



EPL342 –Databases

Lecture 1: Introduction I

Introduction to Databases

(Chapter 1: Elmasri-Navathe, 5ED)

Διδάσκων: Παναγιώτης Ανδρέου

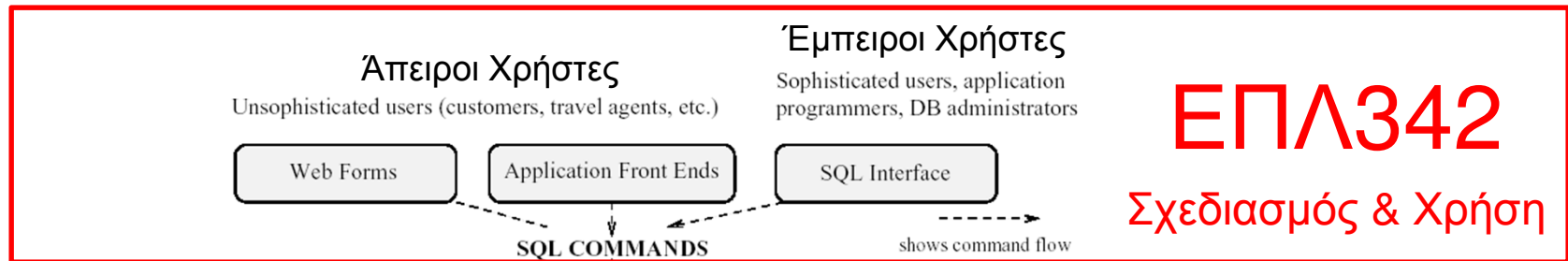
<http://www.cs.ucy.ac.cy/courses/EPL342>



Στόχοι Μαθήματος

- **ΕΠΛ342: Εφαρμογές των Βάσεων Δεδομένων**
 - Εισαγωγή στις **βασικές έννοιες διαχείρισης δεδομένων**
 - **Σχεδιασμός και Χρήση** συστημάτων βάσεων δεδομένων
 - **Πρακτική Εξάσκηση** στην εφαρμογή αυτών των εννοιών χρησιμοποιώντας **βιομηχανικό σύστημα διαχείρισης βάσης δεδομένων.**
- **ΕΠΛ446: Προχωρημένες Βάσεις Δεδομένων**
 - **Κατανόηση και Υλοποίηση** προχωρημένων εννοιών που αφορούν την **εσωτερική λειτουργία** μιας σχεσιακής βάσης δεδομένων
 - Έκθεση σε **Προχωρημένα και Ανερχόμενα Θέματα** στο πεδίο των βάσεων δεδομένων.

