



Βάσεις δεδομένων

(1^ο μάθημα)

Ηρακλής Βαρλάμης
varlamis@hua.gr



Ιστοσελίδες και επικοινωνία

- **Σελίδα** μαθήματος:
- <http://eclass.hua.gr/courses/DIT105>
 - Γραφτείτε στο μάθημα για
 - να κατεβάζετε τις διαφάνειες διαλέξεων,
 - να βρίσκετε ανακοινώσεις, συνδέσμους, κλπ.
 - να χρησιμοποιείτε τις περιοχές συζητήσεων.
 - **να αποκτήσετε πρόσβαση στη ΒΔ Oracle**
- **Διδάσκοντες:** Η. Βαρλάμης - Β. Ευθυμίου
- **Εργαστήρια:** + Χ. Σαρδιανός
- **Επικοινωνία**
 - Προτιμήστε την κατάλληλη **περιοχή συζητήσεων** διαφορετικά ρωτήστε απευθείας στο μάθημα ή στο εργαστήριο
 - **Γραφείο:** 5.1 (Βαρλάμης)
 - Μηνύματα ηλεκτρονικού ταχυδρομείου **μόνο αν είναι αδύνατη η χρήση των ομάδων συζητήσεων**



Διαλέξεις – Εργαστήρια - Ασκήσεις

- Τα μαθήματα γίνονται
 - κάθε Παρασκευή 9:15-11:45 στο Αμφιθέατρο
- Οι διαφάνειες των διαλέξεων: στη σελίδα του μαθήματος.
- Τα εργαστήρια γίνονται σε 4 τμήματα
 - Πέμπτη 9:00-10:30
 - Πέμπτη 10:30-12:00
 - Πέμπτη 12:00-13:30
 - Πέμπτη 13:30-15:00
 - Θα χρησιμοποιηθεί η Oracle 11gR2 (οδηγίες σύνδεσης στο e-class)
 - Προαιρετικά: Μπορείτε να εγκαταστήσετε σπίτι την Oracle 11g Express Edition κατεβάζοντάς την από εδώ:
<http://www.oracle.com/technetwork/database/express-edition/overview/index.html>
- Εργασίες
 - 2 υποχρεωτικές εργασίες σε ομάδες των τριών, σταθερές για όλες τις εργασίες
 - 1 εξέταση εντός του εργαστηρίου



Βαθμολογία

- Βαθμολογία (ισχύει μέχρι και το Σεπτέμβριο):
 - Τελική εξέταση: 60% [βαθμός ≥ 5]
 - Εργασίες: 40%=10+10+20 [βαθμός ≥ 5 , **υποχρεωτική**]
 - Πιθανό bonus ως 10% σε καλές εργασίες
- Τελική εξέταση: με ανοιχτό βιβλίο και σημειώσεις μαθήματος
- Κατοχύρωση:
 - Εργασίες δίνετε **μόνο** μέσα στο εξάμηνο και ο βαθμός τους κατοχυρώνεται μέχρι και το Σεπτέμβριο
 - Τίποτε δεν κατοχυρώνεται για επόμενη χρονιά!



Βιβλιογραφία

- Βιβλία:
 - Βασικές αρχές για τα Συστήματα Βάσεων Δεδομένων, Jeffrey Ullman, Jennifer Widom, A First Course in Database Systems, (μεταφραστική επιμέλεια Β. Βερύκιος), Εκδόσεις Κλειδάριθμος, 2008
 - Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων, 6η Έκδοση, R. Elmasri S. B. Navathe, Fundamentals of Database Systems, Addison Wesley (μεταφραστική επιμέλεια Μ. Χατζόπουλος), Εκδόσεις Δίαυλος, 2012
- Άλλα προτεινόμενα βιβλία στη σελίδα του μαθήματος, και στη βιβλιοθήκη



Απαιτούμενα και καλές πρακτικές

- **Καλές πρακτικές**

- Παρακολούθηση διαλέξεων και τήρηση σημειώσεων
- Μελέτη διαφανειών και σημειώσεων μετά από κάθε διάλεξη
- Εξάσκηση στον υπολογιστή κατά το διάβασμα (θα μπορείτε να έχετε πρόσβαση και από το σπίτι)
- Ενεργή συμμετοχή στο μάθημα και τα εργαστήρια
- Παρακολούθηση των περιοχών συζητήσεων
- Σοβαρή ενασχόληση με τις εργασίες
- Επανάληψη μέσω των διαφανειών πριν τις εξετάσεις



