



Εισαγωγή στα Συστήματα Βάσεων Δεδομένων



Βασικές Έννοιες

Τι είναι μια βάση δεδομένων;

Βάση Δεδομένων: συλλογή από σχετιζόμενα δεδομένα

Ειδικού σκοπού λογισμικό για τα παραπάνω

Κάποιες χρήσεις ..

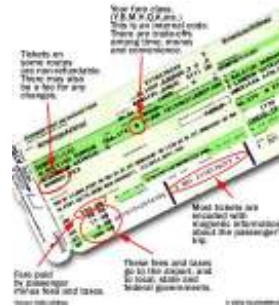


Εύδοξος

Η Ρεγιστρική Υπηρεσία Ολοκλήρωσης Διαχείρισης
Συμμετοχών και Αξιών Βαθμύλων

On-Line Δίγλωση Μαθημάτων

amazon.com.



Βάσεις Δεδομένων 2011-2012

3

Που είναι η βάση δεδομένων;



Βάσεις Δεδομένων 2011-2012

Ευαγγελία Πιπουρά

4

Άλλα είδη δεδομένων



The screenshot shows a web browser window displaying Yahoo! Local Maps. The search term is "uc berkeley". The results list 11 items, including "Uc Berkeley", "UC Berkeley Botanical Garden", "UC Berkeley Art Museum", "UC Berkeley Foundation", "Weinstein, Rhona S PhD - UC Berkeley Psych Clinic", "Capri Motel Berkeley", "University Art Museum", "Enterprise Rent A Car", "Faculty Club", "California Engineering Company", and "Campus Flowers". The map shows Berkeley, CA, with numbered markers (1-11) corresponding to the search results. The browser address bar shows "http://maps.yahoo.com/broadband#mvt=m&tp=1&tt=uc+berkeley&trf=0&lon=-122.256203&lat=37.871939".

Άλλα είδη δεδομένων



The screenshot shows the NCBI Genome browser interface. The search term is "Homo sapiens genome view". The interface displays a chromosome map with 22 autosomes, X, and Y chromosomes. The map shows the relative positions of genes and other genomic features. The browser address bar shows "http://www.ncbi.nlm.nih.gov/genome/".

Lineage: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Hominidae; Homo; Homo sapiens

Βάσεις Δεδομένων 2011-2012 Ευαγγελία Πιπρούρα 6



Τι είναι μια βάση δεδομένων;

Βάση Δεδομένων: συλλογή από σχετιζόμενα δεδομένα

Τεράστιος όγκος διαθέσιμων δεδομένων

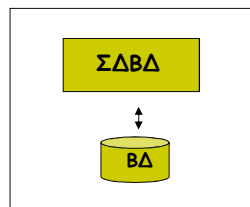
- Πως θα τα μοντελοποιήσουμε, αποθηκεύσουμε, επεξεργαστούμε
- Πως θα διασφαλίσουμε την ορθότητά τους κατά τη διάρκεια αποτυχιών και ταυτόχρονης προσπέλασης από πολλούς χρήστες
- Επίσης, θέματα ασφάλειας, δικαιωμάτων προσπέλασης, ...

Ειδικού σκοπού λογισμικό για τα παραπάνω



Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ): λογισμικό (σύνολο από προγράμματα) για δημιουργία και χρήση μιας βάσης δεδομένων

Database
Management
System (DBMS)



Σύστημα Βάσεων Δεδομένων



Τι θα δούμε σήμερα:

ΕΙΣΑΓΩΓΗ

- Βασικές έννοιες
- Λίγα ιστορικά στοιχεία
- Μια σύνοψη του περιεχομένου του μαθήματος

- Εισαγωγή στο μοντέλο οντοτήτων/συσχετίσεων



Βασικές Έννοιες

Κάποιες λειτουργίες ενός ΣΔΒΔ

- **Ορισμός** μιας βάσης δεδομένων: προδιαγραφή των τύπων, των δομών και των περιορισμών των δεδομένων που θα αποθηκευτούν στη ΒΔ
- **Κατασκευή** μια βάσης δεδομένων: αποθήκευση των ίδιων των δεδομένων
- **Χειρισμός** (manipulation) μιας βάσης δεδομένων: υποβολή ερωτήσεων για την ανάκτηση δεδομένων, ενημέρωση (νέες εισαγωγές, διαγραφές ή τροποποιήσεις)
- **Άλλες λειτουργίες**: Διαμοιρασμός, προστασία από αστοχίες υλικού και λογισμικού, ασφάλεια