Τι είναι μια DBMS; Ανατομία ενός SQL Server



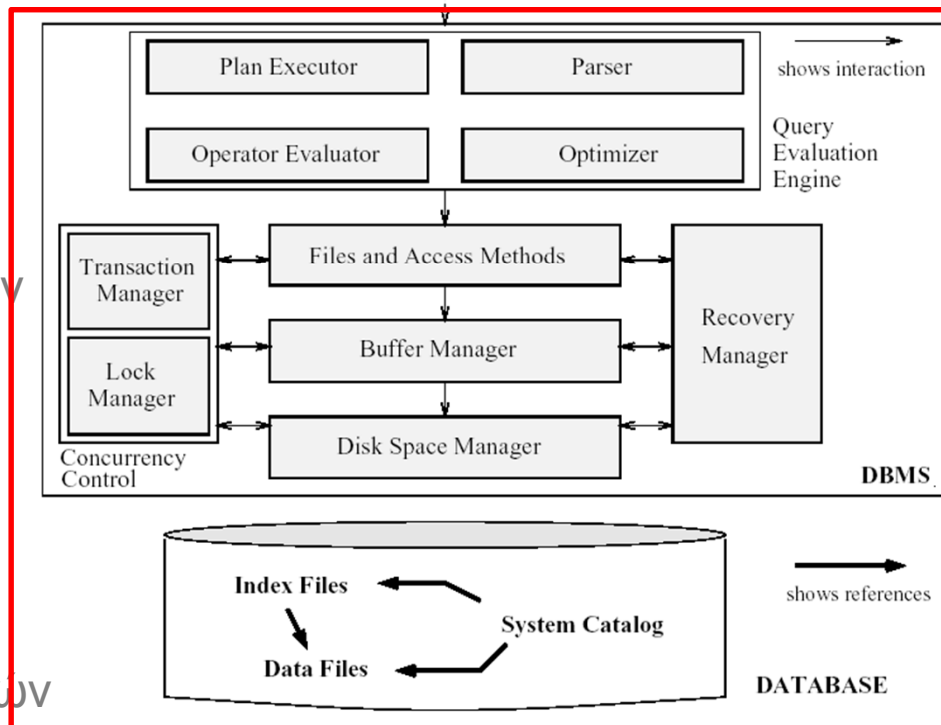
ΕΠΛ342

Σχεδιασμός & Χρήση

Δοσοληψιών και
Τεχνικές Ελέγχου
Ταυτοχρονίας

Δομή Ευρετηρίων
Δευτερεύουσας
Μνήμης (Hash,
B+)

Αποθήκευση
Δεδομένων και
Μετα-πληροφοριών



Βελτιστοποίησης
Επερωτήσεων

Τεχνικές
Ανάκαμψης (σε
περιπτώσεις
σφαλμάτων)

Ενδόμνημη
Διαχείριση
Δεδομένων

ΕΠΛ446



Συμβόλαιο Μαθήματος

- **Επίπεδο:** Προπτυχιακό
 - Υποχρεωτικό Μάθημα (Γενική Κατεύθυνση & Υπ. Συστημάτων και Δικτύων)
- **Πίστωση:** 7.5 μονάδες ECTS
- **Προαπαιτούμενα:**
 - ΕΠΛ231: Δομές Δεδομένων & Αλγόριθμοι
 - ΕΠΛ111: Διακριτές Δομές στην Πληροφορική και τον Υπολογισμό
 - ΕΠΛ132: Αρχές Προγραμματισμού II
- **Μέθοδοι Διδασκαλίας**
 - Διαλέξεις (3 ώρες εβδομαδιαίως): Παράδοση Διδασ. Ύλης
 - Φροντιστήριο (1 1/2 ώρα εβδομαδιαίως): Θεωρητική Εξάσκηση
 - Εργαστήριο (1 ώρα εβδομαδιαίως): Πρακτική Εξάσκηση



Συμβόλαιο Μαθήματος

- **Αξιολόγηση**

- 50% Τελική Εξέταση (1)

- 25% Ενδιάμεση Εξέταση (1)

- **Ημερ.: Πέμπτη, 21 Μαρτίου 2013! (9^η Εβδ.)**

- 25% Ασκήσεις - Exercises

- 10% Θεωρητικές ή/και Προγραμματιστικές.
Ασκήσεις (4)

- 15% Project Μαθήματος σε Εμπορική Βάση
Δεδομένων.



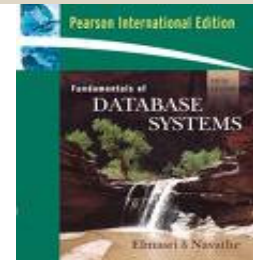
Βιβλιογραφία

Βασική Βιβλιογραφία

A1) Fundamentals of Database Systems, 5/E Ramez Elmasri, Shamkant B. Navathe ISBN: 0-321-36957-2 Publisher: AW 1168 pages, 2007.

Η' Ελληνική Μετάφραση

A2) Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων - 5η έκδοση (1ος τόμος), γραμμένο από τους Ramez Elmasri και Sham B. Navathe (μετάφραση Μ. Χατζόπουλος), Δίαυλος, 2007.

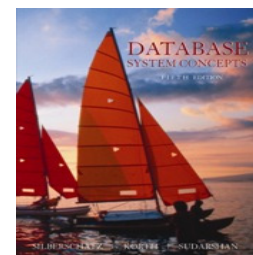
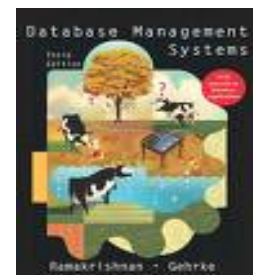


Βοηθητική Βιβλιογραφία

B) Σημειώσεις Μαθήματος και Συνοδευτικό Υλικό

Γ) Database Management Systems, 3rd Edition R. Ramakrishnan, & Johannes Gehrke, 1104 pp. McGraw-Hill Publisher, ISBN 0-07-123057-2, 2003.

