

---

# Προηγμένα Θέματα Βάσεων Δεδομένων

Διδάσκων: Άγγελος Μιχάλας  
amichalas@uowm.gr

# Λίστα Συγγραμμάτων

---

- **Συστήματα Βάσεων Δεδομένων**, 6<sup>η</sup> Έκδοση, Abraham Silberschatz, Henry F. Korth, S. Sudarshan
- **Θεμελιώδεις αρχές συστημάτων βάσεων δεδομένων**, 5<sup>η</sup> Έκδοση, *Elmasri Ramez, Navathe Shamkant B.*
- **Συστήματα Διαχείρισης Βάσεων Δεδομένων**, 3<sup>η</sup> Έκδοση, Ramakrishnan Raghu, Gehrke Joahannes
- Άλλα βιβλία του χώρου:
  - H. Garcia-Molina, J.D. Ullman, J.D. Widom, "*Database Systems: The Complete Book*", Prentice Hall, 2001.
  - C.J. Date, *Introduction to Database Systems*, Addison Wesley, 2003.

# Περιεχόμενα Μαθήματος

---

## Στόχοι:

- Η κατανόηση των μεθόδων που χρησιμοποιούνται από τα συστήματα βάσεων δεδομένων για την υποστήριξη των λειτουργιών τους.
- Η κατανόηση σύνθετων λειτουργιών που εκτελούνται σε σχεσιακά συστήματα βάσεων δεδομένων
- Η κατανόηση του τρόπου λειτουργίας των βάσεων δεδομένων εγγράφων και NoSQL βάσεων δεδομένων

# Περιεχόμενα Μαθήματος

---

Πιο συγκεκριμένα θα μελετήσουμε:

- Φυσική οργάνωση
  - Τύποι εγγραφών
  - Τύποι σελίδων
  - Ομαδοποίηση εγγραφών
  - Διαχείριση χώρου δίσκου
  - Διαχείριση απομονωτικής μνήμης

# Περιεχόμενα Μαθήματος

---

Πιο συγκεκριμένα θα μελετήσουμε:

- Δομές Ευρετηρίου, Μεθόδους προσπέλασης
  - B-Δένδρα
  - B\*-Δένδρα
  - B<sup>+</sup>-Δένδρα
  - Παραλλαγές των B-Δένδρων
  - Συναρτήσεις κατακερματισμού
  - Δυναμικός κατακερματισμός
  - Επεκτατικός κατακερματισμός
  - Εκθετικός, γραμμικός κατακερματισμός

# Περιεχόμενα Μαθήματος

---

Πιο συγκεκριμένα θα μελετήσουμε:

- Βελτιστοποίηση και επεξεργασία ερωτημάτων
  - Αλγεβρικοί μετασχηματισμοί
  - Ευριστική βελτιστοποίηση
  - Υλοποίηση πράξεων και εκτίμηση κόστους
  - Εξαγωγή αποτελέσματος

# Περιεχόμενα Μαθήματος

---

Πιο συγκεκριμένα θα μελετήσουμε:

- Διαχείριση συναλλαγών
  - Συναλλαγές
  - Ταυτοχρονισμός
  - Επανάκτηση δεδομένων
- Αντικειμενοστρεφή και αντικειμενοσχεσιακά συστήματα
- Σύγχρονες εφαρμογές
- Παράλληλα και καταναμημένα συστήματα

# Περιεχόμενα Μαθήματος

---

Πιο συγκεκριμένα θα μελετήσουμε:

- Procedures
- Functions
- Triggers
- Views
- Indexes

Θα χρησιμοποιηθεί η MySQL



# Περιεχόμενα Μαθήματος

---

- Γραπτή Εξέταση (80%)
  - Εργασίες (20% + bonus 20%)
  - Τελικός Βαθμός =  $0.2 * \text{Εργασίες} + 0.8 * \text{Εξέταση}$
- 
- Θεωρία Τετάρτη 13:00-15:00
  - Εργαστήριο Τετάρτη 15:00-17:00

