

Βάσεις Δεδομένων

1 - Εισαγωγή

Δρ. Χαράλαμπος Καρανίκας

Περιεχόμενα

- Εφαρμογές & Τύποι Βάσεων Δεδομένων
- Βασικοί ορισμοί
- Τυπική Λειτουργικότητα ΣΔΒΔ - DBMS
- Παράδειγμα μιας βάσης δεδομένων (UNIVERSITY)
- Κύρια χαρακτηριστικά της προσέγγισης με βάσεις δεδομένων
- Χρήστες Βάσης Δεδομένων
- Πλεονεκτήματα από τη χρήση της προσέγγισης με βάσεις δεδομένων
- Πότε δεν χρησιμοποιούμε βάσεις δεδομένων
- Ιστορική εξέλιξη της τεχνολογίας βάσεων δεδομένων

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

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
-

Βασικοί ορισμοί 1/3

Βάση δεδομένων (Database):

Μια συλλογή συσχετιζόμενων δεδομένων.

- Μία συλλογή από συστηματικά μορφοποιημένα σχετιζόμενα δεδομένα στα οποία είναι δυνατή η ανάκτηση δεδομένων μέσω αναζήτησης κατ' απαίτηση
- Τα δεδομένα πρέπει να αυτοπροσδιορίζονται και να έχουν μια σχηματική δομή.

Αναφερόμαστε σε οργανωμένες, διακριτές συλλογές σχετιζόμενων δεδομένων ηλεκτρονικά αποθηκευμένων, στο λογισμικό που χειρίζεται τέτοιες συλλογές (Σύστημα Διαχείρισης Βάσεων Δεδομένων, ή DBMS).

Πέρα από την εγγενή της ικανότητα να αποθηκεύει δεδομένα, η βάση δεδομένων παρέχει μέσω του σχεδιασμού και του τρόπου ιεράρχησης των δεδομένων, τα αποκαλούμενα συστήματα διαχείρισης περιεχομένου, δηλαδή τη δυνατότητα γρήγορης άντλησης και ανανέωσης των δεδομένων.

Βασικοί ορισμοί 2/3

Δεδομένα (Data):

Γνωστά γεγονότα τα οποία μπορούν να καταγραφούν και τα οποία με έμμεσο τρόπο έχουν συγκεκριμένη - υπονοούμενη σημασία (implicit meaning).

- Ένα μη αξιολογημένο σύνολο διακριτών στοιχείων, μιας παρατήρησης - αναφοράς, μια συλλογή που αποτυπώνει «τιμές» επί αντικειμένων, προσώπων, γεγονότων κλπ.
- Τα δεδομένα μπορεί να είναι σημεία πληροφοριών επί επιστημονικών παρατηρήσεων ή συμπεριφοράς και να περιλαμβάνουν λέξεις - έννοιες, αριθμούς, σύμβολα, διαγράμματα, σχέδια, φωτογραφίες, μαγνητοταινίες κλπ που περιγράφουν ή αντιπροσωπεύουν ποσότητες, έννοιες, ιδέες, αντικείμενα, γεγονότα, καταστάσεις και λειτουργίες. Ενδεχομένως κάποιοι από τους τύπους δεδομένων που παρατίθενται εμπεριέχουν ήδη εμφανείς πληροφορίες, όχι όμως την πληροφορία στο επίπεδο που συνθέτει μια αξιολόγησή τους.

Βασικοί ορισμοί 2/3 συνέχεια

Τα δεδομένα περιγράφουν μόνο μέρος ενός συμβάντος, δεν περιλαμβάνουν καμία ανάλυση, κριτική ή αξιόπιστη βάση για περαιτέρω ενέργεια. Τα δεδομένα δεν αναφέρουν τίποτα για τη σημαντικότητά τους ή τη σχέση τους προς οτιδήποτε. Σε αντίθεση, τα δεδομένα μιας παρατήρησης έχουν ορισμένο νόημα και είναι οργανωμένα για συγκεκριμένο σκοπό τα οποία και υπόκεινται σε ανάλυση και αξιολόγηση.

Με διαμόρφωση, οργάνωση, μαθηματική ή στατιστική ανάλυση, διόρθωση λαθών ή συμπίεση παράγεται η **πληροφορία**. Τη διαδικασία αυτή ονομάζουμε **επεξεργασία δεδομένων**.