Περιεχόμενα

- Το περιβάλλον της Βάσης Δεδομένων, σύστημα αρχείων,
- Διαδικασία ανάπτυξης της βάσης δεδομένων,
- Μοντέλο Οντοτήτων-Συσχετίσεων (ΟΣ),
- Λογική σχεδίαση βάσης δεδομένων και το Σχεσιακό Μοντέλο,
- Μοντέλο Σχεσιακής Βάσης Δεδομένων, Κανονικοποίηση,
- Φυσικός Σχεδιασμός Βάσεων Δεδομένων,
- Εισαγωγή στην SQL, SQL ενέργειες
- Δημιουργία σχήματος, Εισαγωγή δεδομένων, ερωτήσεις
- NoSQL Μοντέλα
- Διαχείριση της Βάσης Δεδομένων,
- Αντικειμενοστραφή Μοντέλα,
- Το μοντέλο Client / Server και Βάσεις Δεδομένων,
- Κατανεμημένες βάσεις δεδομένων, Αποθήκες δεδομένων, Εξόρυξη γνώσης από δεδομένα



Εισαγωγή – Βασικές έννοιες



Εισαγωγή - Ορισμοί

- Δεδομένα (data)
 - Γνωστά γεγονότα/ περιστατικά που καταγράφουμε
- Βάση δεδομένων (database)
 - Μία συλλογή δεδομένων που συσχετίζονται και συλλέγονται για κάποιο σκοπό
 - Αναπαριστά μέρος των στοιχείων που μπορούμε να συλλέξουμε για μια εφαρμογή (πεδίο αναφοράς)
- Σύστημα διαχείρισης βάσεων δεδομένων (database management system -DBMS)
 - Συλλογή από προγράμματα (λογισμικό γενικής χρήσης) για την δημιουργία και συντήρηση μιας ή περισσοτέρων βάσεων δεδομένων

Παράδειγμα Βάσης Δεδομένων

ΦΟΙΤΗΤΗΣ	Επώνυμο	Αρ_Μητ	Έτος	Τμήμα
	Βασιλείου	23	1	ΠΛΗΡ
	Ιωάννου	26	1	ΠΛΗΡ

ΒΑΘΜΟΛΟΓΙΑ	Αρ_Μητ	Κωδικός_Μαθ	Βαθμός
	23	1	8
	26	2	7

ΜΑΘΗΜΑ	Όνομα_Μαθ	Κωδικός_Μαθ	Μονάδες	Τμήμα
	Βάσεις Δεδομένων	1	5	ΠΛΗΡ
	Γλώσσες Προγραμματισμού	2	4	ΠΛΗΡ

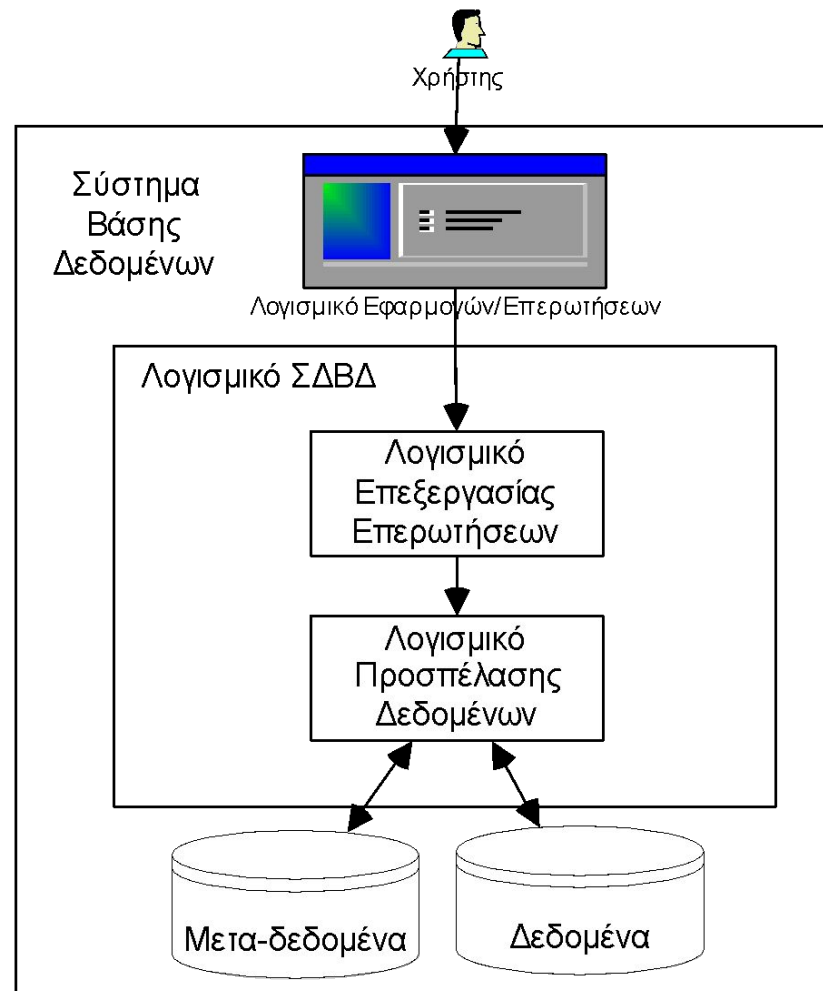
Όνομα	Τύπος	Περιορισμός
Επώνυμο	String	Υποχρ LEN:50
Αρ_Μητ	Integer	MAX:9999
Έτος	Integer	MAX:5
Τμήμα	String	LEN:4

Δεδομένα

Μετα-δεδομένα

Λειτουργίες ΒΔ

- Ορισμός ΒΔ: Προδιαγραφές για τον τύπο, τη δομή και τους περιορισμούς των δεδομένων
- Κατασκευή ΒΔ: Αποθήκευση των δεδομένων στο αποθηκευτικό μέσο
- Χειρισμός ΒΔ: Υποβολή ερωτήσεων (επερωτήσεων) προς τη ΒΔ για εισαγωγή και ανάκτηση δεδομένων, παραγωγή αναφορών, ενημέρωση δεδομένων κλπ.





