

# ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Χειμερινό Εξάμηνο 2014-2015



## ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Δρ. Βαγγελιώ Καβακλή

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ,  
ΤΜΗΜΑ ΠΟΛΙΤΙΣΜΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΑΣ

1

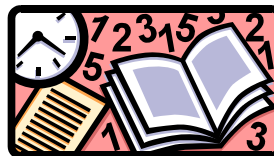
## Στόχοι του μαθήματος

- Η εισαγωγή στις θεμελιώδεις έννοιες που είναι απαραίτητες για το διαχείριση βάσεων δεδομένων
- Στο τέλος του εξαμήνου θα είστε ικανοί να
  - σχεδιάζετε μια Β.Δ.
    - να ορίζετε τη μορφή των δεδομένων που θα αποθηκευτούν στη βάση
  - υλοποιείτε μια Β.Δ.
    - να δημιουργείτε μια Β.Δ. χρησιμοποιώντας ένα σύστημα διαχείρισης Β.Δ.
  - χειρίζεστε μια Β.Δ.
    - να αποθηκεύετε δεδομένα και να υποβάλετε ερωτήσεις προς τη βάση για ανάκτηση συγκεκριμένων δεδομένων



2

# Περιγραφή Μαθήματος



- Διδασκαλία: Πέμπτη 09:00 - 11:00  
Αίθουσα ΣΤ, κτίριο Θάλασσας
- Εργαστήρια: Παρασκευή (Εργαστήριο Ανθρωπογεωγραφίας)  
**Ομάδα Α:** 09:00 - 11:00  
**Ομάδα Β:** 11:00 - 13:00
- Διδακτικά Βιβλία
  - **Σχισιακές Βάσεις Δεδομένων**, Ε. Κεχρής, Εκδόσεις Κριτική, 2005
  - **Βάσεις Δεδομένων και SQL: Μια πρακτική προσέγγιση**, Α. Σταυρακούδης, Εκδόσεις Κλειδάριθμος, 2010
  - **Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων**, R. Elmasri – S. B. Navathe, 6η Έκδοση, Εκδόσεις Δίαυλος, 2012
- Αξιολόγηση
  - Ο τελικός βαθμός υπολογίζεται ως εξής:  
**10% Εργαστηριακές Ασκήσεις + 25% Ασκήσεις + 65% Γραπτή Εξέταση**
  - *Οι φοιτητές από παλιότερα έτη μπορούν να αξιολογηθούν αποκλειστικά βάσει της Γραπτής Εξέτασης.*

3

## Παραδείγματα Συστημάτων Β.Δ.

- Παραδοσιακές Εφαρμογές
  - οι αποθηκευμένες πληροφορίες είναι σε μορφή text ή αριθμών
    - Ενημέρωση τραπεζικού λογαριασμού
    - Κράτηση ξενοδοχείου ή αεροπορικού ταξιδιού
    - Αναζήτηση βιβλιογραφικών στοιχείων σε κατάλογο βιβλιοθήκης
    - Συνδρομή περιοδικού σ' έναν εκδότη
    - Αγορά αντικειμένων
- Πολυμεσικά Συστήματα Β.Δ.
  - αποθηκεύουν και διαχειρίζονται εικόνες / video / ήχο
- Γεωγραφικά Πληροφοριακά Συστήματα (G.I.S.)
  - αποθηκεύουν δεδομένα καιρού / δορυφορικές εικόνες, γεωγραφικά αντικείμενα (2 διαστάσεων), χρονικές πληροφορίες
- Αποθήκη Δεδομένων και On-line Συστήματα Αναλυτικής Επεξεργασίας (OLAP)
  - εξαγωγή και ανάλυση χρησικών πληροφοριών για λήψη αποφάσεων

4

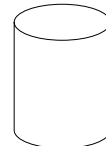
# Περιεχόμενο του μαθήματος

- Σχεδιασμός Βάσεων Δεδομένων
  - Εννοιολογικός Σχεδιασμός
  - Μοντέλο Οντοτήτων – Συσχετίσεων (ΟΣ)
- Υλοποίηση Βάσεων Δεδομένων
  - Σχεσιακές Βάσεις Δεδομένων
  - Σχεσιακή Άλγεβρα
  - Η γλώσσα χειρισμού δεδομένων SQL
- Αποθήκευση Δεδομένων
  - Φυσικές Δομές Αρχείων
  - Μέθοδοι Προσπέλασης
- Συστήματα Διαχείρισης Βάσεων Δεδομένων