Δ) Database Systems Concepts, 5th Edition, Abraham Silberschatz, Henry F. Korth, S. Sudarshan Hardcover, 1168 pages, McGraw-Hill Publisher, ISBN 0072958863, 2006.



Ιστοσελίδα ΕΠΛ342



- Όλες οι πληροφορίες σχετικά με το μάθημα βρίσκονται στο ακόλουθο URL

<http://www.cs.ucy.ac.cy/courses/EPL342>

ΕΠΛ342
Βάσεις Δεδομένων

Οικοσελίδα Ανακοινώσεις Πρόγραμμα Εργαστήρια Ασκήσεις

Καλώς ήλθατε στην ιστοσελίδα του μαθήματος ΕΠΛ342 - Βάσεις Δεδομένων
Δευτέρα, Γενάρης 20, 2013

Το μάθημα στοχεύει στην εισαγωγή βασικών εννοιών που χρειάζονται για το σχεδιασμό και τη χρήση μιας βάσης δεδομένων αλλά και στη πρακτική εξάσκηση στην εφαρμογή αυτών των εννοιών χρησιμοποιώντας ένα βιομηχανικό σύστημα διαχείρισης βάσης δεδομένων. Το περιεχόμενο του μαθήματος περιλαμβάνει την εισαγωγή σε έννοιες βάσεων δεδομένων, το Μοντέλο Οντοτήτων-Συσχετίσεων, το Σχηματικό Μοντέλο και τη Σχεσιακή Άλγεβρα, τη Γλώσσα Δομημένων Επερωτήσεων SQL και αλλά προχωρημένα θέματα όπως Συναρτησιακές Εξαρτήσεις, Κανονικοποίηση και Μεθοδολογία Ανάπτυξης Βάσεων Δεδομένων.

Συμβόλαιο Μαθήματος
Συμβόλαιο Μαθήματος [download \(.pdf\)](#)

Διδάσκοντες Μαθήματος

Διδάσκων Καθηγητής:	Παναγιώτης Ανδρέου
Τηλέφωνο:	+357 22 89 27 49
ΦΑΞ:	+357 22 89 27 01
Γραφείο:	ΘΕΕ01 - Β115
Ηλ. Ταχυδρομείο:	panic-AT-cs.ucy.ac.cy
Ώρες Γραφείου:	Τρίτη και Παρασκευή, 12:00-13:00, ή κατόπιν συνεννόησης
Διαλέξεις:	Δευτέρα και Πέμπτη, 17:30 – 21:00, ΧΩΔ-01 #tbd
Φροντιστήριο:	Τρίτη, 08:00 – 09:00, ΧΩΔ-01 #tbd
Εργαστήριο:	Τετάρτη (18:00-19:30), ΧΩΔ-01 #tbd
Υπεύθυνοι Εργαστηρίου:	Χριστόφορος Παναγιώτου (panchris-AT-cs.ucy.ac.cy), Γραφείο:ΘΕΕ01 Β130, Ώρες Γραφείου: tbd

Τελευταίες Ανακοινώσεις
Καλώς ορίσατε στο ΕΠΛ342. 21/01/2013

Τελευταίες Ασκήσεις

Σημαντικές Πληροφορίες
Το οριστικό υλικό θα αναρτηθεί πριν την αντίστοιχη διάλεξη 03/09/2012

EPL342 Moodle



- Για τις εκπαιδευτικές δραστηριότητες του μαθήματος (υποβολή εργασιών, φόρουμ ανακοινώσεων, ερωτηματολόγια, βαθμολογίες εργασιών, κτλ) θα χρησιμοποιηθεί το Moodle: <http://moodle.cs.ucy.ac.cy/>

A screenshot of a web browser displaying the Moodle course page for EPL342: Databases. The page is titled "EPL342: Databases (ΕΠΛ342: Βάσεις Δεδομένων)" and lists the instructor as Demetris Zeinalipour. The course outline includes six items: Assignment #1, Assignment #2, Assignment #3, Assignment #4, Final Project, and Final Grades. The page also features a sidebar with navigation options like People, Activities, Search Forums, Administration, and My courses. The browser's address bar shows the URL http://moodle.cs.ucy.ac.cy/course/view.php?id=50.