Συστήματα Αρχείων

- Πρόγονοι των ΣΔΒΔ
- Δυσκολία στην πρόσβαση στα δεδομένα
 - Τα δεδομένα αποθηκεύονταν σε αρχεία
 - Μια γλώσσα προγραμματισμού (COBOL) επέτρεπε τη διαχείρισή τους
- Πλεονασμός δεδομένων
 - Κάθε εφαρμογή διατηρούσε δικό της αρχείο δεδομένων
- Απομόνωση δεδομένων
 - Μια ξεχωριστή εφαρμογή προσπέλασης δεδομένων για κάθε αρχείο.
 - Αν άλλαζε η δομή του αρχείου (π.χ. το μέγεθος ενός πεδίου) τότε άλλαζε και η εφαρμογή



Πλεονεκτήματα ΣΔΒΔ

- Αυτοπεριγραφή
 - Τα μεταδεδομένα περιγράφουν τη δομή των αρχείων, των δεδομένων που περιέχουν και των περιορισμών σε αυτά
- Αφαίρεση δεδομένων
 - Μας ενδιαφέρει η λογική οργάνωση των δεδομένων και όχι ο φυσικός τρόπος αποθήκευσής τους σε αρχεία
- Ανεξαρτησία προγραμμάτων/δεδομένων
 - Αν αλλάξει το μέγεθος ενός πεδίου δεδομένων το ΣΔΒΔ προσαρμόζει αυτόματα τη φυσική αποθήκευσή του
- Ανεξαρτησία προγραμμάτων/πράξεων
 - Σε αντικειμενοστραφείς ΒΔ μπορούμε να ενσωματώσουμε στα δεδομένα πράξεις. Στη συνέχεια μπορούμε να καλούμε τις πράξεις αυτές ανεξάρτητα από το πώς υλοποιούνται



Πλεονεκτήματα ΣΔΒΔ

- Πολλαπλές όψεις
 - Στα ίδια δεδομένα ορίζουμε διαφορετικές όψεις (εικονικά δεδομένα) για κάθε χρήστη ή ομάδα χρηστών
- Αυξημένη ασφάλεια
 - Με τον ορισμό χρηστών και δικαιωμάτων πρόσβασης σε δεδομένα και λειτουργίες
- Υποστήριξη δοσοληψιών πολλών χρηστών
 - Επιτυγχάνεται με εφαρμογές ελέγχου ταυτόχρονης προσπέλασης και εφαρμογές επεξεργασίας δοσοληψιών (ταυτόχρονες δοσοληψίες εκτελούνται χωρίς παρεμβολές)
- Μείωση όγκου δεδομένων
 - Αποφεύγονται οι επαναλήψεις

Χρήστες

- Διαχειριστής (DBA): Εξουσιοδοτεί την προσπάθεια στη ΒΔ, διορθώνει θέματα ασφάλειας και χαμηλής απόδοσης, διαμοιράζει τους πόρους υλικού και λογισμικού, συντηρεί τη ΒΔ
- Σχεδιαστής: Επικοινωνεί με τους τελικούς χρήστες, καταγράφει απαιτήσεις και προσδιορίζει τη δομή των δεδομένων που θα αποθηκευτούν, αλλά και τις όψεις που θα δημιουργηθούν
- Τελικοί χρήστες:
 - Περιστασιακοί: Υποβάλλουν διάφορες επερωτήσεις σε SQL (π.χ. για τη δημιουργία ειδικών αναφορών)
 - Απλοϊκοί ή παραμετρικοί: Άνθρωποι ή εφαρμογές που πραγματοποιούν τυποποιημένες ερωτήσεις συχνά (π.χ. για την αναζήτηση με συγκεκριμένα κριτήρια)
 - Αναλυτές και προγραμματιστές: Εντάσσουν τη ΒΔ στις εφαρμογές
 - Εξειδικευμένοι: Επεμβαίνουν στο σχήμα και τις λειτουργίες του ΣΔΒΔ για να ικανοποιήσουν πολύπλοκες απαιτήσεις
 - Μεμονωμένοι: Διατηρούν προσωπικές ΒΔ, χρησιμοποιούν έτοιμα πακέτα προγραμμάτων (π.χ. για προσωπικά φορολογικά δεδομένα)