## Ορισμοί

- **Βάση Δεδομένων** (database)
  - Συλλογή από συσχετιζόμενα δεδομένα
- **Δεδομένα:**
  - γνωστά γεγονότα που μπορούν να καταγραφούν και έχουν κάποια υπονοούμενη σημασία
    - ονόματα, αρ. τηλεφώνων, διευθύνσεις των ανθρώπων που γνωρίζετε, τα οποία έχετε καταγράψει σε ένα ευρετήριο διευθύνσεων
- **Ιδιότητες Β.Δ.**
  - αναπαριστά κάποια άποψη του πραγματικού κόσμου (μικρόκοσμος ή πεδίο αναφοράς)
  - αποτελεί μια λογικά συνεκτική συλλογή δεδομένων που έχει κάποια εγγενή σημασία
  - σχεδιάζεται, χτίζεται και γεμίζει με δεδομένα για κάποιο συγκεκριμένο σκοπό
  - προορίζεται για συγκεκριμένη ομάδα χρηστών



## Μέγεθος Β.Δ.

- Μια Β.Δ. μπορεί να έχει οποιοδήποτε μέγεθος και κυμαινόμενη πολυπλοκότητα
  - λίστα ονομάτων / διευθύνσεων:
    - λίγες εκατοντάδες εγγραφές
  - κατάλογος βιβλιοθήκης:
    - 0,5 εκ. κάρτες αποθηκευμένες υπό διαφορετικές κατηγορίες (όνομα συγγραφέα, θέμα βιβλίου, τίτλο βιβλίου) ταξινομημένες κατά αλφαβητική σειρά σε κάθε κατηγορία
  - Β.Δ. Εφορείας ΗΠΑ:
    - 100 εκ. φορολογούμενοι × 5 φόρμες/ φορολογούμενο × 200 χαρακτήρες / δήλωση × 4 τελευταίες δηλώσεις =  $4 \times 10^{11}$  bytes

