



EPL342 –Databases

Lecture 1: Introduction I

Introduction to Databases

(Chapter 1: Elmasri-Navathe, 7ED)

**Demetris Zeinalipour
(Δημήτρης Ζεϊναλιπούρ)**

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

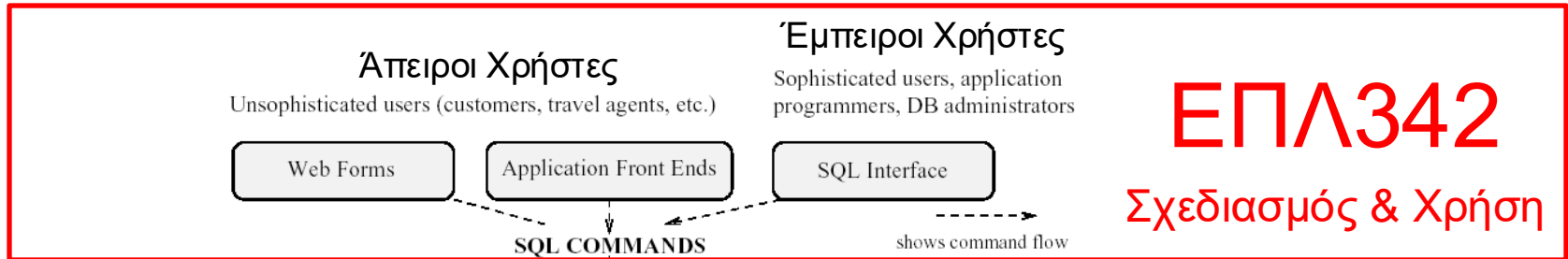


Στόχοι Μαθήματος

- **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/ΕΠΛ646: 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



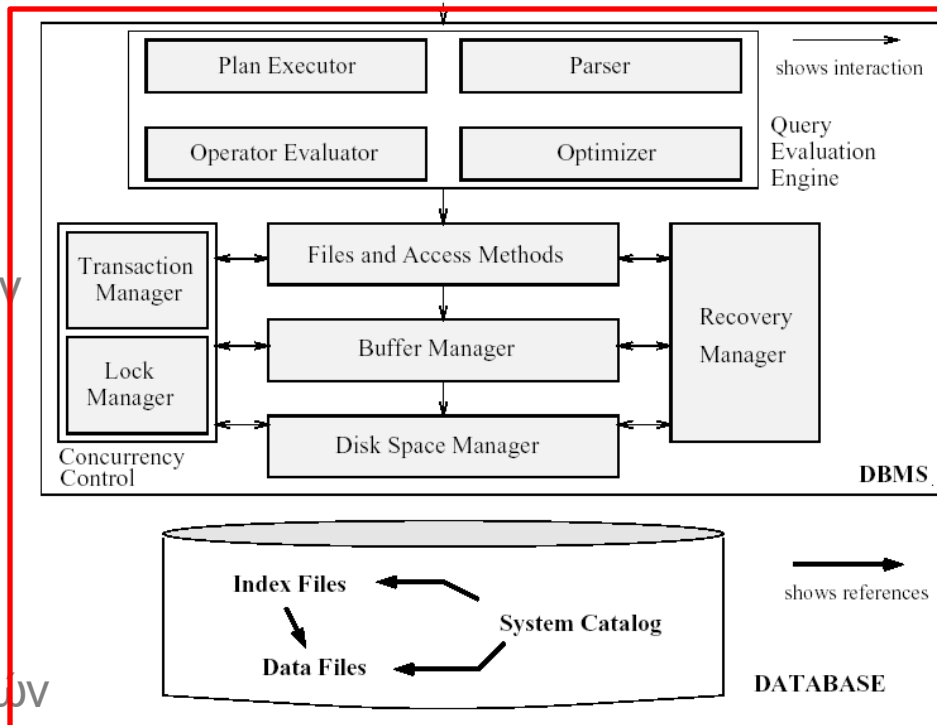
ΕΠΛ342

Σχεδιασμός & Χρήση

Δοσοληπιών και
Τεχνικές Ελέγχου
Ταυτοχρονίας

Δομή Ευρετηρίων
Δευτερεύουσας
Μνήμης (Hash,
B+)