Γιατί ένα ΣΔΒΔ;

- Κοινή λειτουργικότητα ήδη υλοποιημένη
- Σωστή υλοποίηση
- Ανεξαρτησία δεδομένων (θα δούμε περισσότερα σε λίγο)
- Μόνιμη αποθήκευση, έλεγχος της επανάληψης πληροφορίας, ορθότητα, έλεγχος συνδρομικότητας, έλεγχος προσπέλασης, ασφάλεια, και άλλα πολλά



Γιατί όχι;

- Επένδυση σε λογισμικό και υλικό, καθώς και για εκπαίδευση
- Η γενικότητα που παρέχει προκαλεί χρονική επιβάρυνση (overhead)
- Δε σας χρειάζονται όσα προσφέρει



Η θέση των ΣΔΒΔ στη στοίβα του λογισμικού συστημάτων



Εμπορικά

- Oracle
- IBM/DB2
- MS SQL-server
- Sybase
- Informix
- (MS Access, ...)

Ελεύθερο Λογισμικό- Open Source

- Postgres (UCB)
- MySQL, mSQL
- miniBase (Wisc)
- Predator (Cornell)
- ...

Σκοπός του μαθήματος



Θα μάθουμε τι είναι τα ΣΔΒΔ:

- Σχεδιασμός και προγραμματισμός μια βάσης δεδομένων χρησιμοποιώντας ένα ΣΔΒΔ

Τότε να τα χρησιμοποιούμε, πως να μοντελοποιούμε δεδομένα σε αυτά, πώς να αποθηκεύουμε δεδομένα, πώς να κάνουμε ερωτήσεις

- Κάποια θέματα υλοποίησης ενός ΣΔΒΔ (το εσωτερικό του)
- Γενικές τεχνικές/αρχές/αλγόριθμους που διέπουν τη διαχείριση δεδομένων

