

Εισαγωγή στα Συστήματα Βάσεων Δεδομένων

Τι θα δούμε σήμερα

- I. Σύντομη εισαγωγή στις ΒΔ
- II. Περιγραφή σκοπού και περιεχομένου μαθήματος
- III. Ιστορία των ΣΔΒΔ
- IV. Διαδικαστικά θέματα

Βασικές Έννοιες

Τι είναι μια βάση δεδομένων;

Βάση Δεδομένων: συλλογή από σχετιζόμενα δεδομένα

Που υπάρχουν ΒΔ;



amazon.com.

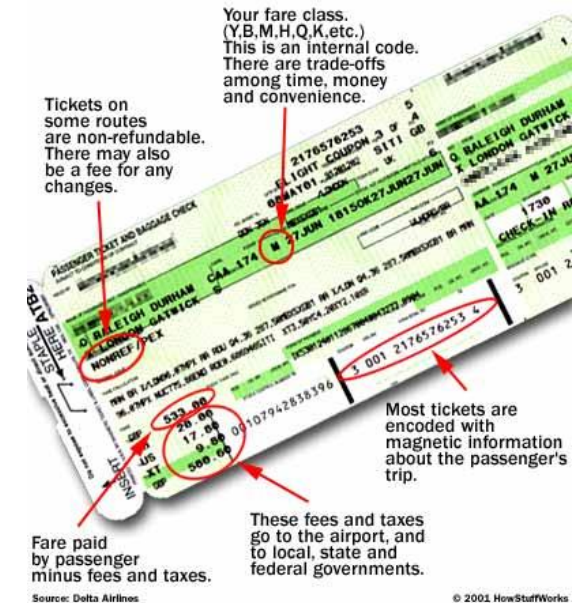


Εύδοξος

Ηλεκτρονική Υπηρεσία Ολοκληρωμένης Διαχείρισης
Συγγραμμάτων και Λοιπών Βοηθημάτων



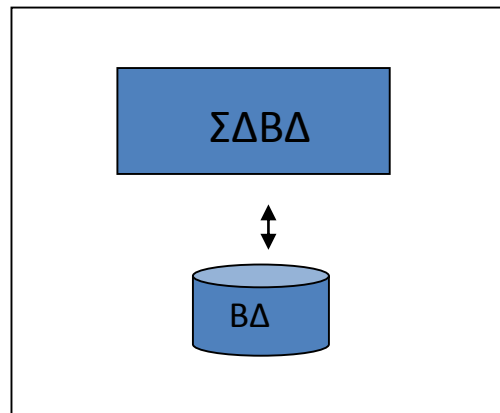
On-Line Δήλωση Μαθημάτων
Students Web



Βασικές Έννοιες

Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ) Database Management System (DBMS)

λογισμικό (σύνολο από προγράμματα) για δημιουργία και χρήση μιας βάσης δεδομένων



**Σύστημα Βάσεων
Δεδομένων**

Μερικά ΣΔΒΔ

Εμπορικά

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

Ελεύθερο Λογισμικό- Open Source

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

Πολλά είδη δεδομένων

Loading "Yahoo! Maps, Driving Directions, and Traffic"

http://maps.yahoo.com/broadband#mvt=m&tp=1&tt=uc+berkeley&trf=0&lon=-122.256203& ncbi genome database

Yahoo! My Yahoo! Mail Make Y! your home page Search: Web Search

YAHOO! LOCAL Maps Sign In New User? Sign Up Maps Home - Dial-Up Map (Original) - Help

▶ GET MAP AND DIRECTIONS Clear
Address, City, State

▼ FIND A BUSINESS ON THE MAP Clear
uc berkeley Search

▶ Browse by Category

- Search Results: uc berkeley Refine
- 1 **Uc Berkeley**
- 2 **UC Berkeley Botanical Garden**
- 3 **UC Berkeley Art Museum**
- 4 **UC Berkeley Foundation**
- 5 Weinstein, Rhona S PhD - **UC Berkeley Pysch Clinic**
- 6 **Capri Motel Berkeley**
- 7 **University Art Museum**
- 8 **Enterprise Rent A Car**
- 9 **Faculty Club**
- 10 **California Engineering Company**
- 11 **Campus Flowers**