Βασικοί ορισμοί 3/3

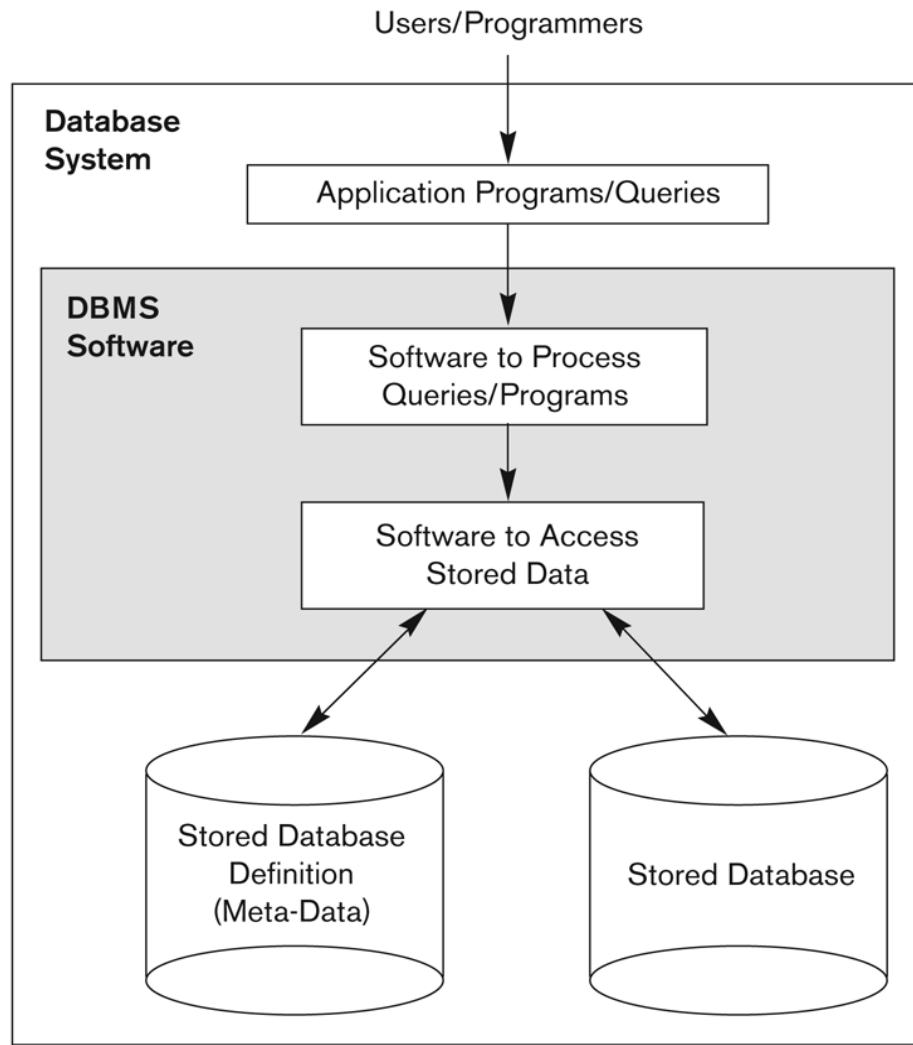
❖ Μικρόκοσμος (Mini - world) ή Πεδίο Αναφοράς:

- Κάποιο τμήμα (υποσύνολο) του πραγματικού κόσμου για το οποίο αποθηκεύονται τα δεδομένα σε μια βάση δεδομένων. Για παράδειγμα, οι εγγραφές φοιτητών ενός πανεπιστημίου.

❖ Σύστημα Διαχείρισης Βάσης Δεδομένων Database Management System (DBMS):

- Λογισμικό για την δημιουργία, τη διατήρηση και επεξεργασία μιας βάσης δεδομένων.
π.χ., Oracle, IBM DB2, SQL-Server, Postgres, MySQL, MS Access, Sybase κλπ
- Κύριος σκοπός των DBMS: Αποθήκευση και Διαχείριση
 - ✓ Εύκολα και βολικά (conveniently)
 - ✓ Γρήγορα και αποδοτικά (efficiently)
 - ✓ Αξιόπιστα και με συνέπεια (reliably and consistently)

Γενικά το περιβάλλον ενός συστήματος βάσης δεδομένων



Σύστημα βάσης δεδομένων (Database System): Το λογισμικό του ΣΔΒΔ μαζί με τα δεδομένα. Μερικές φορές, συμπεριλαμβάνονται και οι εφαρμογές.

Τυπικές Λειτουργίες σε ένα DBMS 1/2

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., δύο διαφορετικοί φοιτητές δεν μπορούν να έχουν τον ίδιο αριθμό ταυτότητας. Αυτός ο κανόνας πρέπει να ισχύει συνέχεια

Τυπικές Λειτουργίες σε ένα DBMS 2/2

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 δεδομένων 1/2

(με Εννοιολογικό Μοντέλο Δεδομένων, Conceptual Data Model)

- **Mini-world παραδείγματος:**

Ένα μέρος του πανεπιστημιακού περιβάλλοντος (UNIVERSITY).

- **Κάποιες Οντότητες του mini-world:**

- STUDENTs
- COURSEs
- SECTIONs (of COURSEs)
- DEPARTMENTs (academic)
- INSTRUCTORs

Παράδειγμα μιας βάσης δεδομένων 2/2

(με Εννοιολογικό Μοντέλο Δεδομένων, Conceptual Data Model)

- **Κάποιες Σχέσεις του mini-world:**

- SECTIONS *are of specific* COURSEs
- STUDENTs *take* SECTIONs
- COURSEs *have prerequisite* COURSEs
- INSTRUCTORs *teach* SECTIONs
- COURSEs *are offered by* DEPARTMENTs
- STUDENTs *studying in* DEPARTMENTs

Σημείωση: Οι παραπάνω οντότητες (entities) και σχέσεις (relationships) εκφράζονται τυπικά σε ένα εννοιολογικό μοντέλο δεδομένων, όπως το Μοντέλο Οντοτήτων Συσχετίσεων που θα δούμε στην συνέχεια

Παράδειγμα μιας απλής βάσης δεδομένων

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

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

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

Μια βάση δεδομένων που αποθηκεύει φοιτητές και πληροφορίες μαθήματος

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Κύρια χαρακτηριστικά της προσέγγισης με βάσεις δεδομένων 1/4

1. Αυτό-περιγραφική φύση (Self-describing nature) ενός database system:

- Ο Κατάλογος (catalog) του DBMS αποθηκεύει την περιγραφή μιας συγκεκριμένης database (π.χ., δομή δεδομένων, τύπους δεδομένων, και περιορισμούς)
- Η περιγραφή λέγεται μετα-δεδομένα (meta-data).
- Αυτό επιτρέπει στο λογισμικό του DBMS να λειτουργεί με διαφορετικές εφαρμογές βάσεων δεδομένων.

