10
1237	DB2	9

Students

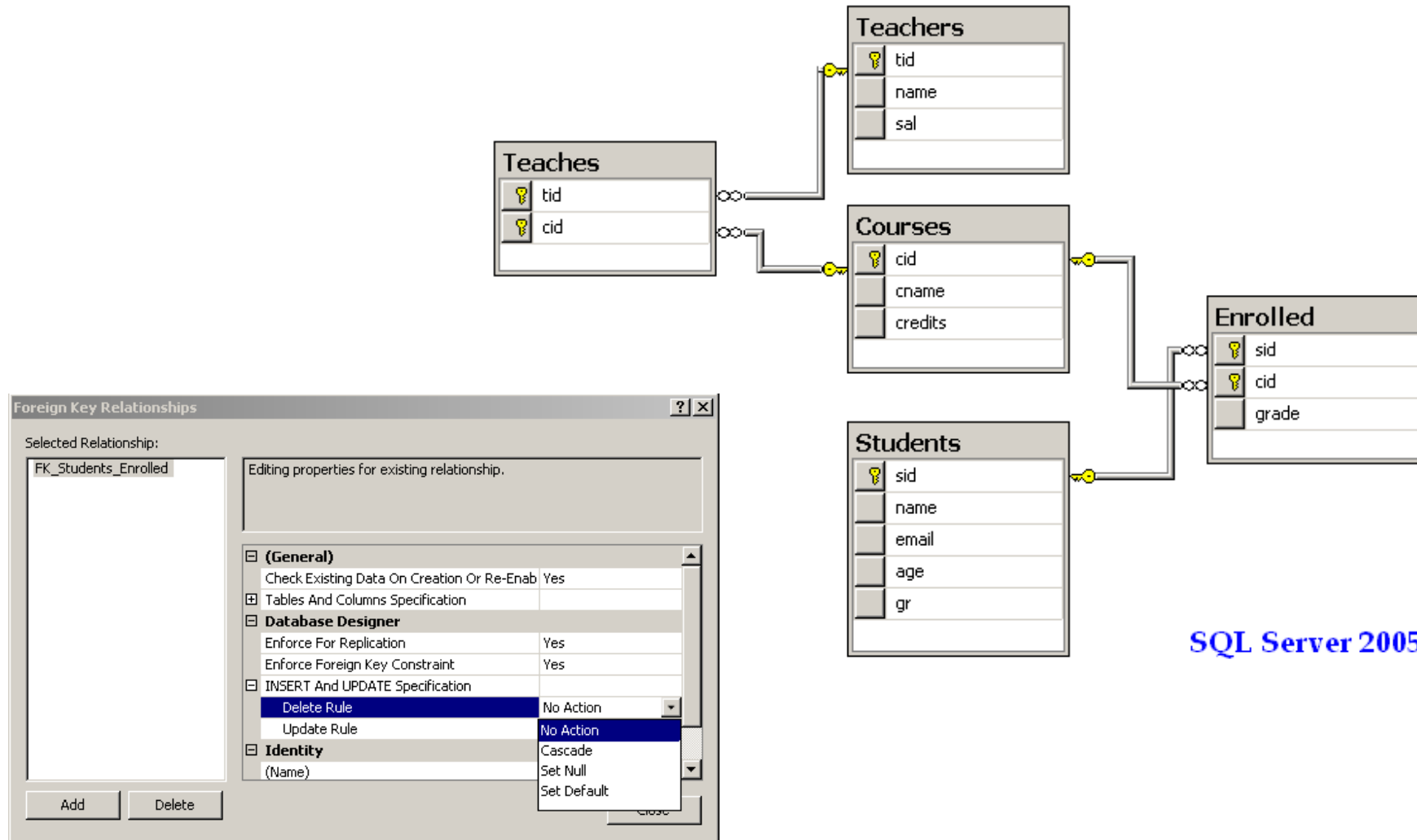
<i>sid</i>	<i>name</i>	<i>email</i>	<i>age</i>	<i>gr</i>
1234	John	j@cs.ro	21	331
1235	Smith	s@cs.ro	22	331
1236	Anne	a@cs.ro	21	332

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



SQL Server 2005