Printable Version Send Save Live Traffic Help

Map Hybrid Satellite

Πολλά είδη δεδομένων

NCBI

PubMed Nucleotide Protein Genome Gene Structure PopSet Taxonomy Help

Search for on chromosome(s) assembly All

Show related entries Help FTP Map Viewer home

Entrez Genomes

MapViewer Home

Map Viewer Help Human Maps Help Mouse Maps Help NCBI Handbook

Related Resources Human Genome Guide Genomic Biology Gene OMIM UniGene

Sequence Data Human Genome Sequencing Mouse Genome Sequencing RefSeq

***Homo sapiens* genome view** [BLAST search the human genome](#)
Build 35.1 statistics

1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 X Y MT

Lineage: [Eukaryota](#); [Metazoa](#); [Chordata](#); [Carniata](#); [Vertebrata](#); [Euteleostomi](#); [Mammalia](#); [Eutheria](#); [Euarchontoglires](#); [Primates](#); [Catarrhini](#); [Hominoidea](#); [Homo](#); [Homo sapiens](#)

Γιατί ΣΔΒΔ;

ΒΔ υπάρχουν παντού

- Κοινά θέματα – τι προσφέρουν:
 - ✓ μοντελοποίηση, έλεγχος πλεονασμού, περιορισμοί ακεραιότητας
 - ✓ αποδοτική επεξεργασία ερωτήσεων (queries) (ευρετήρια, βελτιστοποίηση)
 - ✓ μόνιμη αποθήκευση (persistent storage)
 - ✓ ενημέρωση δεδομένων
 - ✓ ορθότητα λειτουργίας: Πως θα διασφαλίσουμε την ορθότητά τους κατά τη διάρκεια αποτυχιών και ταυτόχρονης προσπέλασης από πολλούς χρήστες, ανάκαμψη από σφάλματα
 - ✓ Επίσης: θέματα ασφάλειας, δικαιωμάτων/εξουσιοδότηση προσπέλασης, ...

Γιατί ΣΔΒΔ;

Γιατί χρειαζόμαστε ειδικό λογισμικό;

- Κοινή λειτουργικότητα ήδη υλοποιημένη
- Σωστή υλοποίηση
- Αποδοτική προσπάθεια
- Γρήγορη ανάπτυξη εφαρμογών
- Ανεξαρτησία δεδομένων (θα δούμε περισσότερα)
- Μόνιμη αποθήκευση

Γιατί όχι ΣΔΒΔ;

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

Βασικές Έννοιες

Η θέση των ΣΔΒΔ στη στοίβα του λογισμικού συστημάτων



Σκοπός του Μαθήματος

Τρεις Βασικοί Στόχοι:

1. Να μπορείτε να σχεδιάσετε και υλοποιήσετε ένα σύστημα βάσεων δεδομένων χρησιμοποιώντας ένα ΣΔΒΔ

Τι σημαίνει αυτό:

- Μοντελοποίηση
- Προγραμματισμός (σε SQL)

Καθώς και τη *σχετική θεωρία* στην οποία τα παραπάνω βασίζονται (κανονικός μορφές, σχεσιακή άλγεβρα)

Σκοπός του Μαθήματος

2. Θέματα υλοποίησης ενός ΣΔΒΔ (το εσωτερικό του)
3. Γενικές γνώσεις και δεξιότητες για τη διαχείριση δεδομένων

Μια γενική εικόνα: δημιουργία ΣΒΔ

ΒΗΜΑ 1: Μοντελοποίηση

- Εννοιολογικό Μοντέλο (Μοντέλο Οντοτήτων/Συσχετίσεων)
- Μοντέλο Υλοποίησης (Σχεσιακό μοντέλο)