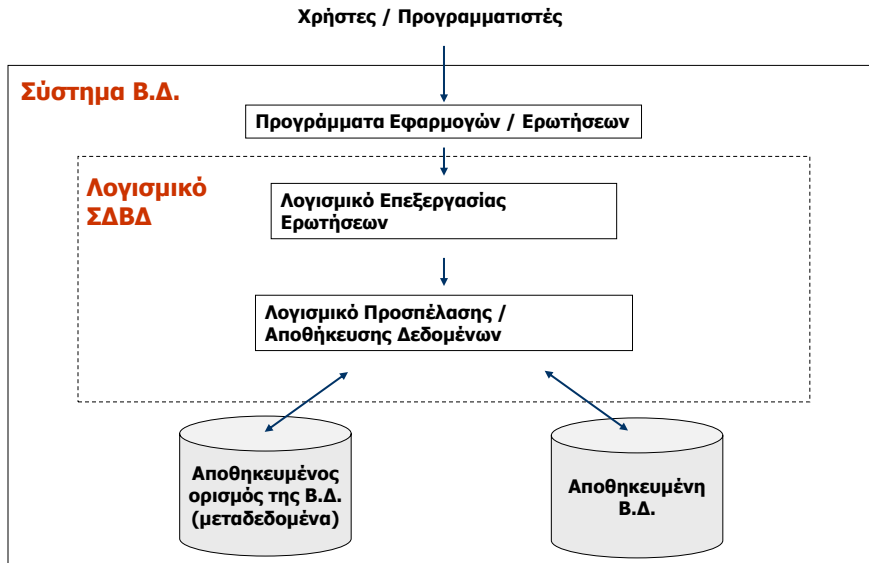
7

## Σύστημα Βάσης Δεδομένων

- Μια Β.Δ. μπορεί να δημιουργηθεί και να συντηρηθεί είτε χειρόγραφα, είτε με χρήση μηχανών
- **Σύστημα Διαχείρισης Βάσεων Δεδομένων - ΣΔΒΔ** (database management system - DBMS)
  - μια συλλογή από προγράμματα που επιτρέπουν στους χρήστες να δημιουργήσουν και να συντηρήσουν μια βάση δεδομένων
  - Ένα ΣΔΒΔ μπορεί να είναι:
    - **γενικής χρήσης** (general purpose): διευκολύνει τις διαδικασίες ορισμού, κατασκευής και χειρισμού Β.Δ. για διάφορες εφαρμογές
    - **ειδικού σκοπού** (special purpose): υλοποιεί μια συγκεκριμένη Β.Δ.
- **Σύστημα Βάσης Δεδομένων (database system):**
  - μια Β.Δ. μαζί με το αντίστοιχο λογισμικό ΣΔΒΔ

8

# Περιβάλλον Συστήματος Β.Δ.



9

## Διαχείριση Β.Δ.

- Περιλαμβάνει τις εξής διαδικασίες:
  - ορισμός: προδιαγραφή των τύπων των δομών και των περιορισμών των δεδομένων που θα αποθηκευτούν στη βάση
  - κατασκευή: αποθήκευση των δεδομένων σ' ένα αποθηκευτικό μέσο που ελέγχεται από το ΣΔΒΔ
  - χειρισμός: περιλαμβάνει λειτουργίες όπως:
    - υποβολή ερωτήσεων προς τη βάση για ανάκτηση συγκεκριμένων δεδομένων
    - ενημέρωση της βάσης ώστε να αντανakλά αλλαγές στο μικρόκοσμο
    - παραγωγή αναφορών από τα δεδομένα

10

## Παράδειγμα: ΠΑΝΕΠΙΣΤΗΜΙΟ

- Στόχος: Διατήρηση πληροφοριών που αφορούν
  - φοιτητές
  - μαθήματα
  - βαθμολογίες
- Οργάνωση Β.Δ. σε 4 αρχεία εγγραφών
  - ΦΟΙΤΗΤΗΣ
  - ΜΑΘΗΜΑΤΑ
  - ΔΙΔΑΣΚΑΛΙΑ
  - ΒΑΘΜΟΛΟΓΙΑ

11

ΦΟΙΤΗΤΗΣ	Όνομα	Αριθ_Μητρώου	Έτος	Κατεύθυνση
	Παπαδόπουλος	2000/0001	3	Μουσειολογία
	Δημητρέλια	2001/0005	3	Εκπ Τεχνολογία

ΜΑΘΗΜΑ	Όνομα_Μαθήματος	Κωδικός_Μαθ	Διδακ_μονάδες	Τμήμα
	Διαδραστικά Πολυμέσα	HY101	3	ΤΠΤΕ
	Δικτυακά Πολυμέσα Ι	HY201	3	ΤΠΤΕ
	Φύλο και Πολιτισμός	ΔΙΑ300	2	ΠΕΡΙΒΑΛΟΝΤΟΣ

ΔΙΔΑΣΚΑΛΙΑ	Κωδ_Διδασκ	Κωδικός_Μαθ	Εξάμηνο	Έτος	Διδάσκων
	11	HY101	Χειμερινό	2000	Καβακλή
	12	HY101	Χειμερινό	2001	Ρόκου
	21	HY201	Εαρινό	2001	Καβακλή
	22	HY201	Χειμερινό	2002	Δημαράκη

ΒΑΘΜΟΛΟΓΙΑ	Αριθ_Μητρώου	Κωδ_Διδασκ	Βαθμός
	2000/0001	11	7.5
	2000/0001	21	6.5
	2001/0005	22	8