Γενική Εικόνα του Μαθήματος



ΜΕΡΟΣ 1

Σχεδιασμός ΒΔ

Προγραμματισμός ΒΔ

Δημιουργία/Κατασκευή

Εισαγωγή Δεδομένων

Επεξεργασία Δεδομένων

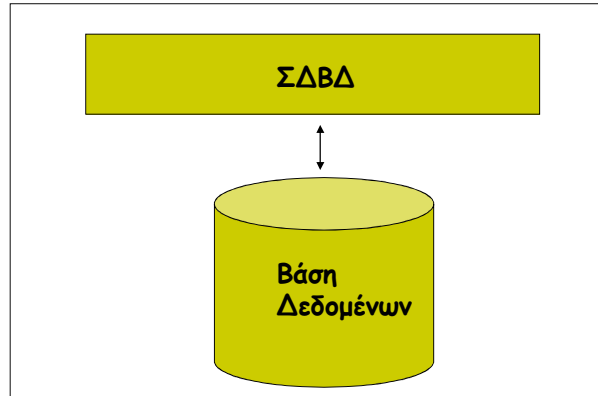
Με χρήση ΣΔΒΔ

ΜΕΡΟΣ 2

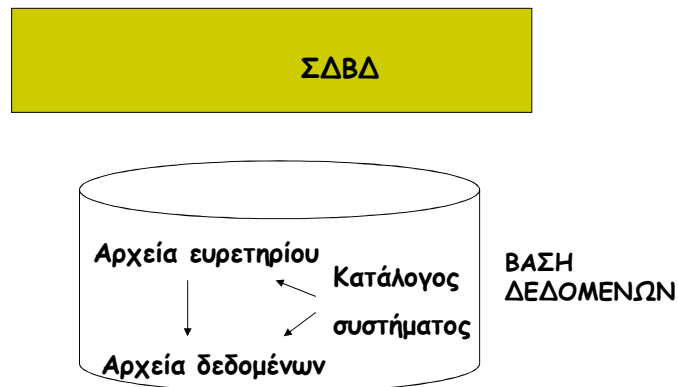
Υλοποίηση ΣΔΒΔ

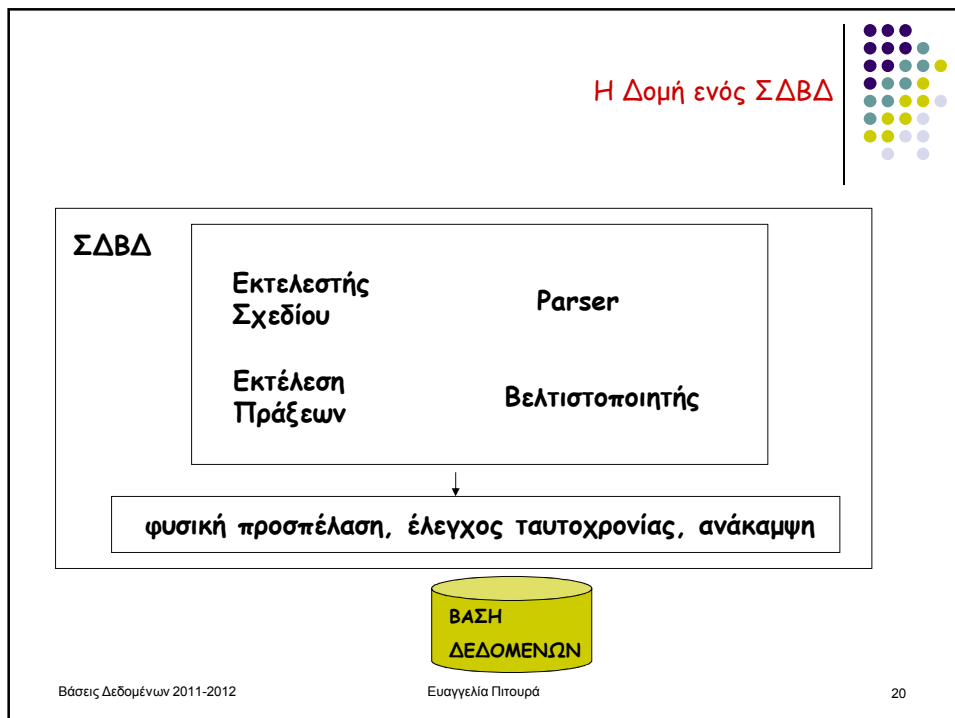
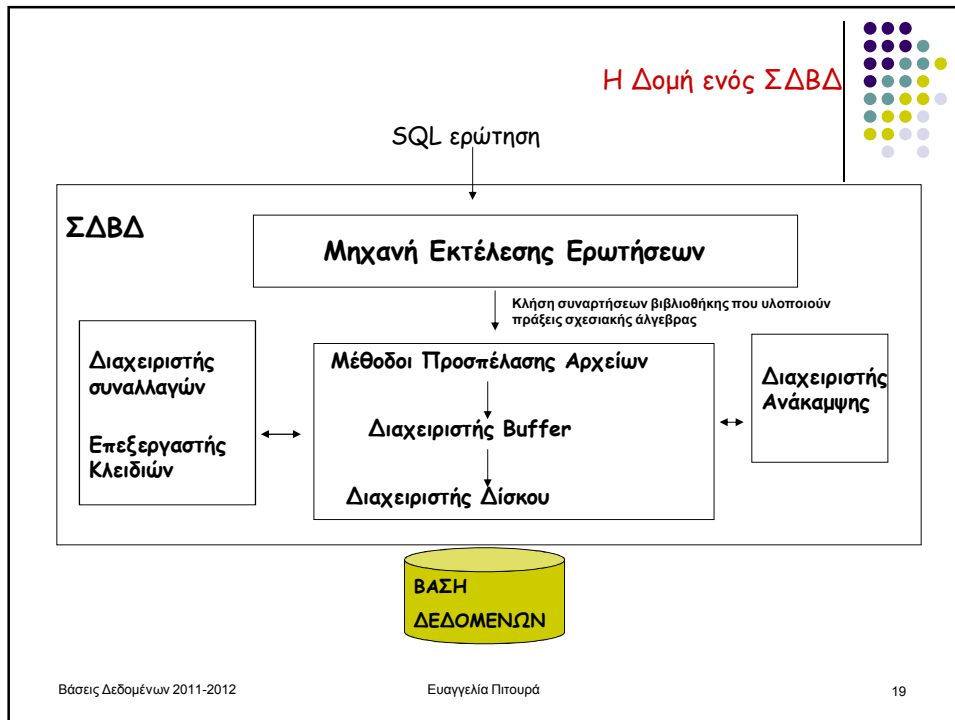
Το εσωτερικό ενός ΣΔΒΔ