ΒΗΜΑ 2: Προγραμματισμός/Υλοποίηση

Μοντελοποίηση

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

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

Μοντελοποίηση

- Υψηλού επιπέδου (εννοιολογικά) μοντέλα

Υψηλού επιπέδου, περισσότερο αφηρημένη περιγραφή της δομής

Μοντέλο Οντοτήτων/Συσχετίσεων

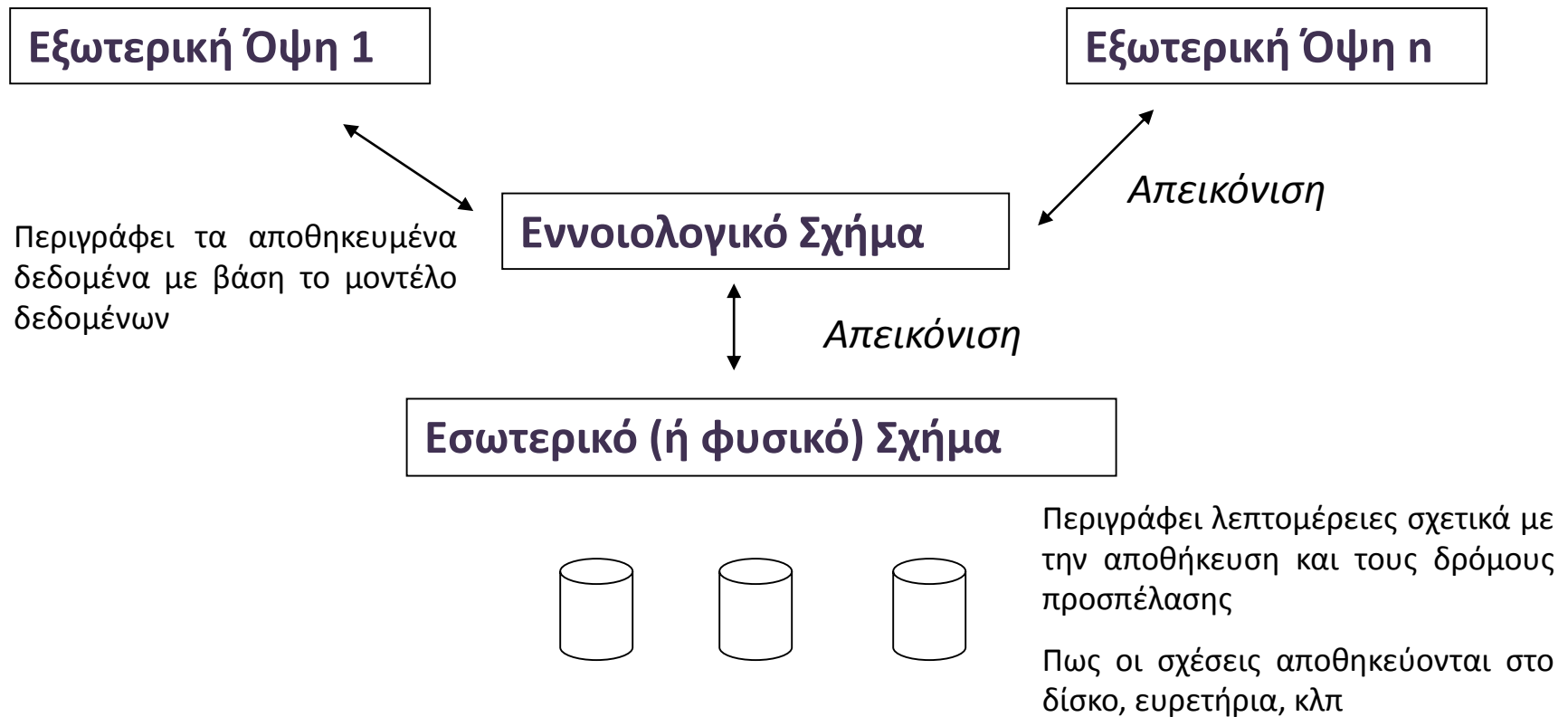
- Παραστατικά μοντέλα ή μοντέλα υλοποίησης ή λογικά μοντέλα

Σχεσιακό Μοντέλο, Ιεραρχικό Μοντέλο, Δικτυωτό Μοντέλο

- Χαμηλού επιπέδου ή φυσικά μοντέλα

Μοντέλα αποθήκευσης

Αρχιτεκτονική Τριών Επιπέδων



- Η περιγραφή της βάσης δεδομένων περιλαμβάνει ένα σχήμα για καθένα από τα επίπεδα αφαίρεσης

Ανεξαρτησία Δεδομένων

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 και Στιγμιότυπο

Σχήμα της Βάσης

Πρόθεση (intension)

Μοντέλο (δομικά στοιχεία + περιορισμοί ακεραιότητας)

Ανάπτυξη (extension)

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

(αρχική κατάσταση, έγκυρη κατάσταση)

Δημιουργία ΣΒΔ

ΒΗΜΑ 1: Μοντελοποίηση (ορισμός σχήματος)

ΒΗΜΑ 2: Υλοποίηση του μοντέλου

Χρήση ΒΔ:

- Εισαγωγή στοιχείων (δημιουργία του αρχικού στιγμιότυπου)
- Διατύπωση ερωτήσεων
- Εισαγωγή/διαγραφή/τροποποίηση δεδομένων

Γλώσσες ΣΔΒΔ (SQL)

Γλώσσα Ορισμού

create table R(A1 T1, A2, T2, ...)