Αποθήκευση
Δεδομένων και
Μετα-πληροφοριών



Βελτιστοποίησης
Επερωτήσεων

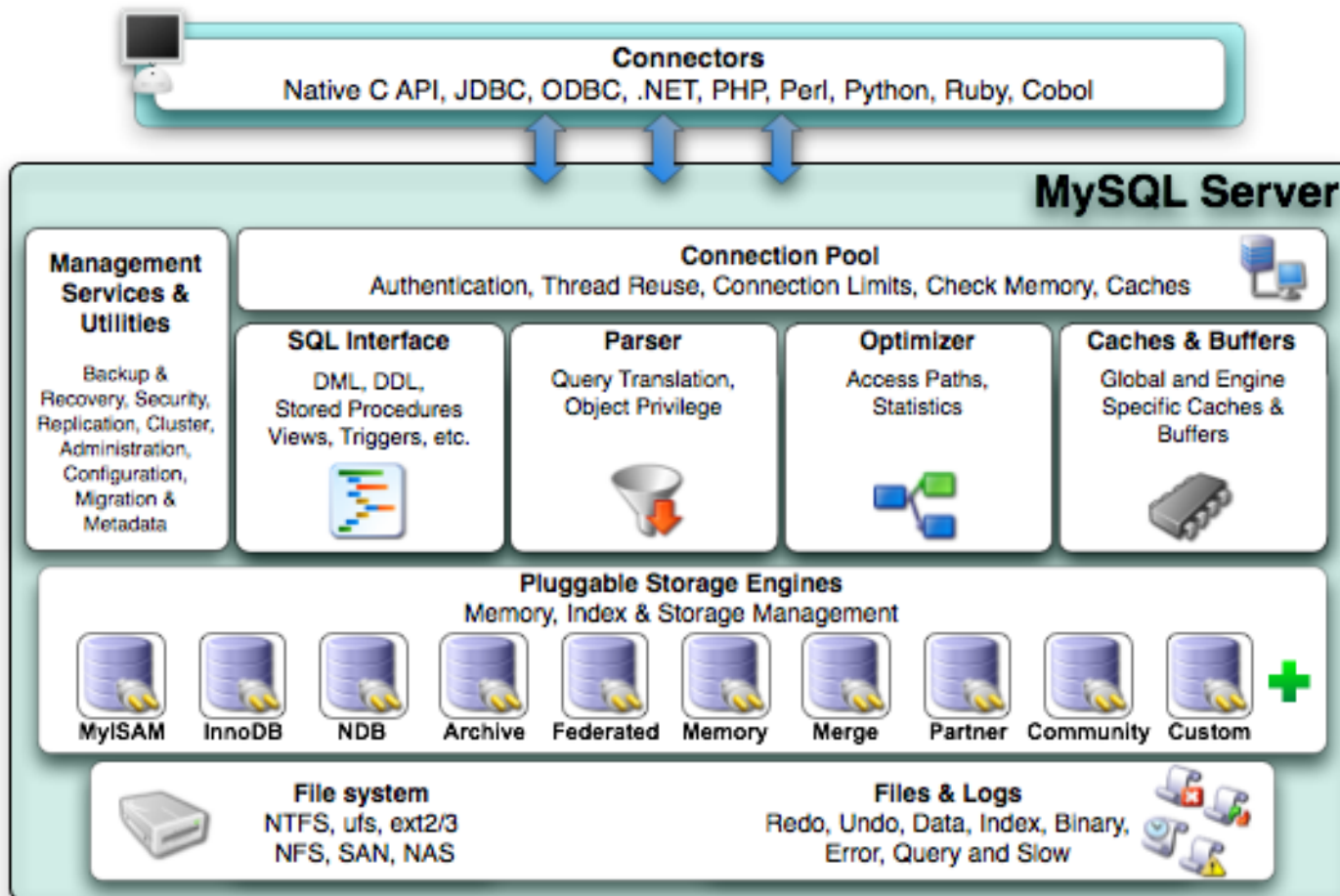
Τεχνικές
Ανάκαμψης (σε
περιπτώσεις
σφαλμάτων)