Η Δομή ενός ΣΔΒΔ



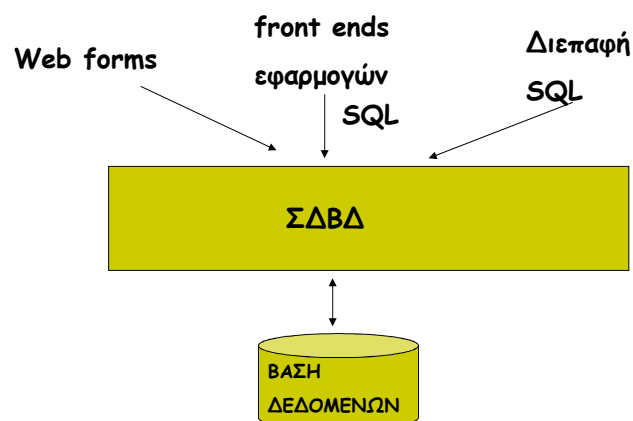
Η Δομή ενός ΣΔΒΔ







- Βασιζόμενες σε μενού (κατάλογο από επιλογές)
- Γραφικών
- Βασιζόμενες σε φόρμες
- Φυσικής γλώσσας
- Για παραμετρικούς χρήστες
- Για το ΔΒΔ





Δεκαετία του 1950

Κάρτες και ταινίες (σειριακή επεξεργασία) - Batch processing

Αρχή του 1960

Γενικευμένη χρήση δίσκων

πρώτο γενικού-σκοπού ΣΔΒΔ: Integrated Data Store (GE)

Charles Bachman (Recipient of the 1st Turing Award, 1973)

network data model (δικτυωτό)

Τέλη του 1960

Information Management System (IMS) IBM

hierarchical data model (Ιεραρχικό)

SABRE Airline Reservation System (AA+IBM, travelocity!!)



1970

Edgar Codd (IBM, San Jose) σχεσιακό μοντέλο δεδομένων (relational data model)

(Recipient of the Turing Award, 1981)

Ερευνητικά Προγράμματα: System R, INGRES - Γλώσσες: SEQUEL, QBE, QUEL

Δεκαετία του 1980

SQL (μέρος του System R)

transaction management (Jim Gray, Turing Award, 1999)

υποσημείωση: Jim Gray gone missing

[Τάσεις: αντικειμενοστραφή, αρχιτεκτονική πελάτη-εξυπηρετή, κατανεμημένες, έμπειρα]



Δεκαετία του 1990

εμπορικά αντικειμενοστραφή συστήματα

[Τάσεις: πολυβάσεις, χωρικές & χρονικές, πολυμέσα, συμπερασματικές, αποθήκες δεδομένων (αναλυτική επεξεργασία), προγραμματισμό πόρων της επιχείρησης (ERP - Enterprise Resource Planning) και της διαχείρισης τους (MRP - Management Resource Planning), Internet]



Δεκαετία του 2000

Σύστημα Διαχείρισης Χρωμοσωμάτων (Human Genome Project)

Σύστημα Παρατήρησης της Γης (Earth Observation System)



Σήμερα

Οι (παραδοσιακές) βάσεις δεδομένων Πέθαναν! Ζήτω οι βάσεις δεδομένων!

NoSQL/CLOUD

Web

Μεγάλος όγκος δεδομένων (+ κυρίως ανάγνωση + κοινωνικά δίκτυα)

Αλλαγές σε υλικό (επεξεργαστές με πολλούς πυρήνες, ιεραρχία αποθήκευσης)

cloud computing, software as service

Μερικά «διαχειριστικά» θέματα

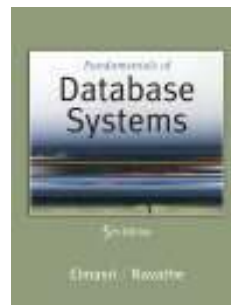


- web σελίδα <http://www.cs.uoi.gr/~pitoura>

«Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων»,
Elmasri&Navathe



«Συστήματα Διαχείρισης Βάσεων Δεδομένων»
Ramakrishnan&Gehrke



Hank Korth, Avi Silberschatz, and S. Sudarshan, *Database System Concepts*, 5th Edition, McGraw-Hill, 2005.