Γλώσσα Χειρισμού/Επεξεργασίας Δεδομένων

εισαγωγή, διαγραφή, τροποποίηση δεδομένων

insert/delete/update

διατύπωση ερωτημάτων

select Γνωρίσματα

from Πίνακες

where Συνθήκη

Γλώσσες ΣΔΒΔ

Γλώσσες Ερωτήσεων (Query Languages)

- δεν είναι γλώσσα προγραμματισμού (δυνατότητα εμφύτευσης σε μια γλώσσα υψηλού επιπέδου)
- δηλωτικές (μη διαδικαστικές)

Βασικές Έννοιες (ανασκόπηση)

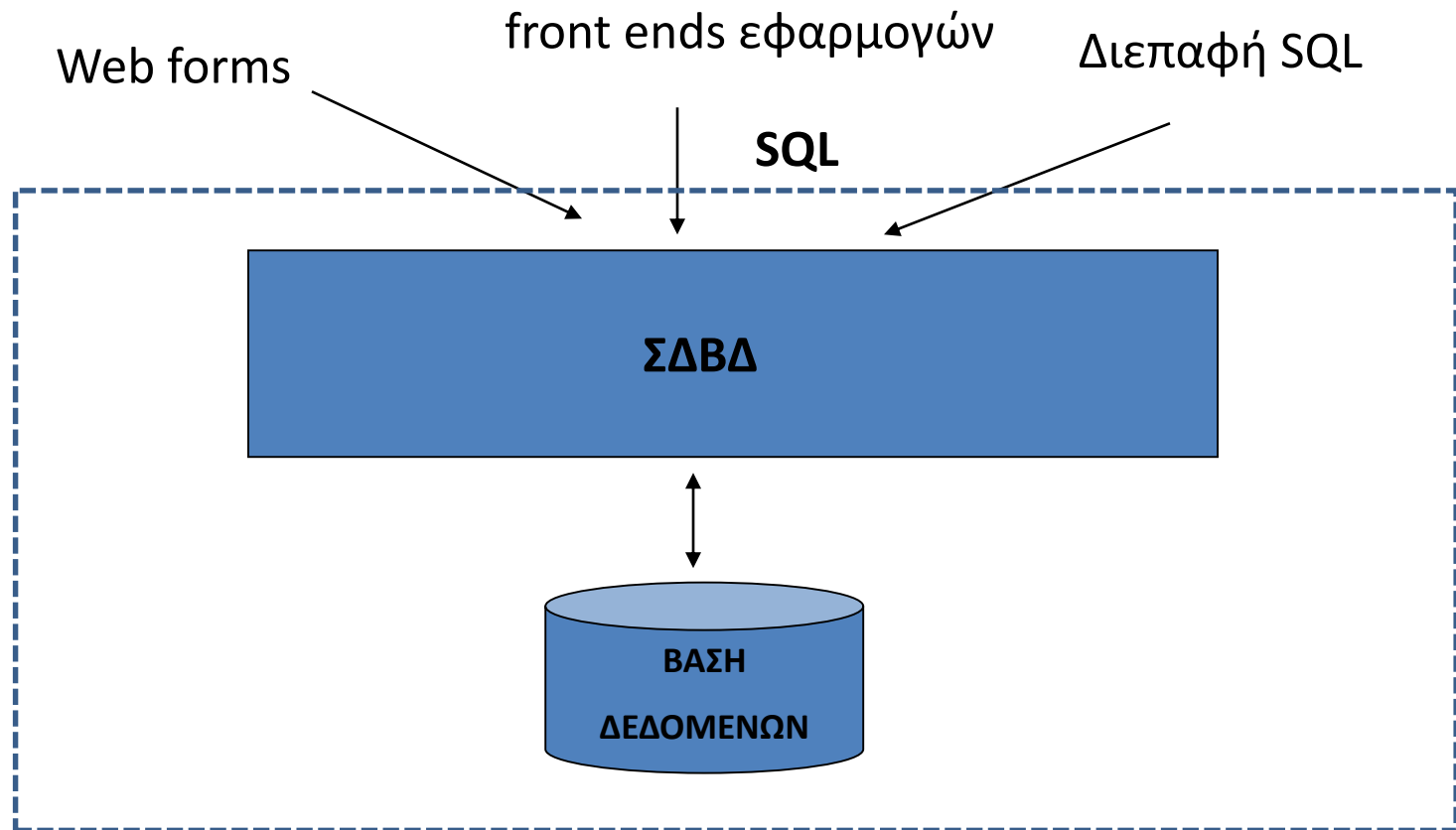
Κάποιες λειτουργίες ενός ΣΔΒΔ

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

Χρήστες

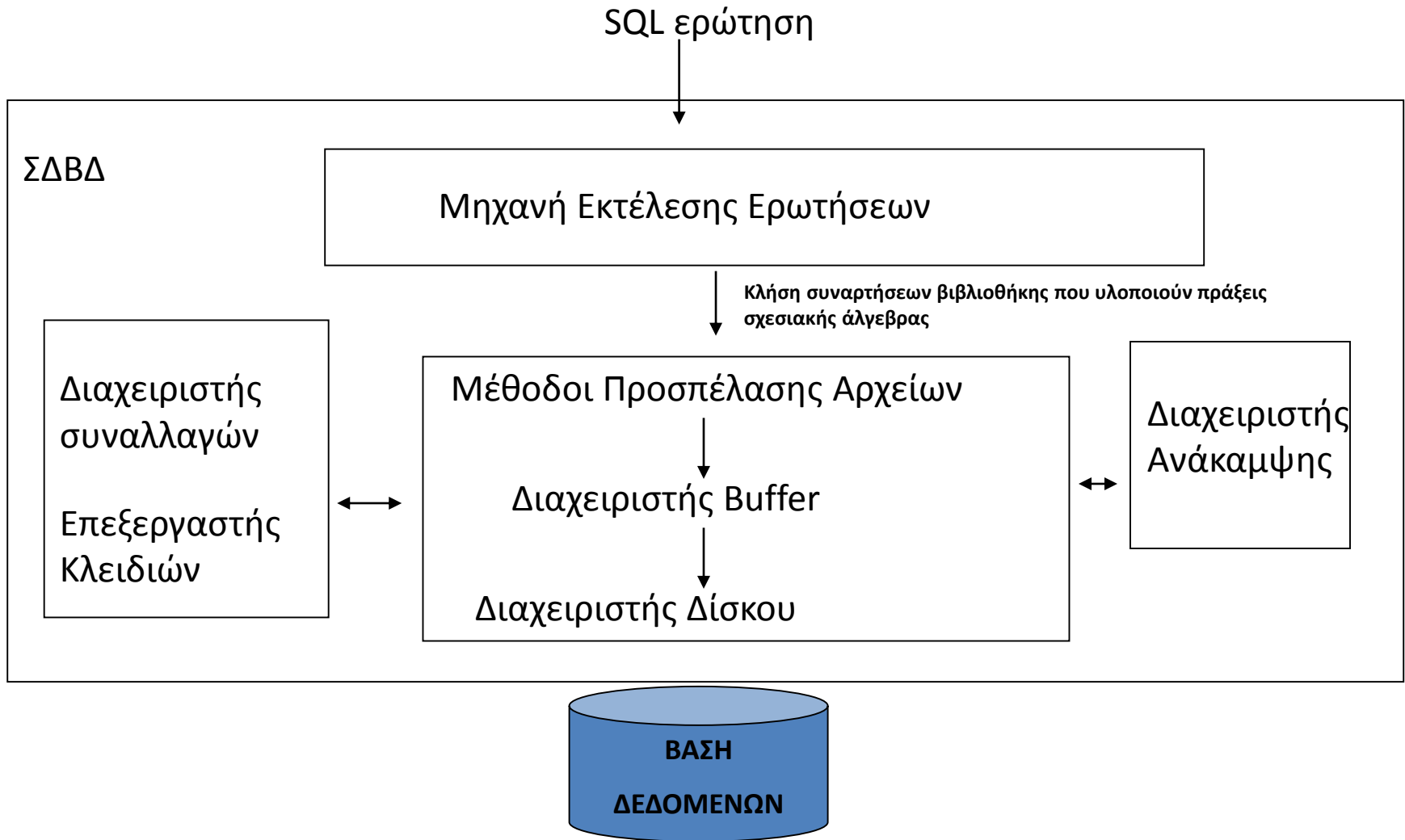
- Απλοί Χρήστες
- Προγραμματιστές εφαρμογών
- Σχεδιαστές βάσεων δεδομένων
- Διαχειριστές συστήματος
- Δημιουργοί ΣΔΒΔ