Σκοποί της ΒΔ

- Έλεγχος των πλεονασμών (redundancy)
- Έλεγχος της συμβατότητας (consistency) των δεδομένων
- Έλεγχος ορθότητας (integrity) δεδομένων με χρήση περιορισμών
- Περιορισμός της μη εξουσιοδοτημένης προσπέλασης
- Μόνιμη αποθήκευση αντικειμένων προγραμμάτων και δομών δεδομένων (π.χ. δρόμων προσπέλασης)
- Παραγωγή συμπερασμάτων με χρήση επαγωγικών κανόνων (που αποθηκεύονται μαζί με τη ΒΔ)
- Παράσταση συσχετίσεων μεταξύ δεδομένων π.χ. ένας μαθητής έχει ένα ή περισσότερους βαθμούς
- Τήρηση αντιγράφων και παροχή μηχανισμών ανάκαμψης



Αρχιτεκτονική ΒΔ

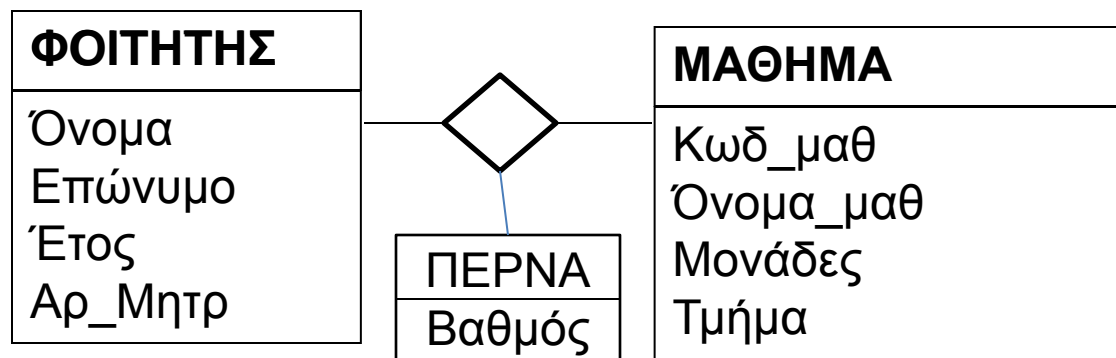


Μοντέλο δεδομένων

- Ορίζει
 - Έννοιες για την περιγραφή της δομής της ΒΔ: τύποι δεδομένων, συσχετίσεις, περιορισμοί για τα δεδομένα
 - Βασικές πράξεις για ανάκτηση και ενημέρωση δεδομένων
 - Πράξεις ορισμένες από το χρήστη πάνω στα δεδομένα (π.χ. μέσος όρος κλπ)

Κατηγορίες μοντέλων δεδομένων

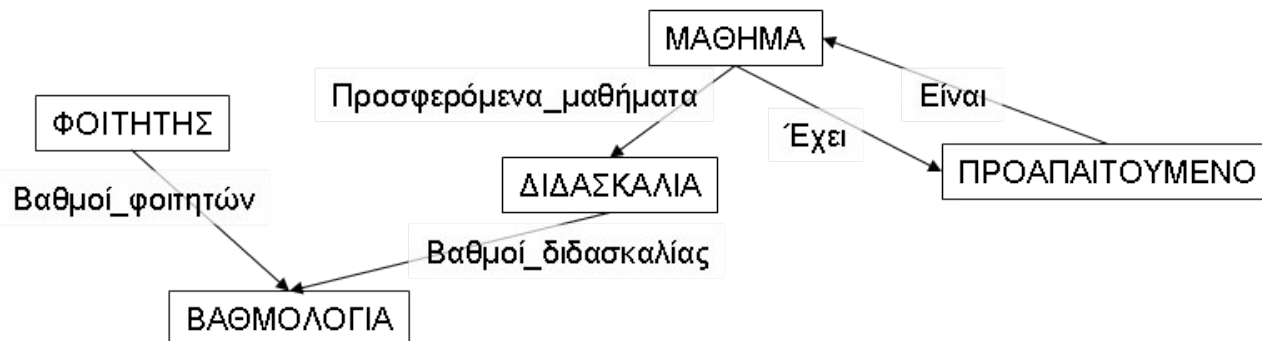
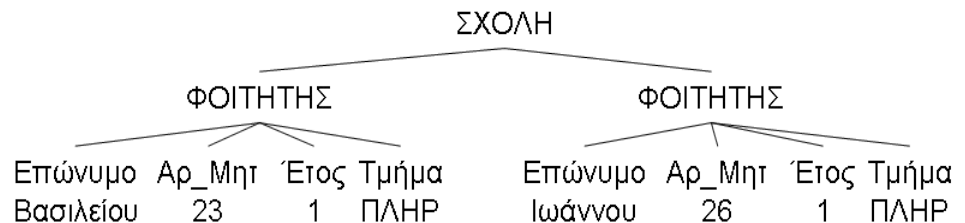
- Υψηλού επιπέδου (εννοιολογικά) μοντέλα
 - Οντότητα (entity): Ένα αντικείμενο, μια έννοια του πραγματικού κόσμου
 - Γνώρισμα (attribute): Μια ενδιαφέρουσα ιδιότητα που περιγράφει την οντότητα
 - Συσχέτιση (relationship): Μια αλληλεπίδραση μεταξύ δύο ή περισσότερων οντοτήτων