Περιεχόμενο Διάλεξης 1

- Τύποι Βάσεων και Εφαρμογές τους
- Βασικοί Ορισμοί
- Τυπικές Λειτουργίες μιας DBMS
- Παράδειγμα Βάσης (UNIVERSITY)
- Βασικά Πλεονεκτήματα της Προσέγγισης των Βάσεων (σε αντίθεση με Αρχεία)
- Χρήστες Βάσεων
- Άλλα Πλεονεκτήματα Χρήσης Βάσεων
- Πότε να μην χρησιμοποιούμε Βάσεις

Εφαρμογές Βάσεων Δεδομένων



- Αεροπορικές κρατήσεις (expedia.com)
- Νοσοκομεία, Χημεία
- Τραπεζικές Εφαρμογές (WellsFargo.com)
- Λογιστήρια
- Σύστημα εγγραφής φοιτητών
- Κοινωνικές Ασφαλίσεις, Φόροι
- Πωλήσεις – Αποθήκες (Amazon.com, eBay)
- Πιστωτικές κάρτες
- Χρηματιστήριο
- Τηλεπικοινωνίες
- ... και άλλες πολλές!

Τύποι Βάσεων Δεδομένων



- **Παραδοσιακές Εφαρμογές:**

- Αριθμητικές Βάσεις (Numerical) and Βάσεις Κειμένου (Text)

- Παράδειγμα Sprint Telecom (Αμερική): TOP-5 Database.

- Καταγραφή στοιχείων συνδιάλεξης (π.χ., ο Χ τηλεφώνησε του Υ για 10 λεπτά και 20 δευτερόλεπτα)
- 2,85 τρισεκατομμύρια ($\times 10^{12}$) εγγραφές!
- 385 εκατομμύρια ($\times 10^6$) νέες εγγραφές την ημέρα!
- Σε εξαιρετικές περιπτώσεις μέχρι και 70,000 νέες εγγραφές το Δευτερόλεπτο.

- **Πρόσφατες Εφαρμογές:**

- Βάσεις Πολυμέσων και Γεωγραφικών Πληροφ. Συστημάτων (GIS)

- Αποθήκες Δεδομένων (Data Warehouses)

- Βάσεις Ροών (Stream Databases), Βάσεις Νέφους (Cloud Databases), Βάσεις XML, κτλ

- Εμείς θα επικεντρωθούμε στις Παραδοσιακές Εφαρμογές

Βασικοί Ορισμοί



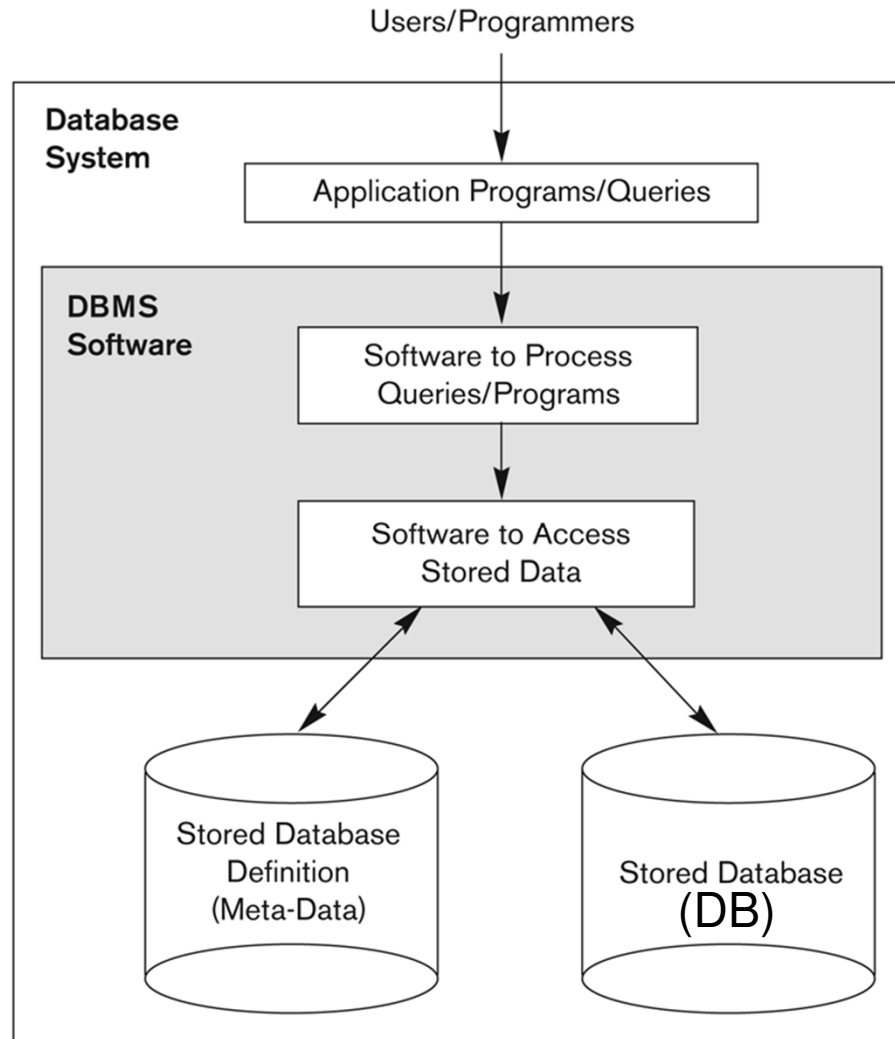
- **Βάση Δεδομένων (Database - DB):**
 - Μια συλλογή από **συσχετιζόμενα** δεδομένα.
 - Τυχαία συλλογή δεδομένων δεν αποτελούν βάση
 - Π.χ., όλες οι λέξεις ενός βιβλίου παρόλο που είναι συλλογή δεδομένων δεν αποτελούν βάση δεδομένων.
- **Δεδομένα (Data):**
 - **Γνωστά γεγονότα** τα οποία μπορούν να καταγραφούν και τα οποία έχουν **συγκεκριμένη σημασία**
 - Π.χ., Ο Κώστας είναι φοιτητής του ΕΠΛ342.
- **Mini-world (Σύνολο Αναφοράς – Universe of Discourse)**
 - Ένα μέρος (υποσύνολο) του **πραγματικού** κόσμου το οποίο **αντιπροσωπεύει μια DB.**
 - Π.χ., Οι εγγραφές των φοιτητών ενός Πανεπιστημίου