RELATIONS

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

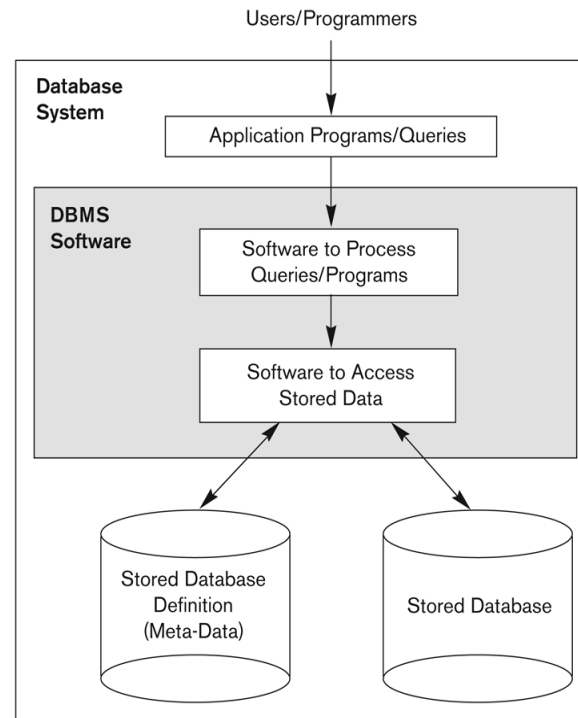
COLUMNS

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

Κύρια χαρακτηριστικά της προσέγγισης με βάσεις δεδομένων 2/4

2. Ανεξαρτησία μεταξύ Προγραμμάτων και Δεδομένων(program-data independence):

Επιτρέπει να αλλάξουμε δομή και οργάνωση αποθήκευσης των δεδομένων χωρίς να χρειαστεί να αλλάξουμε τα προγράμματα DBMS access programs



Π.χ., εάν χρειαστεί η DB να χωριστεί στα δύο αυτό δεν επηρεάζει τον χρήστη

Κύρια χαρακτηριστικά της προσέγγισης με βάσεις δεδομένων 3/4

3. Αφαίρεση (Data Abstraction):

- Ένα μοντέλο δεδομένων (**data model**) χρησιμοποιείται για απόκρυψη των λεπτομερειών αποθήκευσης και για παρουσίαση στους χρήστες μιας εννοιολογικής όψης της βάσης δεδομένων.
- Τα προγράμματα αναφέρονται στο μοντέλο δεδομένων και όχι στις λεπτομέρειες αποθήκευσης των δεδομένων.

4. Υποστήριξη πολλαπλών όψεων (views) των δεδομένων:

- Κάθε χρήστης μπορεί να βλέπει μια διαφορετική όψη της βάσης δεδομένων, που περιγράφει μόνο τα δεδομένα που τον ενδιαφέρουν.
π.χ., η γραμματεία της πληροφορικής βλέπει μόνο όσες εγγραφές φοιτητών σχετίζονται με τις εγγραφές των φοιτητών του τμήματος πληροφορικής.

Κύρια χαρακτηριστικά της προσέγγισης με βάσεις δεδομένων 4/4

5. Διαμοιρασμός δεδομένων (Sharing of data) και επεξεργασία συναλλαγών πολλαπλών χρηστών:

- Επιτρέπει ένα σύνολο ταυτόχρονων χρηστών (concurrent users) να ανακτούν αλλά και να ενημερώνουν δεδομένα στη βάση.
- Concurrency control στο DBMS εγγυάται ότι κάθε συναλλαγή στην βάση εκτελείται ή απορρίπτεται ορθά.
- Recovery subsystem διασφαλίζει ότι κάθε ολοκληρωμένη συναλλαγή έχει μόνιμα καταγράψει το αποτέλεσμα της στην βάση δεδομένων.
- OLTP (Online Transaction Processing) είναι ένα σημαντικό μέρος των εφαρμογών βάσεων δεδομένων. Αυτό επιτρέπει τις εκατοντάδες των ταυτόχρονων συναλλαγών να εκτελούνται ανά δευτερόλεπτο.

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

Θεωρήστε:

Account[7] = €100

T1: Deposit1(100, 7)

T2: Deposit2(50, 7)

Εκτέλεση 1

T1 T2

Read1
Write1

Read2
Write2

Εκτέλεση 2

T1 T2

Read1
Write1

Read2
Write2

Concurrency control:

Διασφαλίζεται
ότι η Εκτέλεση 2
δε θα συμβεί

Χρήστες Βάσης δεδομένων 1/2

Οι χρήστες μπορούν να χωριστούν σε:

- Εκείνους που πραγματικά χρησιμοποιούν και ελέγχουν το περιεχόμενο της βάσης δεδομένων, καθώς και εκείνους που σχεδιάζουν, αναπτύσσουν και συντηρούν της εφαρμογές βάσης δεδομένων (που ονομάζονται «Εργαζόμενοι στο Προσκήνιο»), και
- Όσους σχεδιάζουν και αναπτύσσουν το λογισμικό του ΣΔΒΔ και των σχετικών εργαλείων (που ονομάζονται "Οι εργαζόμενοι πίσω από τη σκηνή»).