Κατηγορίες μοντέλων δεδομένων

- Παραστατικά μοντέλα ή μοντέλα υλοποίησης
 - Βρίσκονται ανάμεσα στο εννοιολογικό και το φυσικό επίπεδο
 - Σχεσιακό, Δικτυωτό, Ιεραρχικό
 - Αντικειμενοστραφές

ΦΟΙΤΗΤΗΣ	Επώνυμο	Αρ_Μητ	Έτος	Τμήμα
	Βασιλείου	23	1	ΠΛΗΡ
	Ιωάννου	26	1	ΠΛΗΡ



Κατηγορίες μοντέλων δεδομένων

- Χαμηλού επιπέδου (φυσικά) μοντέλα
 - Περιγράφουν πώς αποθηκεύονται τα δεδομένα στον υπολογιστή

ΦΟΙΤΗΤΗΣ	Επώνυμο	Αρ_Μητ	Έτος	Τμήμα
	Βασιλείου	23	1	ΠΛΗΡ
	Ιωάννου	26	1	ΠΛΗΡ



Φοιτητής_Αρ_Μητ

23	0	26	30
----	---	----	----

Φοιτητής

23	Βασιλείου	1	ΠΛΗΡ	26	Ιωάννου	1	ΠΛΗΡ
----	-----------	---	------	----	---------	---	------

Σχήμα - Στιγμιότυπο

- **Σχήμα ΒΔ:** Προσδιορίζεται κατά το σχεδιασμό της βάσης δεδομένων και δεν αναμένεται να αλλάζει συχνά.
- Δίνει τη δομή των δεδομένων, ονόματα τύπων, τύπους

ΦΟΙΤΗΤΗΣ	ΠΡΟΠΟΙΗΜΕΝΟΝ	Αρ_Μητ	Έτος	Τμήμα
----------	--------------	--------	------	-------

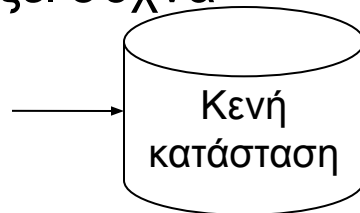
ΒΑΘΜΟΛΟΓΙΑ	Αρ_Μητ	Κωδικός_Μαθ	Βαθμός
------------	--------	-------------	--------

ΜΑΘΗΜΑ	Όνομα_Μαθ	Κωδικός_Μαθ	Μονάδες	Τμήμα
--------	-----------	-------------	---------	-------

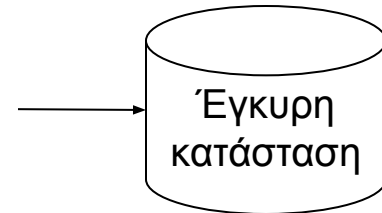
Δομικά στοιχεία του σχήματος

- **Στιγμιότυπο:** Τα πραγματικά δεδομένα που αποθηκεύονται μια συγκεκριμένη στιγμή στη ΒΔ (**εγγραφές**). Η κατάσταση της ΒΔ, η οποία αλλάζει συχνά