# Η ανάγκη διαχείρισης δεδομένων

---

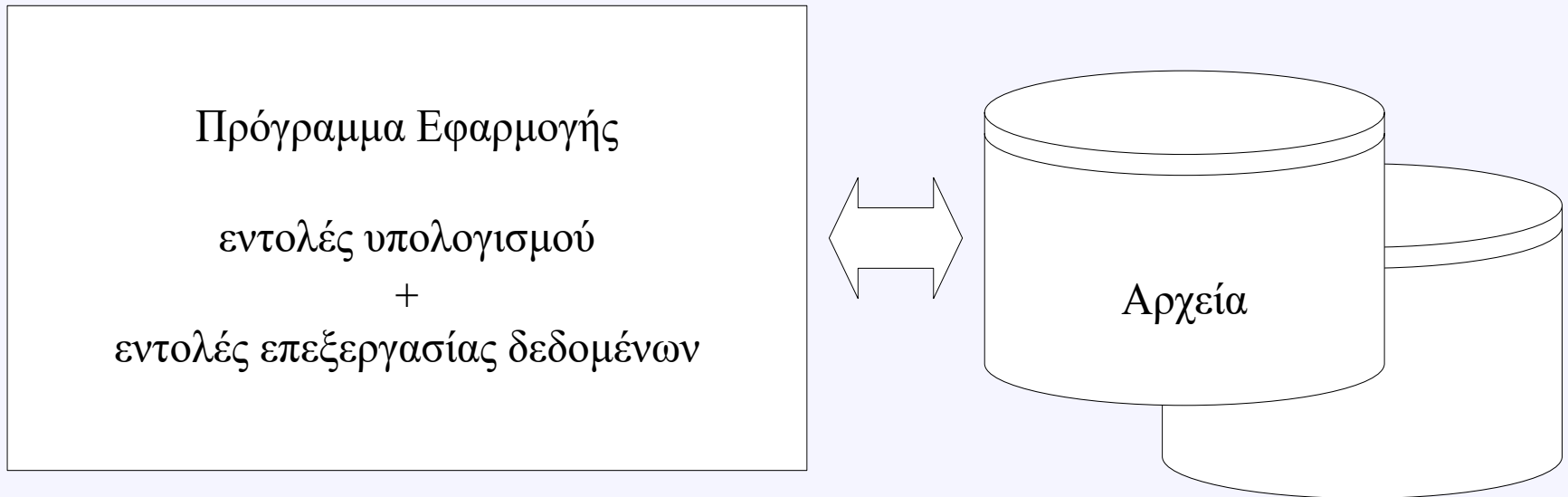
- Τα **δεδομένα** υπάρχουν παντού
  - Τα πάντα καταγράφονται.
  - Εφαρμογές με τεράστιο όγκο δεδομένων.
- Οι μεγάλες ποσότητες δεδομένων επιβάλλουν την εύρεση αποτελεσματικών μεθόδων αποθήκευσης.
- Τα δεδομένα πρέπει να
  - έχουν συγκεκριμένη δομή και
  - να είναι οργανωμένα
  - ώστε να διευκολύνεται η αναζήτηση και η ενημέρωσή τους.
- Η **Διαχείριση Δεδομένων** (data management) μελετά μεθόδους οργάνωσης με βασικό στόχο την αποτελεσματική και αποδοτική αναζήτηση.

# Συστήματα Αρχείων

---

- Προ Βάσεων Δεδομένων (ΒΔ) εποχή
  - Χρησιμοποιούνταν τα αρχεία του λειτουργικού συστήματος για την αποθήκευση δεδομένων.
  - Η αποθήκευση και η διαχείριση των δεδομένων ήταν αποκλειστικά στην ευθύνη των εφαρμογών.
- Η ανάγνωση, αποθήκευση, η ενημέρωση και η επεξεργασία των δεδομένων ανήκε στην ευθύνη του προγραμματιστή της εφαρμογής.
- Άρα τόσο η **επιχειρησιακή λογική** όσο και η **διαχείριση δεδομένων** ήταν στην ευθύνη της εφαρμογής.