Χρήστες Βάσης δεδομένων 2/2

Εργαζόμενοι στο Προσκήνιο

Διαχειριστές Βάσεων (Database Administrators - DBA)

- Έλεγχος πρόσβασης, επίδοσης, ορθότητας, backups, κτλ.
- Συνήθως κατέχουν επαγγελματικά πιστοποιητικά: π.χ., Microsoft MCDBA, Oracle Certified Master (OCM), κτλ.,
- Είναι οι υπεύθυνοι για την πληροφορία σε ένα οργανισμό.

Σχεδιαστές / Προγραμματιστές Βάσεων (DB Designers / Developers)

- **Εννοιολογική Αναπαράσταση** των Απαιτήσεων του Χρήστη
- Παραγωγή **Διαγραμμάτων** και **Δομών** της Βάσης για αναπαράσταση της πληροφορίας (Συνήθως από Αναλυτές)

Κατηγορίες των τελικών χρηστών (End-users) 1/2

Εργαζόμενοι στο Προσκήνιο

End-users: Χρησιμοποιούν τα δεδομένα για SQL ερωτήσεις, αναφορές και μερικοί ενημερώνουν το περιεχόμενο της βάσης. Οι τελικοί χρήστες μπορούν να ταξινομηθούν σε:

- ✓ **Περιστασιακούς (Casual):** προσπέλαση στη βάση περιστασιακά όταν χρειάζεται
- ✓ **Naïve or Parametric:** απαρτίζουν ένα μεγάλο τμήμα του πληθυσμού των τελικών χρηστών.
 - Χρησιμοποιούν καλά ορισμένες λειτουργίες με την μορφή "προκατασκευασμένων συναλλαγών» στη βάση δεδομένων.
 - Με παραδείγματα ταμίες ή υπαλλήλους κρατήσεων.

Κατηγορίες των τελικών χρηστών (End-users) 2/2

✓ Έμπειροι (Sophisticated):

- Περιλαμβάνουν αναλυτές, επιστήμονες, μηχανικούς και άλλους που είναι εξοικειωμένοι με τις δυνατότητες του συστήματος.
- Πολλοί χρησιμοποιούν εργαλεία με την μορφή πακέτων λογισμικού που λειτουργούν πάνω από την υφιστάμενη βάση.

✓ Αυτόνομοι (Stand-alone):

- Συνήθως διατηρούν προσωπικές βάσεις δεδομένων χρησιμοποιώντας έτοιμα προς χρήση πακέτα εφαρμογών.
- Παράδειγμα είναι ένα φορολογικό πρόγραμμα που δημιουργεί μια βάση για έναν συγκεκριμένο χρήστη.
- Ένα άλλο παράδειγμα είναι ο χρήστης που δημιουργεί ένα βιβλίο διευθύνσεων

Πλεονεκτήματα από τη χρήση της προσέγγισης με Βάση Δεδομένων 1/4

Αναφέραμε ήδη **4 πλεονεκτήματα** των Βάσεων σε σχέση με **συμβατικά αρχεία**.

1. Αυτό-περιγραφική Φύση της ΒΔ (Κατάλογος)
2. Ανεξαρτησία Προγράμματος-Δεδομένων
3. Παροχή Πολλαπλών Όψεων
4. Διαμοιρασμός Φόρτου και Επεξεργασία Δοσοληψιών.

Πλεονεκτήματα από τη χρήση της προσέγγισης με Βάση Δεδομένων 2/4

Έλεγχος των πλεονασμών (redundancy)

- Π.χ., Στο παράδειγμα του University μπορεί να υπάρχουν πολλές ανεξάρτητες βάσεις: π.χ., LIBRARY-DB, CS-DB, UCY-DB, etc.
- Αυτό δημιουργεί προβλήματα:
 - Σπατάλη Χώρου (Storage Space Waste)
 - Περιπτώσεις Ασυνέπειας (Inconsistencies)
 - Σπατάλη Χρόνου (Duplication of effort)
- Ιδανικά θα θέλαμε η πληροφορία να αποθηκεύεται σε μια μόνο τοποθεσία.

Πως; Κατά τη φάση της σχεδίασης και κανονικοποίησης.

Πλεονεκτήματα από τη χρήση της προσέγγισης με Βάση Δεδομένων 3/4

Περιορισμός της Μη-Εξουσιοδοτημένης Προσπέλασης

- Δυνατότητα δημιουργίας ρόλων, ομάδων, κτλ. όπως σε συμβατικά Λειτουργικά Συστήματα.

Παροχή Εφεδρικών Αντιγράφων (Backup) και Μηχανισμών Ανάκαμψης (Recovery)

- Π.χ., Εάν έχουμε διακοπή ρεύματος (χωρίς UPS) δεν θα βρεθεί σε ασυνεπή κατάσταση η βάση δεδομένων.