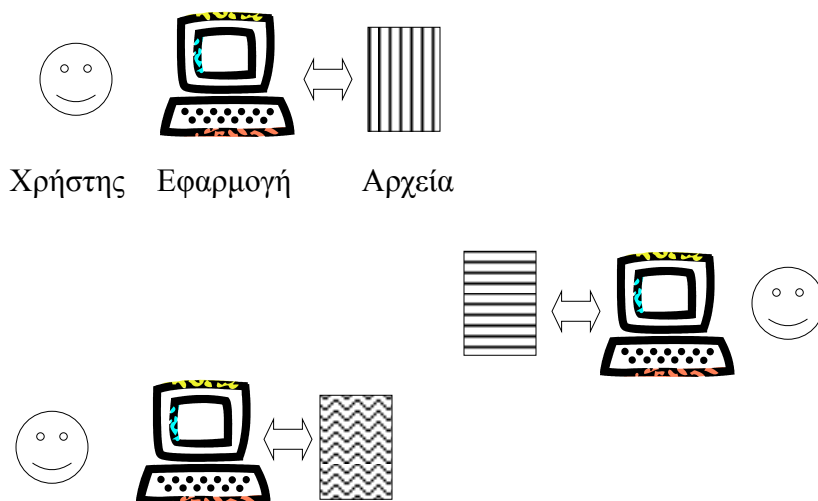
12

## Διαχείριση της Β.Δ. ΠΑΝΕΠΙΣΤΗΜΙΟ

- Ορισμός της Β.Δ.
  - προσδιορισμός των **στοιχειωδών δεδομένων** που πρόκειται να αποθηκευτούν σε κάθε εγγραφή
  - προσδιορισμός του **τύπου δεδομένων** για κάθε στοιχειώδες δεδομένο
- Κατασκευή
  - αποθήκευση δεδομένων στο αντίστοιχο αρχείο
  - οι εγγραφές στα διάφορα αρχεία μπορεί να σχετίζονται η μια με την άλλη!
- Χειρισμός
  - διενέργεια επερωτήσεων και ενημερώσεων με τη χρήση κάποιας «γλώσσας» του ΣΒΔ
    - «βρες τα ονόματα και τις βαθμολογίες των φοιτητών που παρακολούθησαν τη διδασκαλία του μαθήματος Δικτυακά Πολυμέσα Ι κατά το Εαρινό εξάμηνο 1992»
    - «δημιούργησε μια νέα διδασκαλία για το μάθημα Δικτυακά Πολυμέσα Ι γι' αυτό το εξάμηνο»
    - «καταχώρησε το βαθμό 10 στον Παπαδόπουλο για τη διδασκαλία του μαθήματος Φύλο και Πολιτισμός του προηγούμενου εξαμήνου»

13

## Σύστημα διαχείρισης αρχείων (file system)



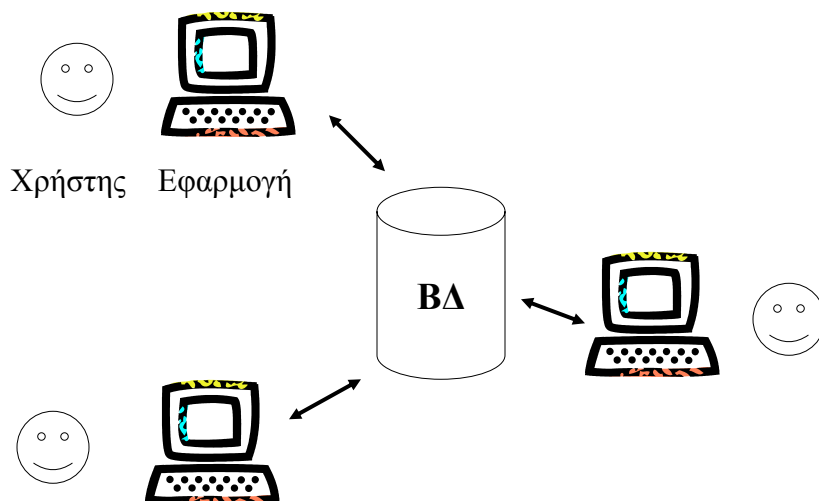
14

## Μειονεκτήματα συστήματος διαχείρισης αρχείων

- Πλεονασμός δεδομένων (data redundancy)
- Ασυνέπεια δεδομένων (inconsistency)
- Δύσκολη αναζήτηση (search)
- Δυσκολία διαμοιρασμού (data sharing)
- Αδυναμία ταυτόχρονης πρόσβασης (concurrent access)
- Ανομοιομορφία (heterogeneity)
- Δυσκολία στην εφαρμογή κανόνων ακεραιότητας (integrity rules)
- Δυσκολία στην τήρηση αντιγράφων ασφαλείας (backup)