# Συστήματα Αρχείων



- Μπορείτε να σκεφτείτε
  - Παραδείγματα εφαρμογών διαχείρισης δεδομένων;
  - Παραδείγματα καλής και κακής λειτουργίας αυτού του τρόπου διαχείρισης;

# Μειονεκτήματα: ευθύνη του προγραμματιστή

---

- Τα πάντα είναι **ευθύνη του προγραμματιστή**.
- Σωστή ενημέρωση των αρχείων των δεδομένων, ανάλογα με τις εισαγωγές και τις διαγραφές που πραγματοποιούνται.
- Αναζήτηση και επεξεργασία δεδομένων
  - Προσδιορισμός απαραίτητων αρχείων
  - Ανάγνωση δεδομένων
  - Συνδυασμός πληροφορίας που περιέχεται σε διαφορετικά αρχεία.
- Η κατάσταση δυσκολεύει όσο αυξάνεται
  - Το πλήθος των δεδομένων
  - Ο αριθμός των αρχείων δεδομένων
  - Η πολυπλοκότητα των ερωτημάτων.

# Μειονεκτήματα: η δομή του αρχείου

---

- **Η δομή των αρχείων καθορίζεται από την εφαρμογή.**
  - Καθίσταται πολύ δύσκολη η αλλαγή της δομής ενός αρχείου.
- Πολλές φορές παρουσιάζεται η **ανάγκη** να προστεθεί **ένα νέο χαρακτηριστικό στα δεδομένα**.
  1. Αυτό συνεπάγεται μεταβολή της δομής του αρχείου και
  2. Μεταβολή της λειτουργικότητας της εφαρμογής ώστε να ληφθεί το νέο πεδίο υπόψη κατά την αναζήτηση, εισαγωγή και διαγραφή δεδομένων.
- Οι αλλαγές αυτές είναι χρονοβόρες και επικίνδυνες για την ακεραιότητα των δεδομένων.
- Μεγάλη πιθανότητα ανθρώπινου λάθους.

# Μειονεκτήματα: επικίνδυνος πλουραλισμός

---

- Η **μορφοποίηση** των **αρχείων εξαρτάται**
  - από την εκάστοτε **γλώσσα προγραμματισμού**.
- Πρόβλημα ομογενοποίησης
  - σε περιπτώσεις όπου απαιτείται η **χρήση αρχείων από διαφορετικές εφαρμογές** ή διαφορετικούς προγραμματιστές
  - που χρησιμοποιούν **διαφορετικές γλώσσες προγραμματισμού** σε διαφορετικά λειτουργικά συστήματα.
- Το πρόβλημα είναι ιδιαίτερα έντονο όταν τα δεδομένα αποθηκεύονται σε δυαδική (binary) μορφή και όχι σε αλφαριθμητική (alphanumeric, text) μορφή.

# Μειονεκτήματα: ταυτόχρονες λειτουργίες

---

- Εφαρμογές που εξυπηρετούν πολλούς χρήστες ταυτόχρονα.
- Διασφάλιση της προστασίας και της **ακεραιότητας των δεδομένων** κατά την ταυτόχρονη αναζήτηση, εισαγωγή και διαγραφή στοιχείων είναι αρκετά περίπλοκη
- Σε συγκεκριμένες λειτουργίες πρέπει να διασφαλίζεται η **ατομικότητα της εκτέλεσης**



# Μειονεκτήματα: δικαιώματα χρηστών

---

- Γενικά, κάθε χρήστης της εφαρμογής έχει διαφορετικά δικαιώματα σχετικά με την προσπέλαση των δεδομένων.
- Η εφαρμογή των κανόνων προστασίας των δεδομένων ανάλογα με τα δικαιώματα του χρήστη στηριζόμενοι στα εργαλεία του λειτουργικού συστήματος είναι δύσκολα επιτεύξιμο.
- Βασική μονάδα τα δεδομένα και όχι το αρχείο
  - **Είναι πολύ δύσκολο να ορίσουμε διαφορετικά δικαιώματα μέσα στο ίδιο αρχείο.**