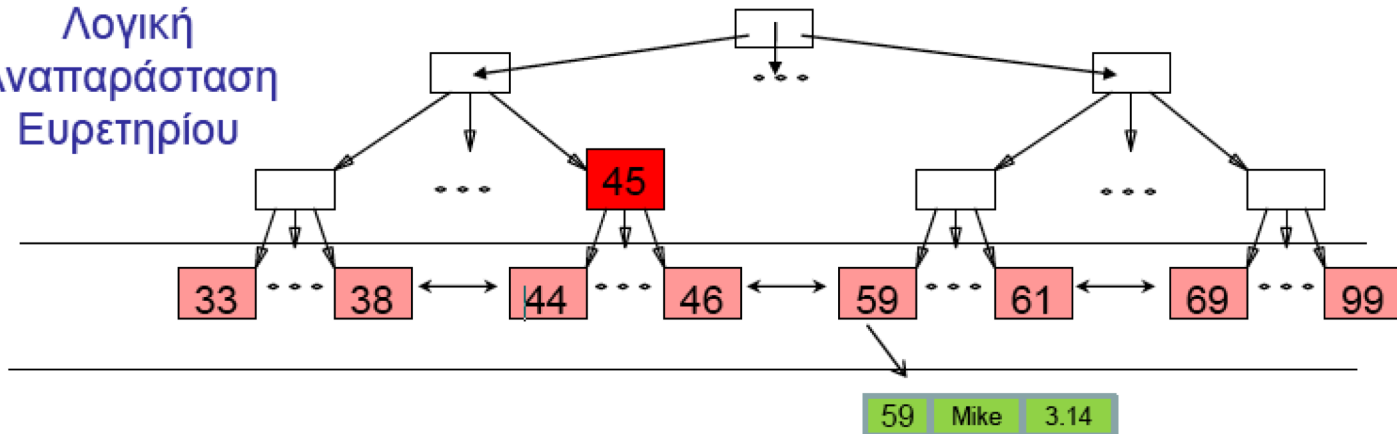
Υποστήριξη Δομών Αποθήκευσης για Αποτελεσματική Επεξεργασία Επερωτήσεων

- Π.χ., Ευρετήρια Κατακερματισμού, Δενδρικά Ευρετήρια, Χωρικά Ευρετήρια και πολλά άλλα.

2.

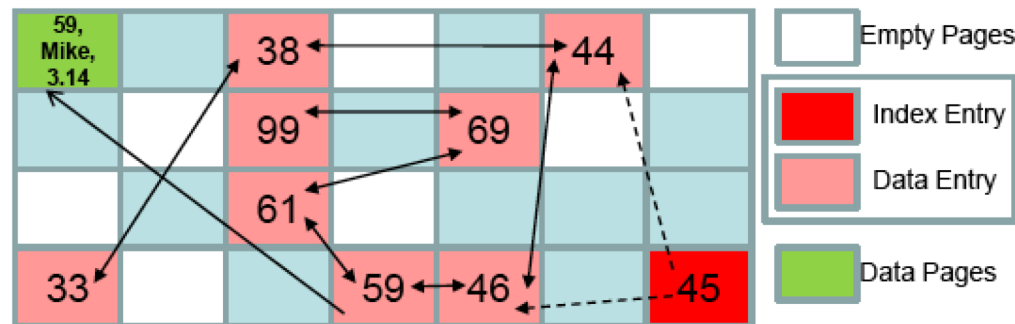
Παράδειγμα Δενδρικού Ευρετηρίου

Λογική
Αναπαράσταση
Ευρετηρίου



Φυσική
Αναπαράσταση
Ευρετηρίου στη
Δευτερεύουσα
Μνήμη

Physical Layout (on Disk)



Πλεονεκτήματα από τη χρήση της προσέγγισης με Βάση Δεδομένων 4/4

Δυνατότητα Επιβολής Κανόνων Ακεραιότητας

- Π.χ., Επιβολή κανόνα: «Σε κανένα μάθημα δεν μπορούν να εγγραφούν πάνω από 99 φοιτητές»
- Παράδειγμα σε SQL-DDL*:
CREATE TABLE COURSE
(CourseID INT NOT NULL DEFAULT 1
Positions INT NOT NULL **CHECK (Pos<100)....)**

Δυνατότητα Εκτέλεσης Σκανδαλών (Triggers)

- Π.χ., Όποτε αλλάξει κάτι στον λογαριασμό του χρήστη στείλε του SMS!
- ```
CREATE TRIGGER change AFTER INSERT OR DELETE OR
UPDATE ON CustomerAccount
FOR EACH ROW EXECUTE PROCEDURE sendSMS ();
```

# Πρόσθετες επιπτώσεις από τη χρήση της προσέγγισης Βάσης Δεδομένων 1/2

## **Δυνατότητα Επιβολής Κανόνων στην Αναπαράσταση Δεδομένων σε ένα οργανισμό:**

- Π.χ., Σχεδιάζοντας και Υλοποιώντας τη βάση UNIVERSITY προτυποποιείται ότι το STUDENT.Name είναι μέχρι 30 χαρακτήρες.

## **Μειωμένος χρόνος ανάπτυξης εφαρμογών**

- Υπολογίζεται ότι ο χρόνος ανάπτυξης εφαρμογών με βάσεις μειώνεται από **75-84%** του χρόνου που απαιτείται.

## **Ευελιξία**

- Εύκολη μετάβαση σε άλλη κωδικοποίηση βάσης (με εργαλεία EXPORT).

# Πρόσθετες επιπτώσεις από τη χρήση της προσέγγισης Βάσης Δεδομένων 2/2

## **Διαθεσιμότητα επίκαιρων πληροφοριών :**

- Εξαιρετικά σημαντικό για on-line συστήματα συναλλαγών, όπως αεροπορικά, ξενοδοχεία, κρατήσεις αυτοκινήτων κ.λ.π.

## **Οικονομίες κλίμακας:**

- Άσκοπη επικάλυψη των πόρων και του προσωπικού μπορεί να αποφευχθεί με ενοποίηση δεδομένων και εφαρμογών μεταξύ των τμημάτων.