Ενδόμνημη
Διαχείριση
Δεδομένων

ΕΠΛ446/Ε

ΠΛ646¹⁻³

Τι είναι μια DBMS; Ανατομία του MySQL Server





Συμβόλαιο Μαθήματος

- **Επίπεδο:** Προπτυχιακό
 - Υποχρεωτικό Μάθημα
- **Πίστωση:** 7.5 μονάδες ECTS
- **Προαπαιτούμενα:**
 - ΕΠΛ231: Δομές Δεδομένων & Αλγόριθμοι
 - ΕΠΛ111: Διακριτές Δομές στην Πληροφορική και τον Υπολογισμό
 - ΕΠΛ133: Αντικειμενοστρεφής Προγραμματισμός
- **Μέθοδοι Διδασκαλίας**
 - Διαλέξεις (3 ώρες εβδομαδιαίως): Παράδοση Διδασ. Ύλης
 - Φροντιστήριο (1 ώρα εβδομαδιαίως): Θεωρητική Εξάσκηση
 - Εργαστήριο (1.5 ώρες εβδομαδιαίως): Πρακτική Εξάσκηση



Συμβόλαιο Μαθήματος

- **Αξιολόγηση**

- 50% Τελική Εξέταση (1)

- 25% Ενδιάμεση Εξέταση (1)

- **Ημερ.: Τρίτη, 22 Οκτωβρίου 2024! (8^η Εβδ.)**

- *** Εναλλακτική Ημερ. Παρασκευή, 25 Οκτωβρίου 2024 (7:30-9:00)! (8^η Εβδ.)**

- 25% Ασκήσεις ή/και Quizzes

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

- 15% Project Μαθήματος σε Εμπορική Βάση Δεδομένων.



Βιβλιογραφία

Βασική Βιβλιογραφία

A1) [Fundamentals of Database Systems](#), 7/E Ramez Elmasri, Shamkant B. Navathe, ISBN-10: 1-292-09761-2, ISBN-13: 978-1-292-09761-9, Pearson Press, 2017.

Η' Ελληνική Μετάφραση

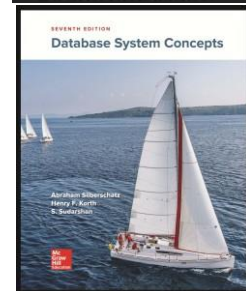
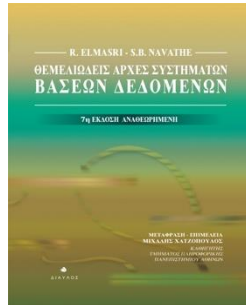
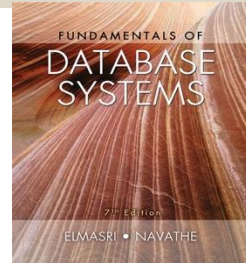
A2) [Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων](#) - 7η έκδοση (1ος τόμος), γραμμένο από τους Ramez Elmasri και Sham B. Navathe (μετάφραση Μ. Χατζόπουλος), Δίαυλος, 978-960-531-343-2, 2016.

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](#), 7th Edition, Abraham Silberschatz, Henry F. Korth, S. Sudarshan Hardcover, 1168 pages, McGraw-Hill Publisher, ISBN 9780078022159, 2017.



Ιστοσελίδα ΕΠΛ342



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

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

The screenshot shows the course page for EPL342: Databases. The page header includes the University of Cyprus logo and navigation links for EPL342, News, Schedule, Labs, Assignments, Moodle, and TA. The breadcrumb trail is 'Zeinalipour > Courses > EPL342'. The course title is 'EPL342: Databases'. The instructor is Demetris Zeinalipour. The course type is 'Upper-division Undergraduate (Compulsory)'. The prerequisite is 'EPL231 - Data Structures and Algorithms (or equivalent)'. The course is held on Tuesdays and Fridays. There are two lecture sections: #1 (MAM: 10184) from 09:00 to 10:15 in room XΩΔ002-#B105, and #2 (MAM: 16405) from 10:30 to 11:45 in room XΩΔ002-#119. There are two recitation sections on Wednesdays: #1 (MAM: 10184) from 09:00 to 09:50 in room XΩΔ02-009, and #2 (MAM: 16405) from 10:00 to 10:50 in room XΩΔ01-102. The laboratory is on Fridays from 10:30 to 18:00, with 1 1/2 hours each, in room GEE01-B121. The teaching assistant is Christoforos Panayiotou. A COVID-19 notice states that instruction will be carried out both physically and online in a hybrid manner, with two groups for physical classes and one group for online instruction. More information on the rotation schedule and teleconferencing connection URL is provided in the email. The overview section is partially visible at the bottom.