15

## Προσέγγιση Βάσεων Δεδομένων



16



## Πλεονεκτήματα Β.Δ.

- Ταυτόχρονη Προσπέλαση από πολλούς χρήστες
  - το ΣΔΒΔ περιλαμβάνει λογισμικό για έλεγχο ταυτόχρονης προσπέλασης (concurrency control)
  - εξασφαλίζει σωστό αποτέλεσμα ενημέρωσης
- Έλεγχος Πλεονασμών
  - οικονομία χώρου και μόχθου
  - συμβατότητα μεταξύ δεδομένων
- Περιορισμός μη εξουσιοδοτημένης προσπέλασης
  - το ΣΔΒΔ περιλαμβάνει λογισμικό ασφάλειας και εξουσιοδότησης (security & authorisation)
  - διαφορετικοί περιορισμοί προδιαγράφονται για διαφορετικούς χρήστες
- Παροχή μηχανισμών τήρησης εφεδρικών αντιγράφων και ανάκαμψης (backup)
  - αντιμετώπιση (ανάκαμψη) μετά από βλάβες υλικού ή λογισμικού

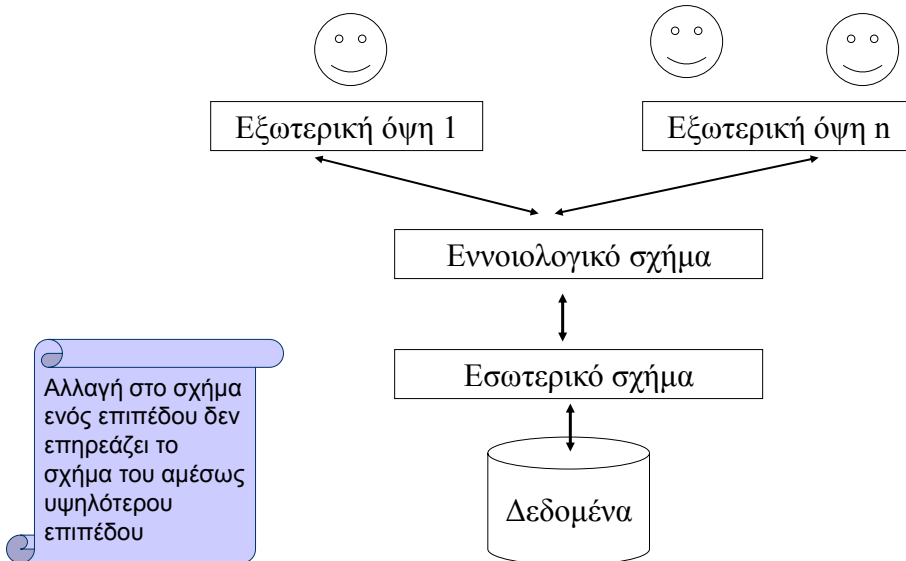
17

## Χρήσιμα χαρακτηριστικά της προσέγγισης Β.Δ.

- Αυτοπεριγραφή
  - ένα ΣΔΒ δεν περιέχει μόνο την ίδια τη βάση, αλλά και τον πλήρη ορισμό της (κατάλογος συστήματος)
  - επιτρέπει το λογισμικό του ΣΔΒΔ να είναι ανεξάρτητο από συγκεκριμένη εφαρμογή
- Ανεξαρτησία Προγραμμάτων - Δεδομένων
  - Η δομή των αρχείων δεδομένων αποθηκεύεται ξεχωριστά από το λογισμικό προσπέλασης
  - αλλαγή στη δομή των αρχείων δεν επηρεάζει το λογισμικό του ΣΔΒΔ
- Υποστήριξη Πολλαπλών Όψεων των Δεδομένων
  - διαφορετικοί χρήστες απαιτούν διαφορετική προοπτική (όψη) της Β.Δ.
  - μια όψη μπορεί να είναι ένα υποσύνολο της Β.Δ. ή να περιέχει εικονικά (virtual) δεδομένα που παράγονται από τα αρχεία αλλά δεν αποθηκεύονται