Βασικοί Ορισμοί



- **Database Management System (DBMS):**
 - Λογισμικό/α για τη **δημιουργία**, τη **διατήρηση**, και **επεξεργασία** μιας DB.
 - π.χ., Oracle, IBM DB2, SQL-Server, Postgres, MySQL, MS Access, Interbase, Sybase, ...
- **Κύριος σκοπός των DBMS:** αποθήκευση & διαχείριση
 - Εύκολα και **βολικά** (conveniently)
 - Γρήγορα και **αποδοτικά** (efficiently)
 - **Αξιόπιστα** και με **συνέπεια** (reliably and consistently)
- **Σύστημα Βάσης (Database System):**
 - DBS = DBMS + DATA
 - Κάποτε περιλαμβάνονται στην πιο πάνω εξίσωση και οι εφαρμογές.

Βασικοί Ορισμοί (Διαγραμματικά)



Τυπικές Λειτουργίες σε μια DBMS



- **A) Ορισμός βάσης:** τύποι δεδομένων (data types), δομές (structures), και περιορισμοί (constraints)
- **B) Δημιουργία ή Φόρτωση** του αρχικού περιεχομένου της βάσης δεδομένων στη δευτερεύουσα μνήμη (secondary storage medium)
- **Γ) (Ταυτόχρονη) Επεξεργασία της Βάσης:**
 - **Ανάκτηση** (Retrieval): Επερωτήσεις (Querying), Δημιουργία Αναφορών (Generating reports),...
 - **Τροποποίηση:** Εισαγωγές, Διαγραφές και Ενημερώσεις
 - **Ταυτόχρονη πρόσβαση** της βάσης από εφαρμογές, web, κτλ από πολλούς χρήστες
 - ... διασφαλίζοντας ότι τα δεδομένα είναι **ορθά (valid)** και **συνεπή (consistent)**
 - π.χ., δυο διαφορετικοί φοιτητές δεν μπορούν να έχουν τον ίδιο αριθμό ταυτότητας. Αυτός ο κανόνας πρέπει να ισχύει συνέχεια.

Τυπικές Λειτουργίες σε μια DBMS



- Άλλες Λειτουργίες σε μια DBMS:
 - Μέτρα **Προστασίας** ή **Ασφάλειας** για περιορισμό μη-εξουδιωτημένων (unauthorized) προσβάσεων στα δεδομένα.
 - **Αναπαράσταση** (Representation) και **Γραφική Απεικόνιση** (Visualization) των δεδομένων
 - **Βελτιστοποίηση Εκτέλεσης Ερωτημάτων**, Χώρου Αποθήκευσης, κτλ
 - Κτλ.