University of Cyprus

EPL342 News Schedule Labs Assignments Moodle TA

Zeinalipour > Courses > EPL342

EPL342: Databases

Instructor: Demetris Zeinalipour

Type: Upper-division Undergraduate (Compulsory)

Prerequisite: [EPL231 - Data Structures and Algorithms](#) (or equivalent)

When: Tuesday and Friday,
#1 (MAM: 10184): 09:00 – 10:15, [XΩΔ002-#B105](#)
#2 (MAM: 16405): 10:30 – 11:45, [XΩΔ002-#119](#)

Recitation: Wednesday,
#1 (MAM: 10184): 09:00 – 09:50, [XΩΔ02-009](#)
#2 (MAM: 16405): 10:00 – 10:50, [XΩΔ01-102](#)

Laboratory: Friday, 10:30-18:00 (physical 1 1/2 hours each / check your Banner slot) in [GEE01-B121](#)

Teaching Assistant: Christoforos Panayiotou

COVID-19 Notice: According to the decisions for teaching at UCY for Fall 2020, instruction will be carried out this semester both physically and online (i.e., in a hybrid manner). Each of the two physical classes of EPL342 will be divided in two groups. The one group will follow instruction physically while the other one will be connected online through a URL that is circulated to students. More information on the rotation schedule for physical attendance as well as the teleconferencing connection URL has been circulated to students to their UCY email (based on Banner subscriptions). Please follow the instructions in the email.

Overview

The main objective of this undergraduate course is to provide an in-depth understanding of concepts related to the design and

EPL342 Moodle



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

The screenshot shows the Moodle interface for the course 'EPL342 - Databases'. The page title is 'EPL342: Databases (ΕΠΛ342: Βάσεις Δεδομένων)' and the instructor is 'Demetris Zeinalipour'. The page is divided into three main sections: 'General', 'AS1', and 'AS2'. The 'General' section contains an 'Important Notices' section with a warning to check the section regularly for updates. The 'AS1' section is currently hidden from students. The 'AS2' section is also hidden. The right sidebar contains a calendar for August 2019, an 'EVENTS KEY', and a 'LATEST ANNOUNCEMENTS' section. The bottom of the page features a 'CSCOLLOQUIUMSERIESUCY' section with a colloquium announcement.

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



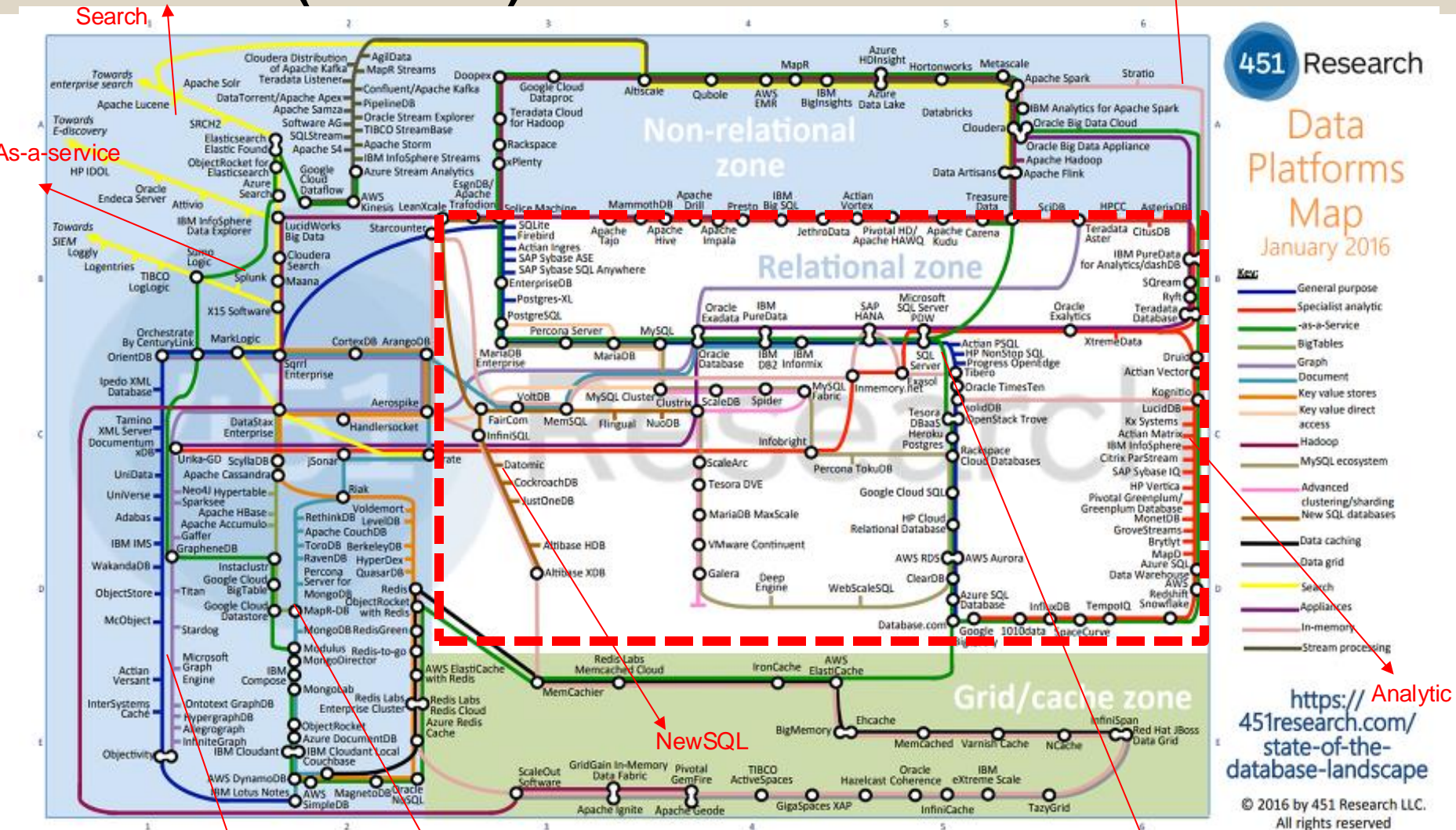
- Αεροπορικές κρατήσεις (expedia.com)
- Νοσοκομεία, Χημεία
- Τραπεζικές Εφαρμογές (WellsFargo.com)
- Λογιστήρια
- Σύστημα εγγραφής φοιτητών
- Κοινωνικές Ασφαλίσεις, Φόροι
- Πωλήσεις – Αποθήκες (Amazon.com, eBay)
- Πιστωτικές κάρτες
- Χρηματιστήριο
- Τηλεπικοινωνίες
- ... και άλλες πολλές!