18

# Ανεξαρτησία Δεδομένων



19

## Παράδειγμα

### Όψη ΒΑΘΜΟΛΟΓΙΑ Φοιτητών

ΒΑΘΜΟΛΟΓΙΑ	Όνομα_Φοιτητή	Αντίγραφο_Φοιτητή				
		Κωδικός_Μαθ	Βαθμός	Εξάμηνο	Έτος	Κωδικός_Διδ
	Παπαδόπουλος	HY101	7.5	Χειμερινό	2000	11
	Παπαδόπουλος	HY201	6.5	Εαρινό	2001	21
	Δημητρέλια	HY201	8	Χειμερινό	2002	22

### Εννοιολογική Αναπαράσταση ΦΟΙΤΗΤΗ

ΦΟΙΤΗΤΗΣ	Όνομα	Αριθ_Μητρώου	Έτος	Κατεύθυνση
	Παπαδόπουλος	2000/0001	3	Μουσειολογία
	Δημητρέλια	2001/0005	3	Εκπ Τεχνολογία

### Εσωτερική Αναπαράσταση ΦΟΙΤΗΤΗ

όνομα στοιχειώδους δεδομένου	θέση αρχής της εγγραφής	μήκος σε χαρακτήρες (bytes)
Όνομα	1	30
Αριθ_Μητρώου	31	4
Έτος	35	4
Κατεύθυνση	39	4

20

## Εργαζόμενοι σ' ένα περιβάλλον Β.Δ.



- **Διαχειριστής Β.Δ.**
  - εξουσιοδοτεί προσπέλαση στη Β.Δ.
  - υπεύθυνος για προβλήματα ασφάλειας, απόδοσης του συστήματος
  - διαχείριση πόρων
- **Σχεδιαστές Β.Δ.**
  - υπεύθυνοι για τον προσδιορισμό δεδομένων
  - αναπτύσσουν τις διαφορετικές όψεις της βάσης που καλύπτουν τις απαιτήσεις διαφορετικών χρηστών
- **Τελικοί χρήστες**
  - άτομα που η εργασία τους απαιτεί προσπέλαση δεδομένων
- **Αναλυτές Συστημάτων και Προγραμματιστές**
  - αναπτύσσουν προδιαγραφές και υλοποιούν προκαθορισμένες δοσοληψίες βάσει των απαιτήσεων των τελικών χρηστών

21

## Συνέπειες Προσέγγισης Β.Δ.



- **Επιβολή τυποποίησης**
  - πρότυπα για τα ονόματα και τη μορφή των δεδομένων
  - καλύτερη συνεργασία
- **Μείωση χρόνου ανάπτυξης εφαρμογών**
  - απαιτείται γύρω στο  $\frac{1}{4}$  -  $\frac{1}{6}$  του χρόνου που απαιτείται με τη χρήση παραδοσιακών συστημάτων αρχείων
- **Ευελιξία**
  - επιτρέπονται αλλαγές στη δομή της Β.Δ. χωρίς να επηρεάζονται τα αποθηκευμένα δεδομένα ή υφιστάμενα προγράμματα εφαρμογών
- **Διαθεσιμότητα ενημερωμένων πληροφοριών**
  - ταυτόχρονα σε όλους τους χρήστες
- **Οικονομία κλίμακας**
  - ελαττώνεται το συνολικό κόστος λειτουργίας και διαχείρισης (δυνατός κεντρικός εξοπλισμός αντί για επιμέρους ανεξάρτητα εξοπλισμό χαμηλότερης απόδοσης)

22

## Μειονεκτήματα Προσέγγισης Β.Δ.

- Δαπανηρή προμήθεια και συντήρηση
- Απαιτήσεις σε υλικό και προσωπικό
- Καθυστέρηση στη διαχείριση δεδομένων



- Τα αρχεία εγγραφών είναι προτιμότερα εάν:
  - δεν αναμένονται αλλαγές στις εφαρμογές και στη δομή της Β.Δ.
  - υπάρχουν πιεστικές απαιτήσεις λειτουργίας σε πραγματικό χρόνο
  - δεν απαιτείται προσπέλαση πολλών χρηστών στα δεδομένα



23