Μερικά «διαχειριστικά» θέματα



Βαθμός

- Ασκήσεις (3 - 4 σύνολα) + Μια προγραμματιστική άσκηση (≥ 4.0)
- Τελικό διαγώνισμα (≥ 4.0)
- Τελικός Βαθμός (≥ 5.0)
- $40\% * (\text{Βαθμός Ασκήσεων} + \text{Προγραμματιστικής}) + 60\% * \text{Βαθμός Τελικού Διαγωνίσματος}$

Όσοι έδωσαν ασκήσεις πέρσι (ακαδημαϊκό έτος 2010-2011) μπορούν να «κρατήσουν» το βαθμό

Οι υπόλοιποι **πρέπει** να τις επαναλάβουν

Μερικές Συμβουλές



- Ναι, πρέπει να **μελετήσετε**
- Καλό θα είναι να *παρακολουθείτε το μάθημα* (τις διαλέξεις, αλλά και το «ρυθμό» του)
- Η ύλη/σειρά στο βιβλίο μπορεί να διαφέρει από το μάθημα - αλλά ό,τι πούμε στο μάθημα και ό,τι υπάρχει στα σχετικά κεφάλαια του βιβλίου αρκεί για να «περάσετε» το μάθημα - *αυτό ισχύει και για τα δύο προτεινόμενα βιβλία*
- Θα πρέπει να σκεφτείτε και να λύσετε προβλήματα «από το μυαλό σας»



- Ανεξαρτησία Δεδομένων
- Αποδοτική Προσπέλαση Δεδομένων
- Περιορισμοί Ακεραιότητας Δεδομένων
- Ταυτόχρονη προσπέλαση
- Ανάρρωση από σφάλματα
- Γρήγορη Ανάπτυξη Εφαρμογών



- Έλεγχος πλεονασμών και συνέπειας
- Εξουσιοδότηση Προσπέλασης
- Ασφάλεια
- Παροχή μόνιμης αποθήκευσης
- Πολλαπλές Διεπαφές και Όψεις



- Απλοί Χρήστες
- Προγραμματιστές εφαρμογών
- Σχεδιαστές βάσεων δεδομένων
- Διαχειριστές συστήματος

Τι θα δούμε σήμερα:

- Βασικές έννοιες
- Λίγα ιστορικά στοιχεία
- Μοντελοποίηση
- Εισαγωγή στο μοντέλο οντοτήτων/συσχετίσεων

Μοντέλα Δεδομένων



Μοντέλο Δεδομένων: ένα σύνολο από έννοιες (δομικά στοιχεία) που μπορούν να χρησιμοποιηθούν για την περιγραφή της δομής της πληροφορίας

- **Υψηλού επιπέδου (εννοιολογικά) μοντέλα**

Υψηλού επιπέδου, περισσότερο αφηρημένη περιγραφή της δομής

Μοντέλο Οντοτήτων/Συσχετίσεων

- **Παραστατικά μοντέλα ή μοντέλα υλοποίησης ή λογικά μοντέλα**

Σχισιακό Μοντέλο, Ιεραρχικό Μοντέλο, Δικτυωτό Μοντέλο

- **Χαμηλού επιπέδου ή φυσικά μοντέλα**

Δρόμος Προσπέλασης

Σχήμα: η περιγραφή της δομής της πληροφορίας που είναι αποθηκευμένη στη βδ με τη χρήση ενός μοντέλου δεδομένων

Η Αρχιτεκτονική Τριών Επιπέδων



Εξωτερική Όψη 1

Εξωτερική Όψη n

Περιγράφει τα αποθηκευμένα δεδομένα με βάση το μοντέλο δεδομένων

Εννοιολογικό Σχήμα

Απεικόνιση

Απεικόνιση

Εσωτερικό (ή φυσικό) Σχήμα



Περιγράφει λεπτομέρειες σχετικά με την αποθήκευση και τους δρόμους προσπέλασης

Πώς οι σχέσεις αποθηκεύονται στο δίσκο, ευρετήρια, κλπ

▪ Η περιγραφή της βάσης δεδομένων περιλαμβάνει ένα σχήμα για καθένα από τα επίπεδα αφάιρησης



Ανεξαρτησία Δεδομένων: αλλαγή του σχήματος ενός επιπέδου **χωρίς** να αλλάξουμε το σχήμα του αμέσως υψηλότερου επιπέδου

- Λογική Ανεξαρτησία Δεδομένων

αλλαγή του εννοιολογικού δεν επηρεάζει τα εξωτερικά σχήματα ή τα προγράμματα εφαρμογών

- Φυσική Ανεξαρτησία Δεδομένων

αλλαγή του εσωτερικού σχήματος χωρίς να χρειάζεται αλλαγή του εννοιολογικού

αλλαγή μόνο της απεικόνισης



Σχήμα της Βάσης

Πρόθεση (intension)

(δομικά στοιχεία + περιορισμοί ακεραιότητας)

Ανάπτυξη (extension)

Στιγμιότυπο της Βάσης (κατάσταση ή σύνολο εμφανίσεων ή σύνολο στιγμιοτύπων)

(αρχική κατάσταση, έγκυρη κατάσταση)