# Πότε να μην χρησιμοποιούμε Βάσεις;

**Υπάρχει συνήθως μεγάλο κόστος αγοράς, π.χ.,**

- Oracle Standard Edition (per CPU): **5,900\$**
- Oracle Enterprise Edition (per CPU): **47,500\$**
- IBM DB2 v9.7 Enterprise: **25,000\$**
- SQL Server 2008 Enterprise: **25,000\$**

• Τα πιο πάνω ΔΕΝ περιλαμβάνουν κόστος αγοράς υλικού (server), λειτουργικού συστήματος, εκπαίδευση, κτλ.!

• Όταν ο προϋπολογισμός ενός project είναι περιορισμός τότε λαμβάνεται σημαντικά υπόψη μια τέτοια παράμετρος.

Ακόμη και εάν χρησιμοποιηθεί DBMS **ανοικτού πηγαίου κώδικα** (π.χ., MySQL, PostgreSQL, κτλ.), οι οποίες είναι **δωρεάν**, τότε υπάρχουν και πάλι λόγοι για τους οποίους δεν θέλουμε να τις χρησιμοποιήσουμε.

# Πότε ένα DBMS δεν αρμόζει?

## ➤ Μειονεκτήματα:

- Ακριβή η αγορά
- Ακριβή η συντήρηση (χρειάζεται διαχειριστή)
- “Ακριβή” η λειτουργία (χρειάζεται πολλούς πόρους)

## ➤ Για αυτό, είναι “υπερβολή” όταν:

- Η βάση είναι μικρή
- Η βάση έχει απλή δομή
- Η εφαρμογή είναι απλή και δεν αναμένετε να αλλάξει
- Μπορούμε να ανεχθούμε σφάλματα
- Δεν απαιτούνται πολλοί και ταυτόχρονοι χρήστες
- Δεν υπάρχει το αναγκαίο προσωπικό

# Εφαρμοσιμότητα: Πότε ένα DBMS χρειάζεταιται?

Για την επιτυχή συνένωση μεγάλων και συσχετιζόμενων δεδομένων

Υποστηρίζει διάθεση και εξόρυξη πληροφοριών

- Εξόρυξη δεδομένων (data mining)
- Συμπερασματικές Βάσεις (deductive databases)
- Ενεργές Βάσεις (active databases)

DBMS εξασφαλίζει πραγματικό Quality of Service (QoS)

QoS είναι η δυνατότητα να παρέχει διαφορετική προτεραιότητα σε διαφορετικές εφαρμογές, χρήστες, ή ροές δεδομένων, ή να εγγυηθεί ένα ορισμένο επίπεδο απόδοσης σε μια ροή δεδομένων.

# Ιστορική εξέλιξη της τεχνολογίας βάσεων δεδομένων 1/6

- ❑ Πρώτο σύστημα βάσεως δεδομένων
  - Πήλινες πλάκες για καταγραφή συναλλαγών αγαθών (πριν 6000 χρόνια)
- ❑ Πηλός → Πάπυρος → Περγαμινή → Χαρτί
- ❑ 1890, Hollerith: διατρητές κάρτες για επεξεργασία δεδομένων
- ❑ 1950s: Η/Υ / Μαγνητικές ταινίες για αποθήκευση δεδομένων
- ❑ 1960s/70s: μαγνητικοί δίσκοι
- ❑ 1980s: εμπορικά σχεσιακά DBMSs
- ❑ Αρχές 1990s: σχεδιασμός του ANSI SQL2
- ❑ Τέλη 1990s: WWW boom
  - Έκδοση του ANSI SQL3
  - Εκτενή ανάπτυξη βάσεων δεδομένων
  - Βάσεις δεδομένων είναι online και διαθέσιμες 24x7

**Technically, Moses**



**was the first person  
with a tablet  
downloading data  
from the cloud**

# Ιστορική εξέλιξη της τεχνολογίας βάσεων δεδομένων 2/6

## □1960: Αρχικές Εφαρμογές «Βάσεων»:

- Ξεκίνησαν με την δημιουργία των Η/Υ τη δεκαετία του '60 και έχουν από τότε αλλάξει σε όλα τα επίπεδα.
- Το πρώτο λογισμικό ΒΔ το **Integrated Data Store (IDS) / CodasyI** το οποίο προσωποποιήθηκε το 1971.

## □1970: Σχεσιακές Βάσεις Δεδομένων

- Το **Σχεσιακό Μοντέλο** (όπου η πληροφορία αναπαριστάται με σχέσεις) **προτάθηκε** το 1969 από τον **E.F. Codd**.
- Στη συνέχεια **ερευνήθηκε** και **δοκιμάστηκε** εκτενώς στο IBM Research και πολλά πανεπιστήμια.
- Προϊόντα Σχες. DBMS δημιουργούνται στις αρχές του 1980
  - Η **Oracle** δημιουργείται το 1977 από ένα project το οποίο επιχορηγείται από το **CIA**. Η Oracle είναι σήμερα ένας από τους κολοσσούς στο χώρο των βάσεων με γραφεία σε **126 χώρες** του κόσμου!