Τύποι Βάσεων Δεδομένων



- **Παραδοσιακές Εφαρμογές (OLTP – Transaction Processing):**
 - 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:**
 - Αποθήκες Δεδομένων (OLAP Analytic Data Warehouses)
 - Βάσεις Πολυμέσων και Γεωγραφικών Πληροφ. Συστημάτων (GIS)
 - Βάσεις Ροών (Stream Databases), Βάσεις Νέφους (Cloud Databases), Βάσεις XML, κτλ
 - **Δες διάγραμμα στην επόμενη διαφάνεια για πιο διευρυμένη εικόνα.**

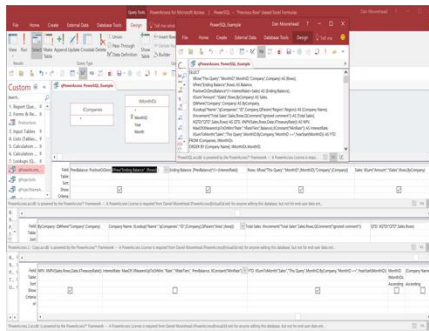
Το Τοπίο των Βάσεων Δεδομένων (2016) – 451 Research



Γρήγορη Γνωριμία με την απλούστερη DBMS στο Lab 1



LAB 1



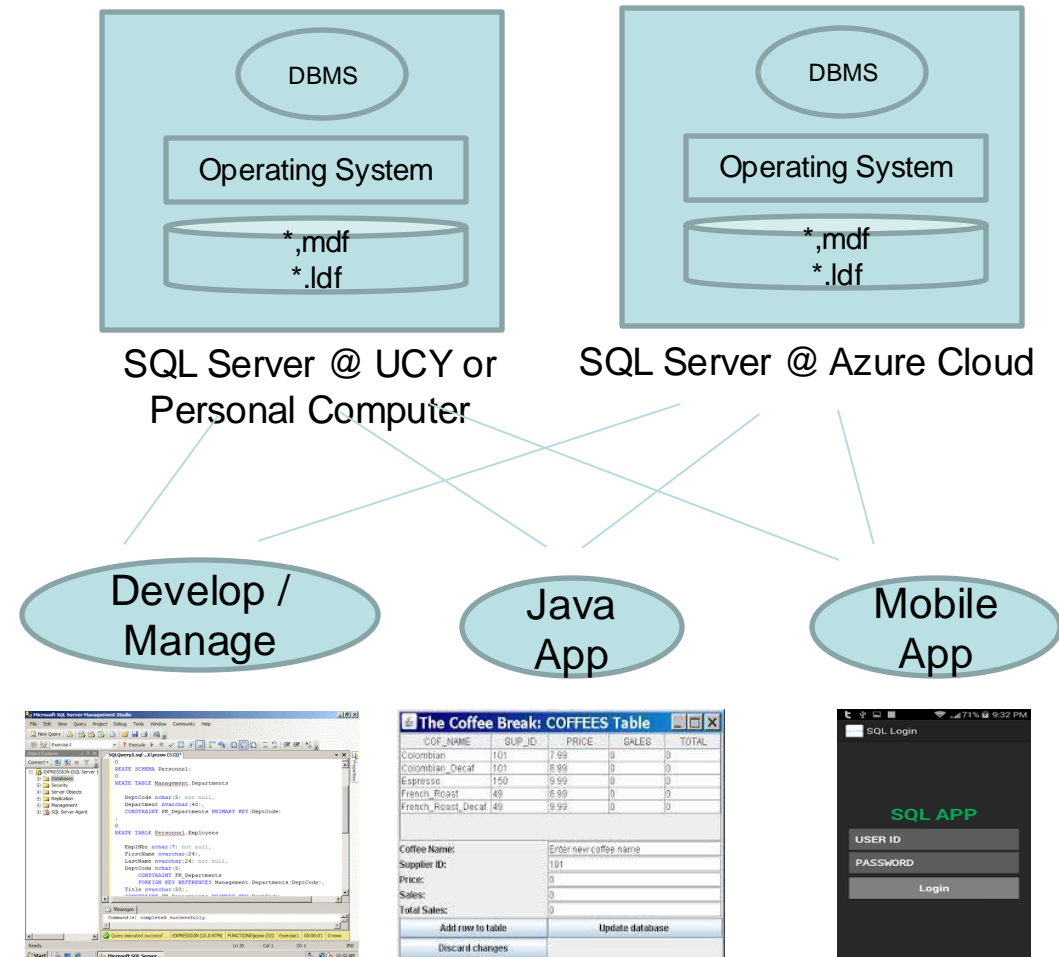
DBMS + App (service-less)

Operating System

*,accdb
*.mdb

MS Access on Lab Machine

Rest Course



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 συσχετιζόμενα δεδομένα 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.**

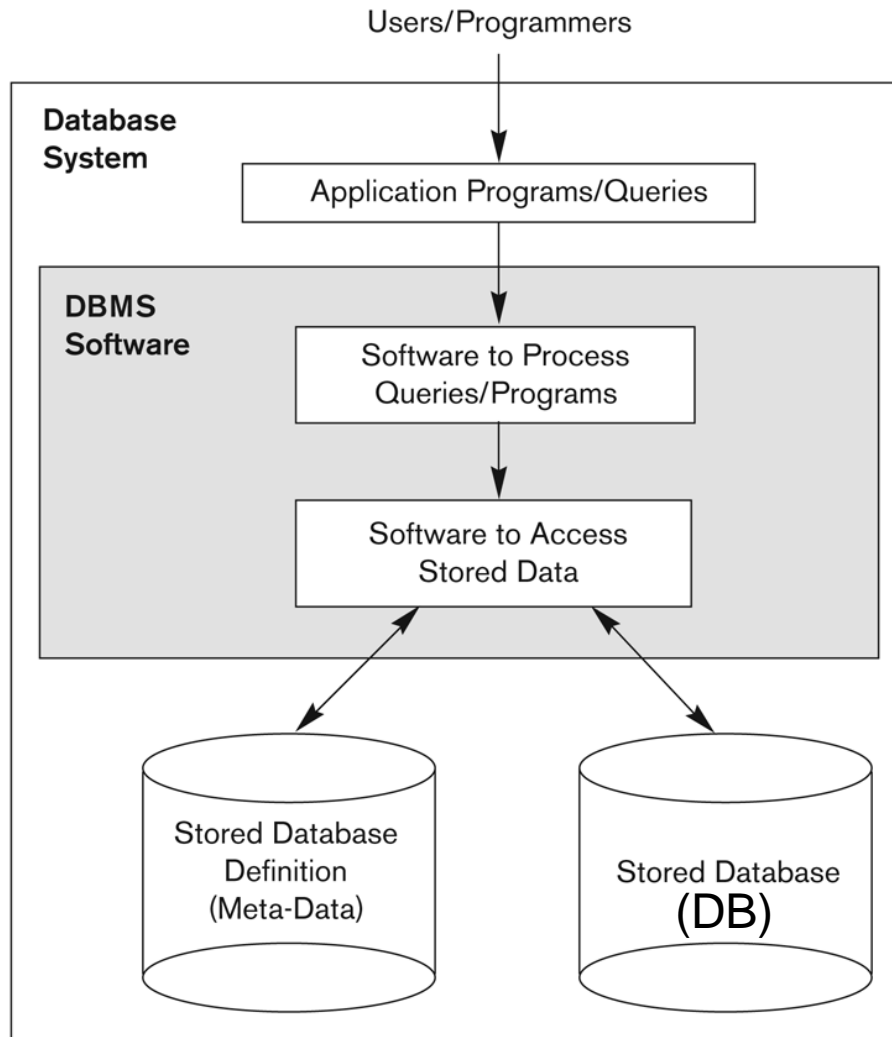
- Π.χ., Οι εγγραφές των φοιτητών ενός Πανεπιστημίου