# Μειονεκτήματα: δυσκαμψία στα ερωτήματα

---

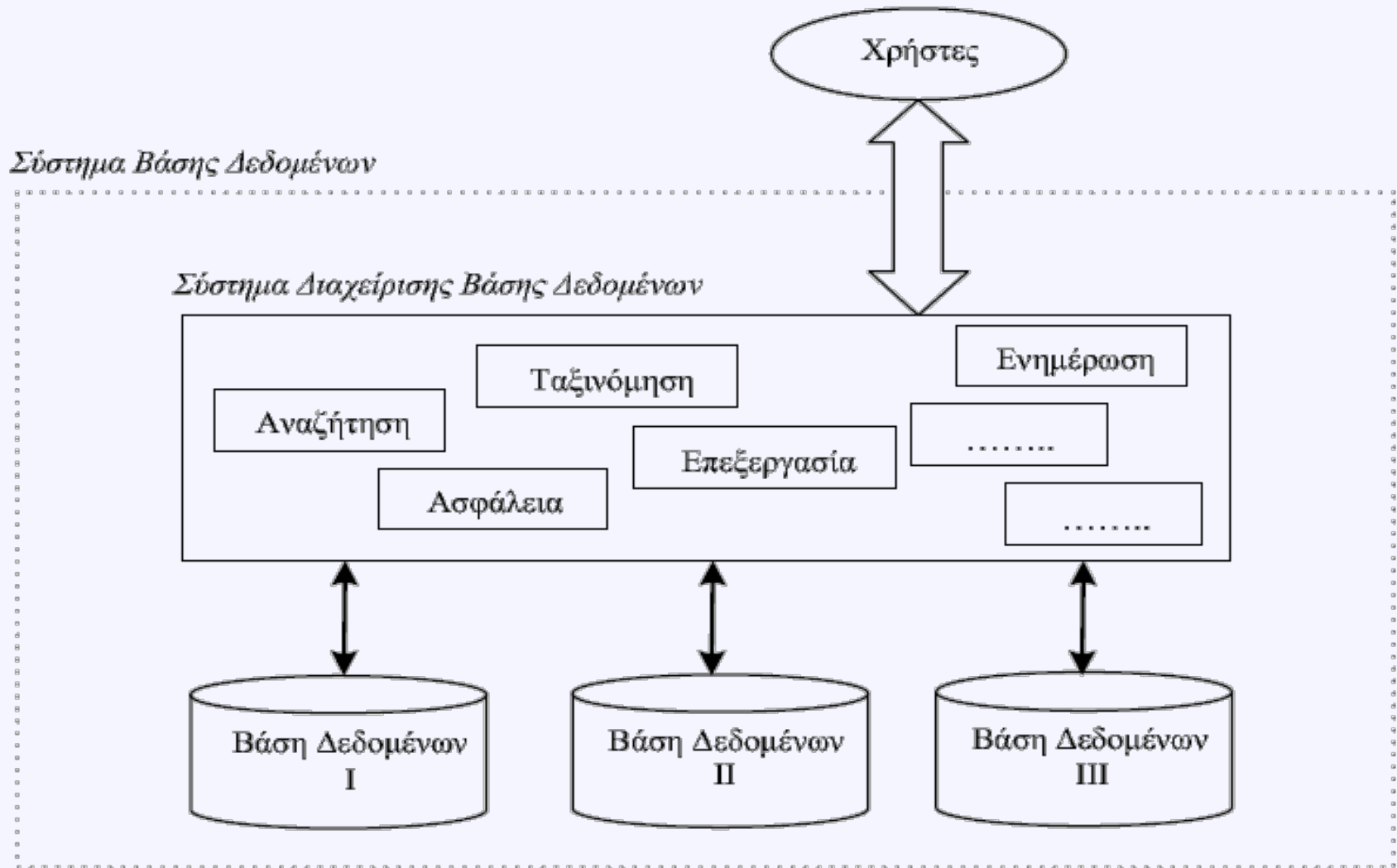
- Η επεξεργασία δεδομένων υλοποιείται στον κώδικα της εφαρμογής,
- Οι **διαφορετικοί τύποι ερωτημάτων** που μπορεί να δεχθεί η εφαρμογή
  - είναι **περιορισμένοι** και
  - **πρέπει να υλοποιηθούν εκ των προτέρων** (ή τουλάχιστον να έχουν προδιαγραφεί).