# Ιστορική εξέλιξη της τεχνολογίας βάσεων δεδομένων 3/6

## □ 1980: Αντικειμενοστρεφείς Βάσεις:

- **Object-Oriented Database Management Systems (OODBMSs)** εισήχθηκαν περί τα τέλη του **1980s** σε μια ανάγκη να αναπαριστούν περίπλοκες ανάγκες διαχείρισης δεδομένων τα οποία προέκυπταν σε **CAD (Computer-Aided design) εφαρμογές**.
- Σε αυτές, τα δεδομένα αναπαριστώνται από **Αντικείμενα (Objects)** κατά αντίστοιχο τρόπο με ΟΟ Προγραμματισμό.
- Σήμερα **δεν** χρησιμοποιούνται **πολύ εκτενώς**
- Ωστόσο πολλές έννοιες εισήχθηκαν σε Σχεσιακές DBMS δημιουργώντας την έννοια **Σχεσιό-Αντικειμενοστρεφών Βάσεων Δεδομένων [ Object-Relational DBMSs (ORDBMSs)]**
  - Π.χ., η PostgreSQL είναι μια τέτοια βάση δεδομένων η οποία επιτρέπει τη δημιουργία οντοτήτων με κληρονομικότητα και πολλά άλλα.

# Ιστορική εξέλιξη της τεχνολογίας βάσεων δεδομένων 4/6

## □ 1990: Δεδομένα στο Web και εφαρμογές Ηλεκτρονικού Εμπορίου:

- Το Web περιέχει περιεχόμενο στη Γλώσσα Μορφοποίησης HTML (Hypertext markup language) με συνδέσμους ανάμεσα σε ιστοσελίδες.
- Αυτό έχει δώσει χώρο σε ένα **νέο σύνολο εφαρμογών στο χώρο των βάσεων δεδομένων**. Συγκεκριμένα:
  - Διαδικτυακές Βάσεις Δεδομένων.
    - Παραγωγή Περιεχομένου HTML από τα δομημένα δεδομένα ΒΔ με χρήση γλωσσών όπως PHP, Java, ASP, κτλ.
    - Επιτρέπει επίσης αλλαγές στα δεδομένα της βάσης
  - Εφαρμογές Business-to-Business (Ανταλλαγή Εγγράφων με χρήση της XML (eXtended Markup Language))

# Επεκτείνοντας της Δυνατότητες των Βάσεων Δεδομένων 1/3

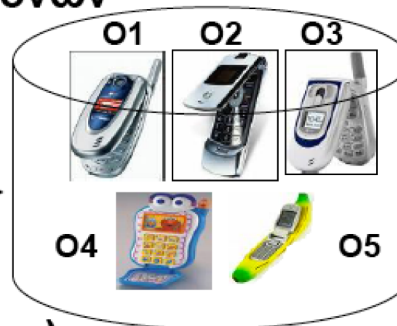
- Νέες λειτουργικότητες έχουν προστεθεί στα ΣΔΒΔ στους ακόλουθους τομείς:
  - Scientific Applications
  - XML (eXtensible Markup Language)
  - Image Storage and Management
  - Audio and Video Data Management
  - Data Warehousing and Data Mining
  - Spatial Data Management
  - Time Series and Historical Data Management
- Το παραπάνω δημιουργούν *νέα έρευνα και ανάπτυξη* στην ενσωμάτωση νέων τύπων δεδομένων, πολύπλοκων δομών δεδομένων, νέων τελεστών και μηχανισμών αποθήκευσης, ευρετηρίασης στα συστήματα βάσεων δεδομένων.

# Επεκτείνοντας της Δυνατότητες των Βάσεων Δεδομένων 2/3

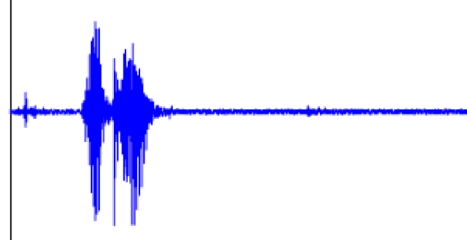
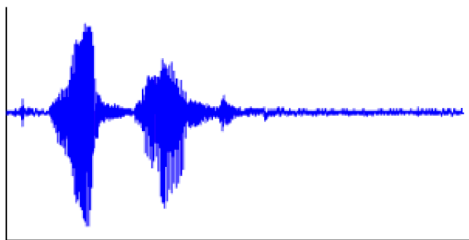
- Υπάρχουν προσθήκες δυνατοτήτων στα ακόλουθα πεδία:
  - Αποθήκευση και Διαχείριση Εικόνων**