Ορισμός
Σχήματος



Εισαγωγή/
Διαγραφή
δεδομένων

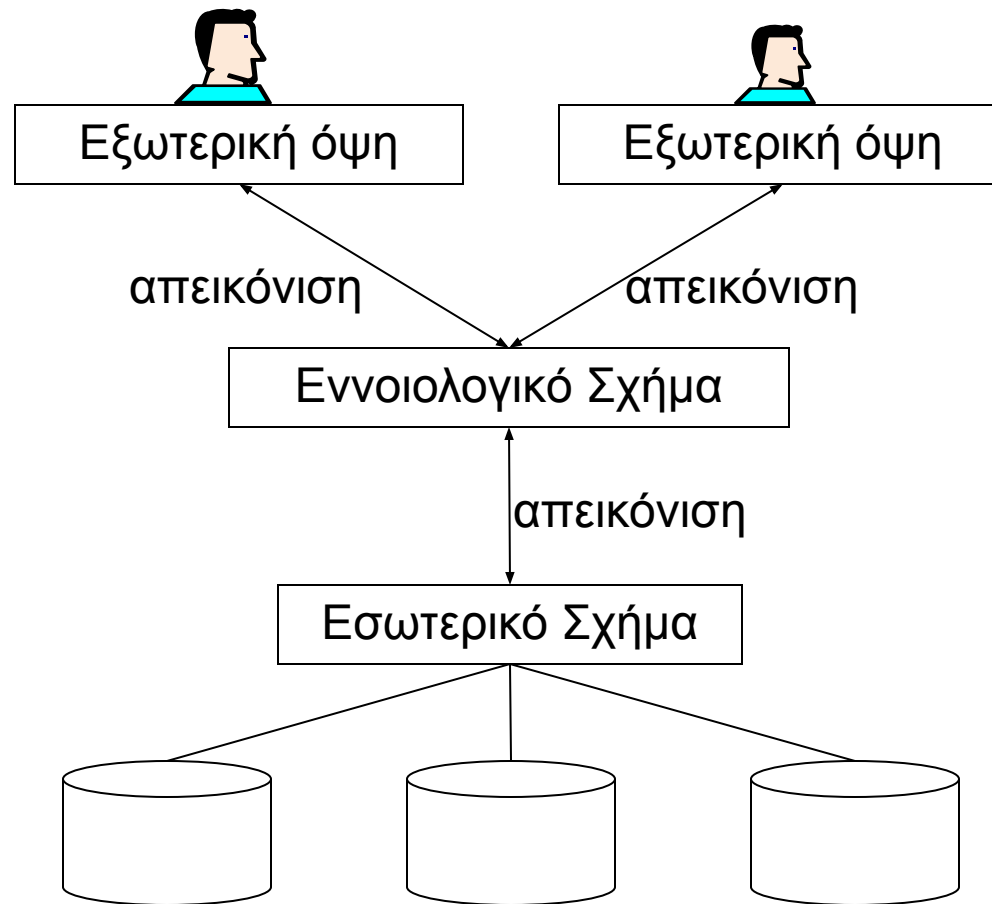




Αρχιτεκτονική τριών σχημάτων

- Εξωτερικό σχήμα ή όψεις χρηστών
 - Μία για κάθε χρήστη ή ομάδα. Αποκρύπτει μέρος της ΒΔ από κάθε ομάδα. Χρησιμοποιεί μοντέλο δεδομένων υλοποίησης ή υψηλού επιπέδου
- Εννοιολογικό σχήμα
 - Περιγράφει τη δομή της ΒΔ για μια ομάδα χρηστών. Χρησιμοποιεί μοντέλο δεδομένων υλοποίησης ή υψηλού επιπέδου (οντότητες, τύπους δεδομένων, συσχετίσεις, περιορισμούς)
- Εσωτερικό σχήμα
 - Περιγράφει τη δομή φυσικής αποθήκευσης της ΒΔ. Χρησιμοποιεί χαμηλού επιπέδου μοντέλο δεδομένων

Σχηματικά





Ανεξαρτησία δεδομένων

- Τα δεδομένα υπάρχουν μόνο μια φορά στο φυσικό επίπεδο της ΒΔ
- Τα υπόλοιπα επίπεδα διαθέτουν πληροφορίες απεικόνισης
- **Λογική ανεξαρτησία δεδομένων**
 - Μπορούμε να αλλάξουμε το εννοιολογικό σχήμα χωρίς να χρειάζεται να αλλάξουμε τα εξωτερικά σχήματα.
Αρκεί να αλλάξει ο ορισμός των όψεων (οι απεικονίσεις).
- **Φυσική ανεξαρτησία δεδομένων**
 - Μπορούμε να αλλάξουμε το εσωτερικό σχήμα χωρίς να χρειάζεται να αλλάξουμε το εννοιολογικό ή τα εξωτερικά.