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.

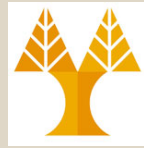
Παράδειγμα μιας Βάσης

(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

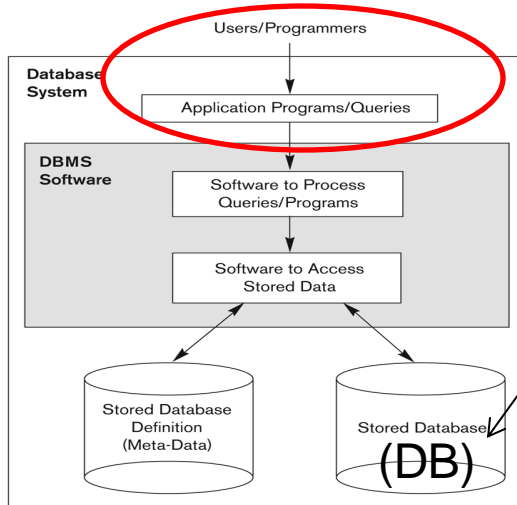
Note: Major_type is defined as an enumerated type with all known majors. XXXXNNNN is used to define a type with four alpha characters followed by four digits

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!

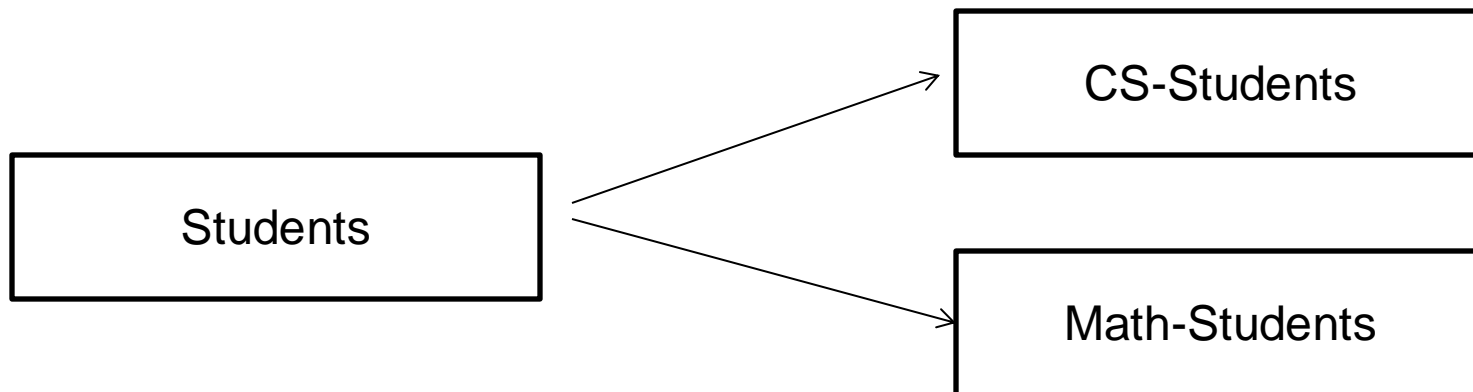
Αφαιρετικότητα (Data Abstraction)!

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)