Παράδειγμα μιας Βάσης (με Εννοιολογικό (Conceptual) Μοντέλο)



- **Mini-world παραδείγματος:**
 - Μέρος ενός Πανεπιστημίου (UNIVERSITY).
- **Μερικές Οντότητες (Entities) του mini-world :**
 - STUDENTs
 - COURSEs
 - SECTIONs (of COURSEs)
 - (academic) DEPARTMENTs
 - INSTRUCTORs

Παράδειγμα μιας Βάσης (με Εννοιολογικό (Conceptual) Μοντέλο)



- **Μερικές Σχέσεις (*relationships*) του mini-world :**
 - SECTIONs *are of specific* COURSEs
 - STUDENTs *take* SECTIONs
 - COURSEs *have prerequisite* COURSEs
 - INSTRUCTORs *teach* SECTIONs
 - COURSEs *are offered by* DEPARTMENTs
 - STUDENTs *major in* DEPARTMENTs
- **Σημείωση:** Οι πιο πάνω **οντότητες (*entities*)** και **συσχετίσεις (*relationships*)** τυπικά αναπαριστώνται σε ένα εννοιολογικό μοντέλο δεδομένων (π.χ., το Μοντέλο Οντοτήτων-Συσχετίσεων που θα δούμε αργότερα)

Παράδειγμα μιας Απλής Βάσης



COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	04	King
92	CS1310	Fall	04	Anderson
102	CS3320	Spring	05	Knuth
112	MATH2410	Fall	05	Chang
119	CS1310	Fall	05	Anderson
135	CS3380	Fall	05	Stone

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Βασικά Πλεονεκτήματα της Προσέγγισης των Βάσεων (σε αντίθεση με Αρχεία)



1) Αυτό-Περιγραφική Φύση μιας Βάσης:

- Μια DBMS περιέχει ένα **Κατάλογο (Catalog)** ο οποίος αποθηκεύει την περιγραφή μιας βάσης (π.χ., δομή πινάκων, τύποι δεδομένων, περιορισμοί, κτλ)
- Η πληροφορία ονομάζεται **meta-data**.
- Αυτό επιτρέπει την **πρόσβαση** στις πληροφορίες χωρίς να **γνωρίζουμε τη δομή της αποθηκευμένης πληροφορίας**.

Παράδειγμα ενός Απλοποιημένου Καταλόγου (Catalog)



RELATIONS

Relation_name	No_of_columns
STUDENT	4
COURSE	4
SECTION	5
GRADE_REPORT	3
PREREQUISITE	2

COLUMNS

Column_name	Data_type	Belongs_to_relation
Name	Character (30)	STUDENT
Student_number	Character (4)	STUDENT
Class	Integer (1)	STUDENT
Major	Major_type	STUDENT
Course_name	Character (10)	COURSE
Course_number	XXXXNNNN	COURSE
...
...
...
Prerequisite_number	XXXXNNNN	PREREQUISITE

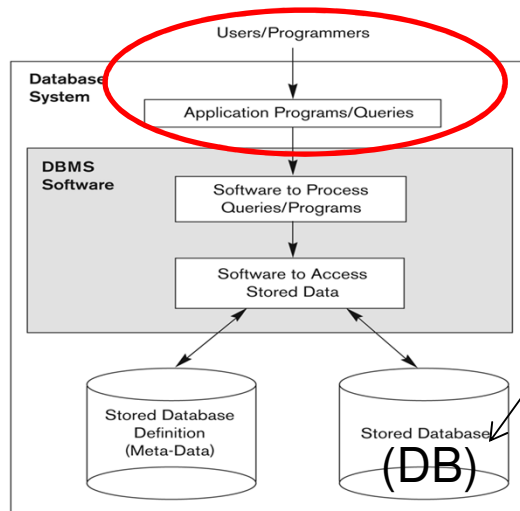
Note: Major_type is defined as an enumerated type with all known majors. XXXXNNNN is used to define a type with four alpha characters followed by four digits



Βασικά Πλεονεκτήματα της Προσέγγισης των Βάσεων (σε αντίθεση με Αρχεία)

2) Ανεξαρτησία Προγράμματος-Δεδομένων (Program-data independence)

- Επιτρέπει να **αλλάξουμε την δομή και φυσική οργάνωση** των δεδομένων χωρίς να χρειαστεί να αλλάξουμε **τα Προγράμματα** (DBMS Application Programs / Queries)



Π.χ., εάν η DB χωριστεί σε 2 επί μέρους DBs τότε αυτό δε χρειάζεται να το γνωρίζει ο χρήστης!