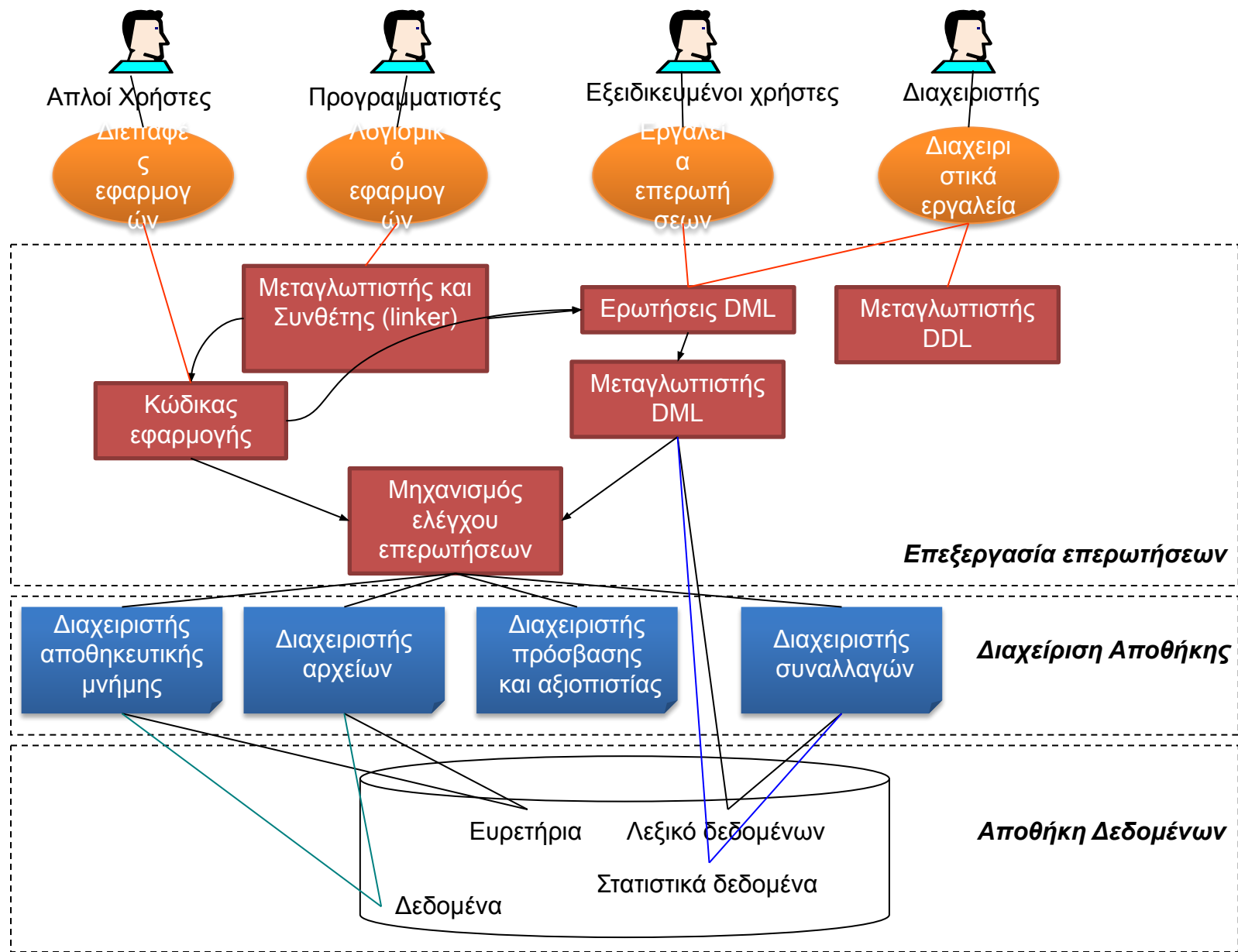


Γλώσσες

- Γλώσσα ορισμού δεδομένων (data definition language – DDL): Για τον ορισμό του εννοιολογικού σχήματος της ΒΔ
- Γλώσσα αποθήκευσης δεδομένων (storage definition language – SDL): Για τον ορισμό του εσωτερικού σχήματος της ΒΔ
- Γλώσσα ορισμού όψεων (view definition language – VDL): Για τον ορισμό των όψεων
- Γλώσσα χειρισμού δεδομένων (data manipulation language – DML): Για διαχείριση των δεδομένων
 - ΓΧΔ υψηλού επιπέδου: Χειρίζεται σύνολα εγγραφών και προσδιορίζει ποια δεδομένα θα χειριστούμε και όχι με ποιο τρόπο θα τα εντοπίσουμε στο φυσικό επίπεδο – **Γλώσσα επερωτήσεων (query language)**
 - ΓΧΔ χαμηλού επιπέδου: Εμφυτεύεται σε μια γλώσσα προγραμματισμού και χειρίζεται μια εγγραφή τη φορά

Διεπαφές

- Καταλόγου επιλογών (menu based interfaces)
 - Οι λειτουργίες επιλέγονται από μενού και οι χρήστες καθοδηγούνται στη συμπλήρωση των παραμέτρων. Δεν απαιτούν τη γνώση μιας γλώσσας επερωτήσεων
- Γραφικών (graphical interfaces)
 - Το σχήμα της ΒΔ εμφανίζεται διαγραμματικά και η επερώτηση δημιουργείται μέσα από το διάγραμμα
- Βασιζόμενες σε φόρμες
 - Κυρίως για εισαγωγή δεδομένων στη ΒΔ
- Φυσικής γλώσσας
 - Χρησιμοποιούν τυποποιημένες λέξεις, λέξεις του σχήματος και τυποποιημένη σύνταξη. Διερμηνεύονται σε μια υψηλού επιπέδου ερώτηση
- Παραμετρικών χρηστών
 - Εντολές σε συντομογραφία που αντιστοιχούν σε βασικές λειτουργίες στη ΒΔ (επιταχύνουν την πληκτρολόγηση)
- Διαχειριστή
 - Επιτρέπουν τη διαχείριση λογαριασμών, παραμέτρων συστήματος, διαχείριση σχήματος και αποθήκευσης των ΒΔ ενός ΣΔΒΔ





Βοηθητικά προγράμματα

- Φόρτωμα αρχείων δεδομένων
- Τήρηση αντιγράφων, για ανάκτηση σε περίπτωση καταστροφής
- Αναδιοργάνωση αρχείων, για βελτίωση της απόδοσης
- Παρακολούθηση της απόδοσης, παρέχει στατιστικά και βοηθά να πάρουμε αποφάσεις για την αναδιοργάνωση των αρχείων
- Εκτεταμένο σύστημα λεξικού δεδομένων, με επιπλέον πληροφορίες για τις σχεδιαστικές αποφάσεις, τους τρόπους προσπέλασης κλπ.