Διεπαφές

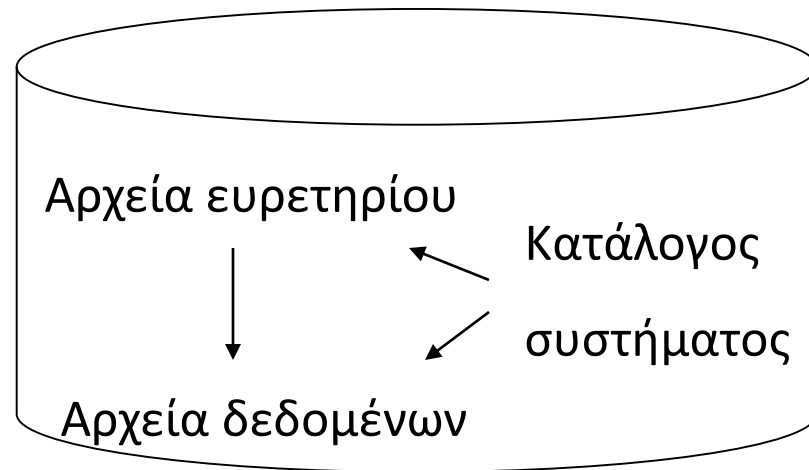


- Βασιζόμενες σε μενού (κατάλογο από επιλογές), Γραφικών, Βασιζόμενες σε φόρμες, Φυσικής γλώσσας, Για παραμετρικούς χρήστες, Για το ΔΒΔ

Το εσωτερικό ενός ΣΔΒΔ



Το εσωτερικό ενός ΣΔΒΔ



ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Ιστορία

Δεκαετία του 1950

Κάρτες και ταινίες (σειριακή επεξεργασία) – Batch processing

Αρχή του 1960

Γενικευμένη χρήση δίσκων

πρώτο **γενικού-σκοπού** ΣΔΒΔ (διαχωρισμός της λειτουργικότητας διαχείρισης δεδομένων από τις εφαρμογές): 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!!)

Ιστορία

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*]

Ιστορία

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)