$Q=(\text{χρώμα, texture, κτλ.})$

$Q$



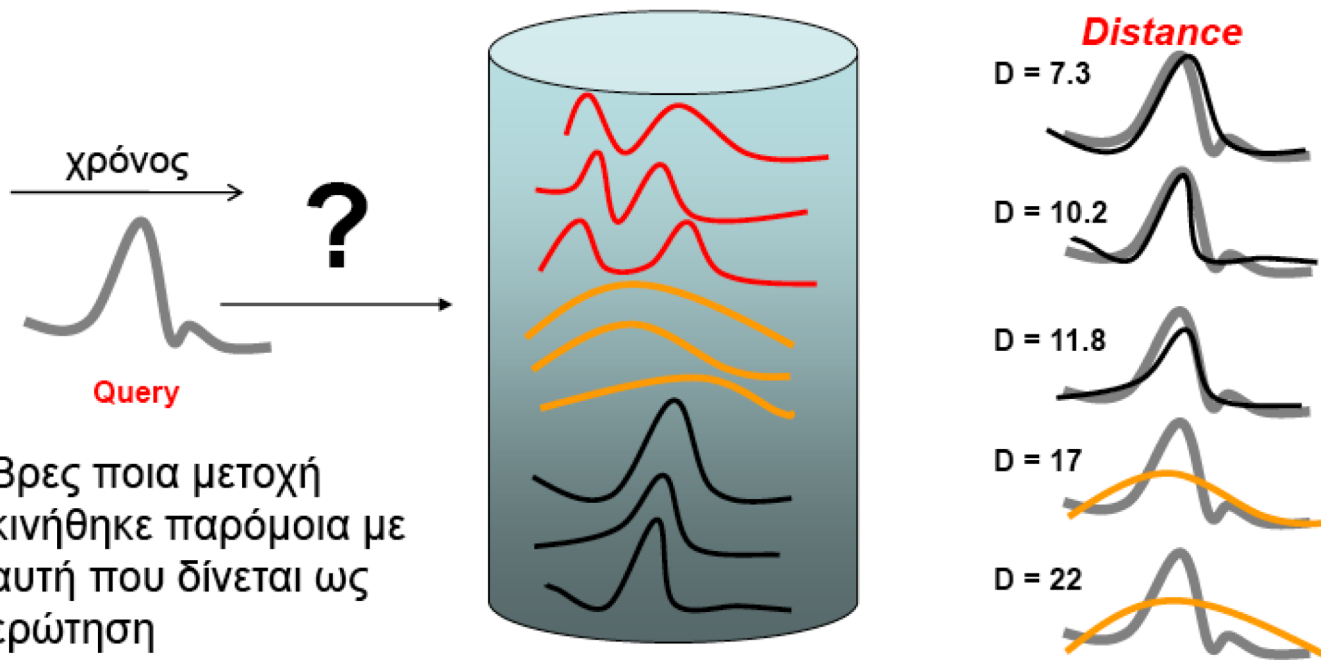
- Διαχείριση Ήχου (αλλά και Βίντεο)**



Σύγκριση Spectrographs για τον εντοπισμό όμοιων λέξεων

# Επεκτείνοντας της Δυνατότητες των Βάσεων Δεδομένων 3/3

- Υπάρχουν προσθήκες δυνατοτήτων στα ακόλουθα πεδία:
  - Διαχείριση Χρονικών (Temporal) Δεδομένων**



“Είναι ένα ταξίδι προς την επίτευξη επιχειρηματικής ευφυΐας (Business Intelligence), που ξεκινά με βάση οργανωμένες συλλογές δεδομένων, αποθηκευμένων με συγκεκριμένο και διακριτό τρόπο που περιέχουν ακριβή, χρήσιμα και κωδικοποιημένα δεδομένα.”

- “Ξεκινάμε, παρουσιάζοντας πρώτα εισαγωγικές έννοιες σχετικά με βάσεις δεδομένων. Τα θέματα που ανήκουν σε αυτή την ενότητα αφορούν την δομή των βάσεων, διαγράμματα σχημάτων (schema diagrams), γλώσσα ορισμού δεδομένων (DDL) και μια αναλυτική παρουσίαση της γλώσσας διαχείρισης δεδομένων (SQL). Χωρίς προαπαιτούμενες γνώσεις, ξεκινάμε από τις απλές έννοιες και δομές της SQL, παρουσιάζοντας με παραδείγματα πώς μπορούμε να διαχειριστούμε τα δεδομένα μας χρησιμοποιώντας εντολές εισαγωγής, διαγραφής, ενημέρωσης, πώς χτίζουμε ερωτήματα, προσθέτοντας βήμα-βήμα μεγαλύτερη πολυπλοκότητα, όπως ένθετα ερωτήματα (nested queries), κοινές υποεκφράσεις (CTE), ερωτήματα υπό συνθήκη (case statements), και τέλος παρουσιάζοντας συγκεντρωτικές λειτουργίες (aggregate functions) και λειτουργίες ταξινόμησης (rank, dense rank). Στη συνέχεια θα αναφερθούμε σε περισσότερο προχωρημένα θέματα βάσεων όπως βελτιστοποίηση ερωτημάτων, ασφάλεια, triggers, loops, cursors κλπ».
- «Το επόμενο βήμα αφορά την εισαγωγή στις αποθήκες δεδομένων (DW). Τις σχετικές έννοιες και τους όρους όπως μοντέλα διαστάσεων (dimensional modelling), πίνακες δεδομένων και μετρήσεις (fact tables & measurements), σχήματα αποθηκών δεδομένων (star, snowflake, fact). Τεχνικές εξαγωγής και μετασχηματισμού δεδομένων (ETL) και έννοιες όπως αντιστοίχιση (Mapping), επεξεργασία σε στάδια (staging), φόρτωση (Loading), μετασχηματισμός (Transform) κλπ. .... DW schema, ETL, Data Marts... ... δημιουργία αναφορών (reporting) και απεικόνισης (visualization), κύβους (OLAP Cubes), dashboards.»”

# Περίληψη

- ✓ Εφαρμογές & Τύποι Βάσεων Δεδομένων
- ✓ Βασικοί ορισμοί
- ✓ Τυπική Λειτουργικότητα ΣΔΒΔ - DBMS
- ✓ Παράδειγμα μιας βάσης δεδομένων (UNIVERSITY)
- ✓ Κύρια χαρακτηριστικά της προσέγγισης με βάσεις δεδομένων
- ✓ Χρήστες Βάσης Δεδομένων
- ✓ Πλεονεκτήματα από τη χρήση της προσέγγισης με βάσεις δεδομένων
- ✓ Πότε δεν χρησιμοποιούνται βάσεις δεδομένων