Ταξινόμηση ΣΔΒΔ

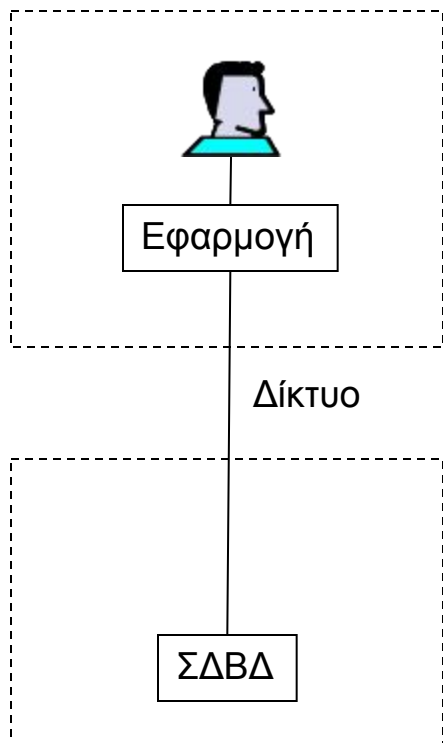
- Με βάση το μοντέλο δεδομένων:
 - **Σχεσιακά**, Δικτυωτά, Ιεραρχικά, **Αντικειμενοστραφή**
- Με βάση τον αριθμό των υποστηριζόμενων χρηστών
 - Ενός χρήστη, **πολλών χρηστών**
- Με βάση τους υπολογιστές στους οποίους κατανέμεται η ΒΔ
 - Κατανεμημένο ΣΔΒΔ
- Με βάση το κόστος
- Με βάση το σκοπό
 - Γενικού και ειδικού σκοπού

Τι χρειαζόμαστε;

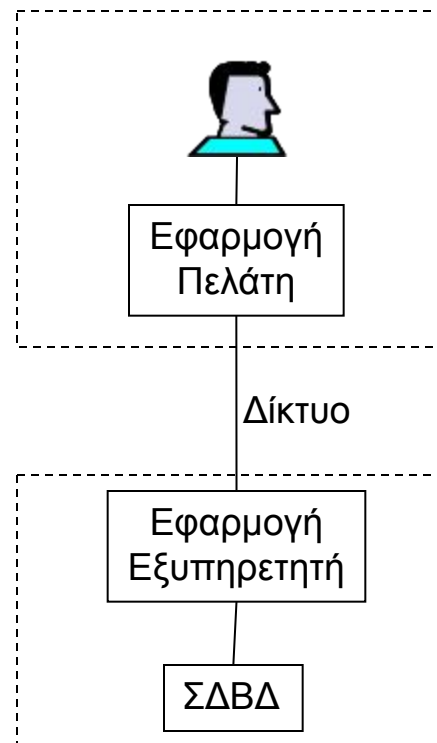
- Προμήθεια του ΣΔΒΔ
- Εκπαίδευση χρηστών
- Ορισμός του σχήματος της ΒΔ
- Εισαγωγή δεδομένων στη ΒΔ
- Υλοποίηση των προγραμμάτων της εφαρμογής
- Αλλαγές στη ΒΔ και στην εφαρμογή

Αρχιτεκτονική εφαρμογών ΒΔ

2 Επίπεδα



3 Επίπεδα





Internet και ΒΔ

- Παγκόσμιος Ιστός (web):
 - Ημιδομημένα (HTML, XML) ή πλήρως αδόμητα δεδομένα.
 - Τεράστιος αριθμός ερωτήσεων
 - Μεγάλος όγκος δεδομένων κρύβεται σε ΒΔ πίσω από κάποιες φόρμες αναζήτησης (π.χ. δικτυακές βιβλιογραφικές ΒΔ)
- Δίκτυα ομοτίμων (peer to peer): πλεονάζοντα δεδομένα διασκορπισμένα σε πολλούς υπολογιστές
- Ρεύματα δεδομένων (data streams): Τα δεδομένα έρχονται σε συνεχή ροή από αισθητήρες ή αντίστοιχες εφαρμογές



Αποθήκες δεδομένων

- OLAP (Online Analytic Processing)
 - Αναλυτική Επεξεργασία – Ανάλυση σε πολυδιάστατο χώρο
π.χ. Πωλήσεις ANA (Προϊόν, Πελάτη, Χρόνο)
προβολές πολυδιάστατου χώρου σε τρεις διαστάσεις (κύβος)
με δυνατότητα ανάλυσης (rollup-drill down) κάθε διάστασης
- Data Warehouses
 - Νέα μορφή για Decision Support Systems (Συστήματα Στήριξης Αποφάσεων)
- DataMarts
 - Μικρότερα Data Warehouses
- Data Mining
 - Εξόρυξη Δεδομένων – Πληροφοριών
π.χ. Βρες ομάδες αγοραστών με κοινά χαρακτηριστικά