Ιστορία

Σήμερα

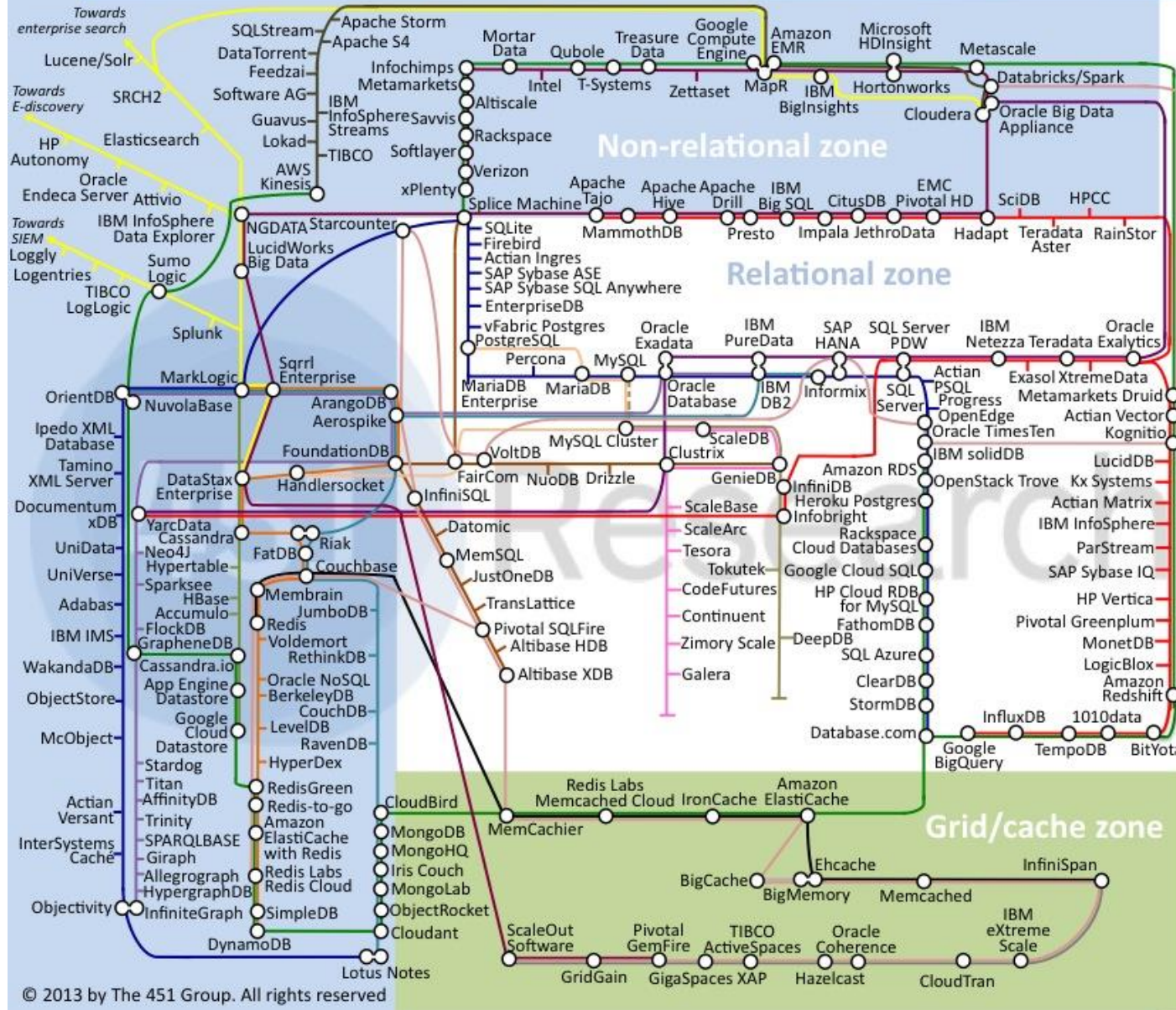
- Μεγάλος όγκος δεδομένων (BIG DATA)
- Αλλαγές σε υλικό (επεξεργαστές με πολλούς πυρήνες, ιεραρχία αποθήκευσης)
- cloud computing, software as service

