

Βάσεις δεδομένων

Εισαγωγή

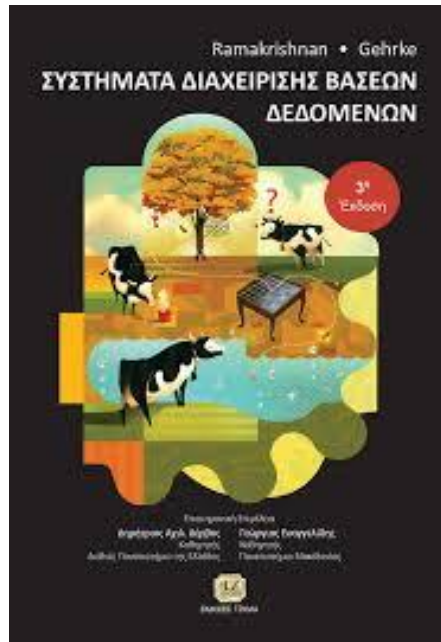
Διδάσκων: Τσούρος Δήμος

Γενικές Πληροφορίες

- Διδάσκων: Τσούρος Δήμος
 - Mail: dtsouros@uowm.gr
- Ώρες μαθήματος: Τετάρτη 09:00 – 12:00
- Ώρες γραφείου:
 - Τρίτη 11:00 – 12:00, 15:00 – 16:00
 - Τετάρτη 15:00 – 16:00

Ύλη μαθήματος:

- Προτεινόμενο Βιβλίο:



- Άλλα βιβλία:



- Διαφάνειες: (θα ανεβαίνουν στο eclass.uowm.gr)

Τρόπος διδασκαλίας

- Διαλέξεις θεωρίας
 - Παραδείγματα υλοποίησης και εφαρμογής
 - Ασκήσεις σε συγκεκριμένες μεθόδους
- Εργαστήριο γλώσσας MySQL
- Ερωτήσεις!!!!!!

Στόχος η κατανόηση των εννοιών και των μεθόδων!! Όχι η αποστήθιση..

Τρόπος αξιολόγησης

- **Γραπτή τελική εξέταση (70%)**

- Θα περιλαμβάνει

- Πολλαπλής επιλογής (με χαμηλή αρνητική βαθμολογία για αποφυγή τυχαίας επιλογής)
 - Ερωτήσεις κατανόησης
 - Ερωτήσεις ανάπτυξης
 - Ασκήσεις

- Στο τελευταίο μάθημα θα δούμε παραδείγματα ερωτήσεων και ασκήσεων

- **Εργασίες (30%)**

- (Σε συνεννόηση μπορούμε να το κάνουμε 50%-50%)

Στόχος η κατανόηση των εννοιών και των μεθόδων!! Όχι η αποστήθιση..

Στόχος

- Κατανόηση των βασικών εννοιών των βάσεων δεδομένων
- Εξοικείωση με τις κύριες τεχνικές σχεδίασης βάσεων δεδομένων.
- Εξοικείωση με την πρακτική ανάπτυξη βάσεων δεδομένων με χρήση SQL
- Ποσόστωση επιτυχίας στις εξετάσεις:
 - Μακάρι 100%!

Εισαγωγή

Τι είναι μια βάση δεδομένων;

- Βάση δεδομένων είναι μια οργανωμένη δομή μέσα στην οποία συλλέγονται δεδομένα και κατηγοριοποιούνται κατάλληλα, με σκοπό την καλύτερη αξιοποίηση τους.
- Για να υλοποιηθεί, χρησιμοποιούνται αλγόριθμοι οι οποίοι έχουν σχεδιαστεί για συστήματα διαχείρισης δεδομένων.
- Πίσω από σχεδόν **κάθε** πρόγραμμα που χρησιμοποιούμε στην καθημερινότητά μας, τα οποία αποθηκεύουν τα δεδομένα μας όπως username ή password, βρίσκεται μια βάση δεδομένων.

Γιατί όμως χρειαζόμαστε βάσεις
δεδομένων;

Ανάγκη Διαχείρισης Δεδομένων

- Τα δεδομένα σήμερα υπάρχουν παντού και σχεδόν τα πάντα καταγράφονται.
 - Χρησιμοποιούνται σε κάθε κυβερνητική, παραγωγική, εμπορική, επιστημονική και καθημερινή δραστηριότητα.



Ηλεκτρονικό εμπόριο



Συναλλαγές στην τράπεζα (κατάθεση, ανάληψη κλπ)



Ηλεκτρονική μίσθωση
καταλύματος



Ενοικίαση αυτοκινήτου



Κράτηση ξενοδοχείου



Κράτηση αεροπορικού
εισιτιοίου

Ανάγκη Διαχείρισης Δεδομένων

Πληροφορία: αποτελείται από δεδομένα τα οποία έχουν υποστεί επεξεργασία και είναι σε μία μορφή που είναι χρήσιμη στους αποδέκτες τους .

Η επεξεργασία των αρχικών δεδομένων, προσθέτει αξία στα δεδομένα.