# ΒΔ: η εναλλακτική προσέγγιση

---

- **Βάση Δεδομένων (database):** αποτελεί μία συλλογή δεδομένων σχετικών μεταξύ τους τα οποία είναι δομημένα και αποθηκευμένα με κατάλληλο τρόπο.
- **Σύστημα Διαχείρισης Βάσεων Δεδομένων (database management system):** είναι ένα λογισμικό, το οποίο υλοποιεί όλες τις λειτουργίες που απαιτούνται, όπως αναζήτηση, εισαγωγή, διαγραφή, συγχρονισμός προσπελάσεων, προστασία και πολλές ακόμη.
- **Συστήματα Βάσεων Δεδομένων (Database System)**
  - Εναλλακτικά αντί αρχείων λειτουργικού συστήματος.
  - Αποτελείται από δύο βασικά μέρη: ΒΔ και ΣΔΒΔ

# ΒΔ: η εναλλακτική προσέγγιση



# ΒΔ: η εναλλακτική προσέγγιση

---

## Πλεονεκτήματα

- **Περιγραφή Δεδομένων** (μεταδεδομένα = δεδομένα για τα δεδομένα).
  - Δυνατότητα μεταβολής της δομής και της οργάνωσης από τους χρήστες.
- **Ανεξαρτησία Δεδομένων και Λειτουργιών**
  - Δεν αλλάζουν τα προγράμματα εφαρμογής όταν αλλάζουν τα δεδομένα.
- **Αποδοτικότερη Διαχείριση Δεδομένων**
  - Ο προγραμματιστής επικεντρώνεται στην εφαρμογή και όχι στα δεδομένα και τους αλγορίθμους επεξεργασίας τους.

# ΒΔ: η εναλλακτική προσέγγιση

---

## Πλεονεκτήματα

- **Προστασία Δεδομένων και Δικαιώματα Χρηστών**
  - για κάθε σύνολο δεδομένων ξεχωριστά.
- **Μηχανισμοί Ταυτόχρονης Προσπέλασης**
  - προσφέρονται από το σύστημα μηχανισμοί κλειδαριών.
- **Επεκτασιμότητα**
  - σε μεθόδους, χρήστες, δεδομένα, εφαρμογές.



# Χρειαζόμαστε Πάντα ΣΔΒΔ;

---

- **Όχι πάντοτε!** Θα πρέπει η φύση της εφαρμογής να είναι τέτοια που η χρήση ΣΔΒΔ θα επιφέρει θετικά αποτελέσματα.
- **Παράδειγμα:** εφαρμογές όπως επεξεργαστές κειμένου δε χρειάζονται ΣΔΒΔ.
- Η αλόγιστη χρήση ΣΔΒΔ αυξάνει τόσο το κόστος του λογισμικού όσο και την πολυπλοκότητα οπότε δε συνιστάται για απλές εφαρμογές που δεν έχουν ιδιαίτερες απαιτήσεις από διαχείριση δεδομένων.



# «Παραδοσιακές» Εφαρμογές

---

- **Άμεση Επεξεργασία Συναλλαγών**  
(On-line transaction processing - OLTP)
- Μισθοδοσία
- Κρατήσεις Θέσεων
- Τραπεζικές Συναλλαγές
- Διαχείριση Εταιρικών Δεδομένων

# Σύγχρονες Εφαρμογές

---

- Πολυμεσικές εφαρμογές (multimedia DBs)
- Διαχείριση γεωγραφικών πληροφοριών (GIS)
- Διαχείριση κινούμενων αντικειμένων (moving objects)
- **Άμεση αναλυτική επεξεργασία** (OLAP)
- **Εξόρυξη δεδομένων** (data mining)
- **Ανάκτηση πληροφορίας** (information retrieval)

# Διαχείριση Μεγάλου Όγκου Δεδομένων

---

- Ταιράστιο πλήθος αισθητήρων
- Internet of Things (IoT)
- Οχήματα διασυνδεδεμένα
- Δίκτυα νέας γενιάς
- Έξυπνες συσκευές
- Κόστος flash μνήμης



**Big Data**