Κίνηση NoSQL

Database Landscape Map – February 2014

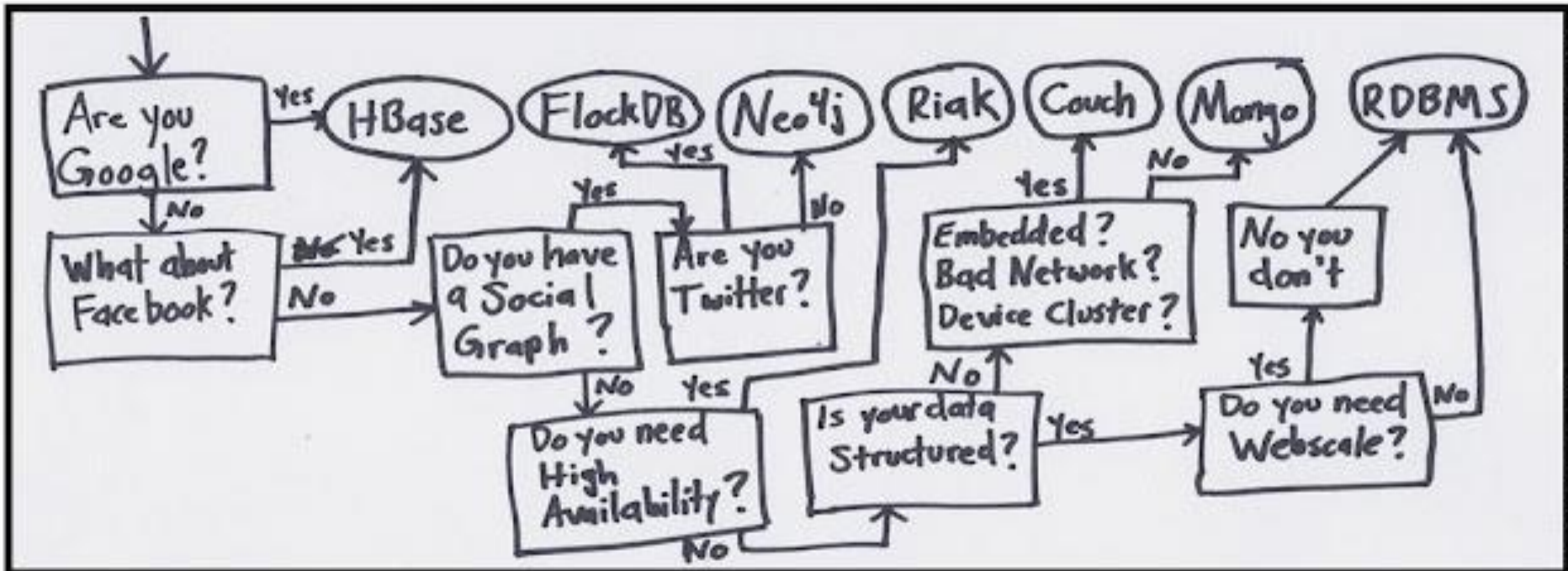
451 Research

- Key:**
- General purpose
 - Specialist analytic
 - -as-a-Service
 - NoSQL extension
 - Big Tables
 - Graph
 - Document
 - Key value stores
 - Key value direct access
 - Hadoop
 - NewSQL extension
 - MySQL storage engines
 - Advanced clustering/sharding
 - New SQL databases
 - Data caching
 - Data grid
 - Search
 - Appliances
 - Off-heap memory
 - In-memory
 - Stream processing



© 2013 by The 451 Group. All rights reserved

Τι ΣΔΒΔ να χρησιμοποιήσω;



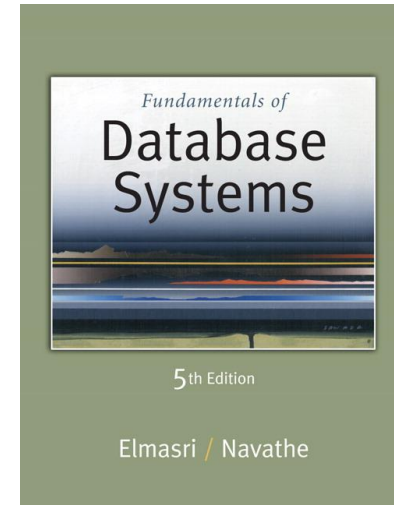
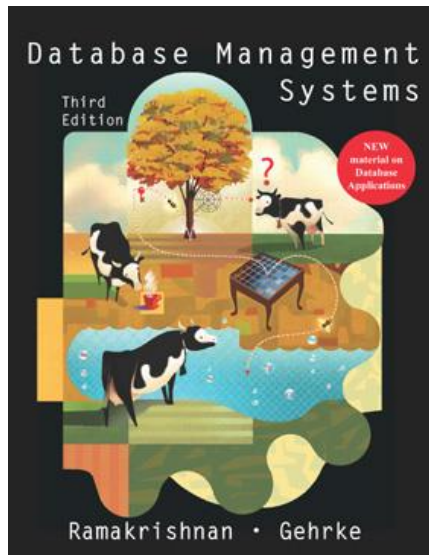
©2011 ERIC REDMOND

CRIPDCOMIC.COM

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.

Διαχειριστικά Θέματα

Βαθμός

- Ασκήσεις (3 σύνολα) + Μια προγραμματιστική άσκηση (≥ 4.0)
- Τελικό διαγώνισμα (≥ 4.0)
- Τελικός Βαθμός (≥ 5.0)
- $35\% * (\text{Βαθμός Ασκήσεων} + \text{Προγραμματιστικής}) +$
 $65\% * \text{Βαθμός Τελικού Διαγωνίσματος}$

Όσοι έδωσαν ασκήσεις τα 2 προηγούμενα Ακαδημαϊκά Έτη μπορούν να «κρατήσουν» το βαθμό τους

Οι υπόλοιποι **πρέπει** να τις επαναλάβουν

1. ΔΕΝ υπάρχει «μεθοδολογία ασκήσεων»
2. Καλό θα είναι να παρακολουθείτε το μάθημα (τις διαλέξεις, αλλά και το «ρυθμό» του)
3. Οι διαφάνειες ΔΕΝ αντικαθιστούν το βιβλίο (είναι συμπληρωματικές σε αυτό)
4. Κάποιες ασκήσεις λύνονται ΜΟΝΟ στον πίνακα σκόπιμα

Και αν δεν περάσω
το μάθημα;

HOW TO WRITE A CV



Leverage the NoSQL boom

Ερωτήσεις;