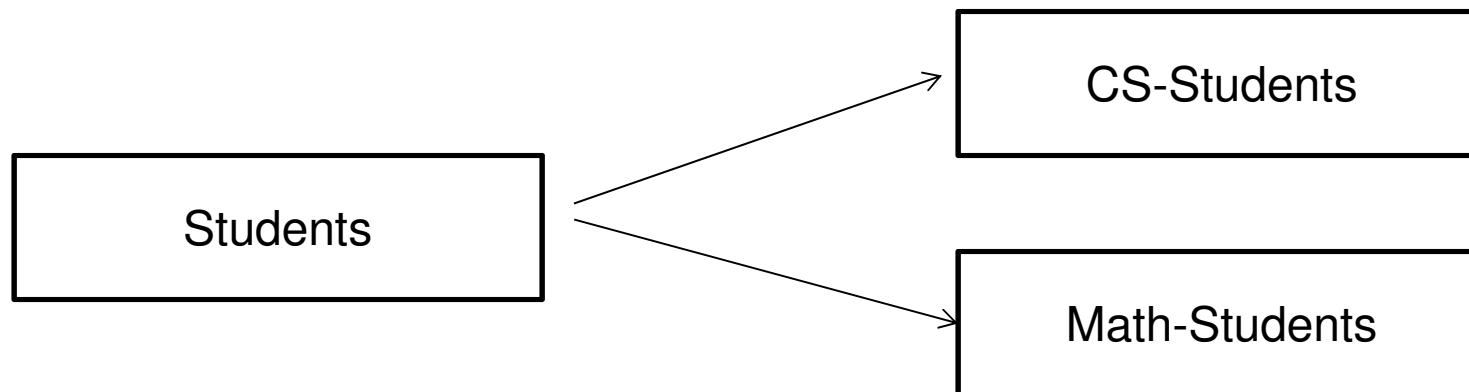
Αφαιρετικότητα (Data Abstraction)!

Βασικά Πλεονεκτήματα της Προσέγγισης των Βάσεων (σε αντίθεση με Αρχεία)



3) Παροχή Πολλαπλών Όψεων (Views) της ίδιας πληροφορίας.

- Π.χ., η γραμματεία της πληροφορικής βλέπει μόνο όσες εγγραφές φοιτητών **συσχετίζονται** με τις εγγραφές των φοιτητών του **Τμήματος Πληροφορικής**.



Βασικά Πλεονεκτήματα της Προσέγγισης των Βάσεων (σε αντίθεση με Αρχεία)



4) Διαμοιρασμός Δεδομένων (Μεταξύ Χρηστών) και Επεξεργασία Δοσοληψιών Πολλαπλών Χρηστών :

- Επιτρέπει σε **ταυτόχρονους χρήστες (concurrent users)** να ανακτούν και να ενημερώνουν την βάση.
- Το **Υποσύστημα Ταυτοχρονίας (Concurrency control subsystem)** της DBMS διασφαλίζει ότι τα δεδομένα θα παραμένουν σε **ορθή (correct)** και **συνεπή (consistent)** κατάσταση.
 - Δες παράδειγμα στην επόμενη διαφάνεια
- Το **Υποσύστημα Ανάκαμψης (Recovery subsystem)** διασφαλίζει ότι κάθε ολοκληρωμένη δοσοληψία (πρόγραμμα βάσης) θα καταγράψει μόνιμα τα αποτελέσματα της στη βάση δεδομένων

Υποσύστημα Ελέγχου Ταυτοχρονίας

(διασφαλίζει ότι η Εκτέλεση 2 δε θα συμβεί)



```

Deposit (amount, account#) {
    x = read(accounts[account#]);
    write(accounts[account#], amount + x);
}
    
```

Θεωρήστε:

Account[7] = €100

T1: Deposit1(100, 7)

T2: Deposit2(50, 7)