Η επεξεργασία των δεδομένων μπορεί να περιλαμβάνει απλές πράξεις όπως:

- τον υπολογισμό μερικών και ολικών συνόλων,
 - τη σύγκριση,
 - ταξινόμηση,
 - κατηγοριοποίηση
- ή σύνθετες πράξεις όπως,
- τον υπολογισμό παραμέτρων πολύπλοκων στατιστικών μοντέλων.

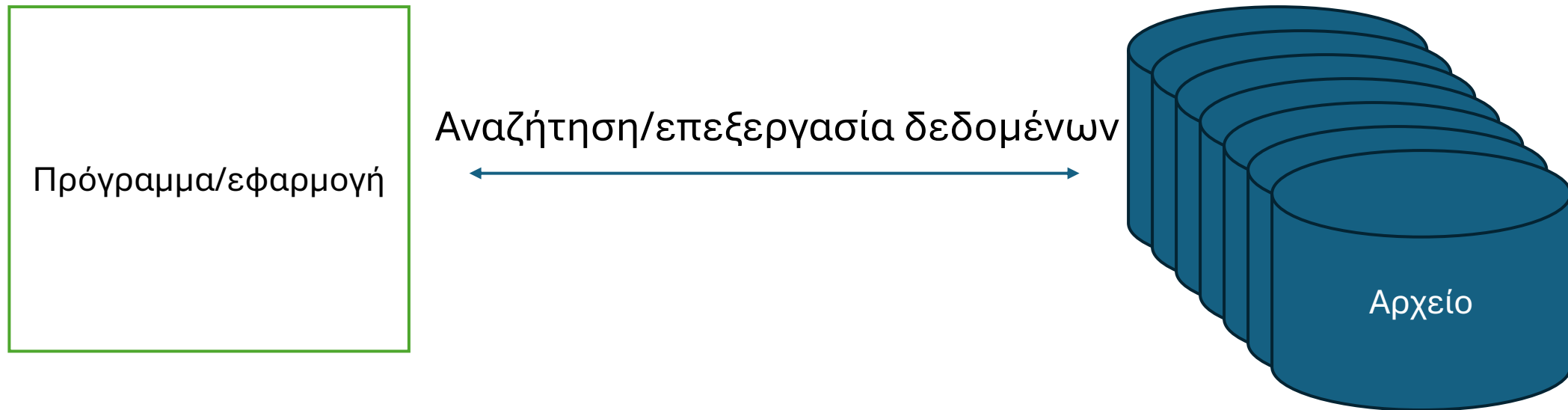
Ανάγκη Διαχείρισης Δεδομένων

- Οι μεγάλες ποσότητες δεδομένων επιβάλλουν την εύρεση αποτελεσματικών μεθόδων αποθήκευσης και χειρισμού τους.
 - Τα δεδομένα πρέπει να είναι οργανωμένα ώστε να διευκολύνεται η αναζήτηση και η ενημέρωσή τους.
- Δεδομένα χωρίς συγκεκριμένη δομή επιφέρουν προβλήματα.
- Η Διαχείριση Δεδομένων ασχολείται με τα προβλήματα αυτά.
- Οι βάσεις δεδομένων φέρνουν την απαιτούμενη δομή στα δεδομένα

Συστήματα Αρχείων: Η αρχική προσέγγιση (1/2)

- Πριν την εμφάνιση των βάσεων δεδομένων, η αποθήκευση και η διαχείριση των δεδομένων ήταν στην ευθύνη του κάθε ξεχωριστού προγράμματος/εφαρμογής.
- Η αποθήκευση των δεδομένων γινόταν σε αρχεία του λειτουργικού συστήματος.
- Η ανάγνωση, αποθήκευση και ενημέρωση των δεδομένων ανήκε στην ευθύνη του προγραμματιστή της εφαρμογής.

Συστήματα Αρχείων: Η αρχική προσέγγιση (2/2)



Μειονεκτήματα – Η ευθύνη του προγραμματιστή

- Ο προγραμματιστής της εφαρμογής είναι υπεύθυνος για τη σωστή ενημέρωση των αρχείων των δεδομένων, ανάλογα με τις εισαγωγές και τις διαγραφές που πραγματοποιούνται.
- Κατά την αναζήτηση στοιχείων πρέπει να προσδιοριστούν τα απαραίτητα αρχεία και να γίνει προσεκτική ανάγνωση, συνδυάζοντας πληροφορίες από διαφορετικά αρχεία.
- Η κατάσταση δυσκολεύει όσο αυξάνεται ο αριθμός των αρχείων δεδομένων.
- Η πολυπλοκότητα των ερωτημάτων επιτείνει το πρόβλημα.

Μειονεκτήματα – Η δομή των αρχείων

- Η δομή των αρχείων καθορίζεται από την εφαρμογή, άρα η αλλαγή της δομής είναι δύσκολη.
- Αν προστεθεί νέο χαρακτηριστικό στα δεδομένα, πρέπει να αλλάξει η δομή του αρχείου και η λογική της εφαρμογής (αναζήτηση, εισαγωγή, διαγραφή).
- Οι αλλαγές είναι υπολογιστικά χρονοβόρες
 - Αλλά και επικίνδυνες για την ακεραιότητα των δεδομένων.

Μειονεκτήματα – Επικίνδυνος «πλουραλισμός»

- Η μορφοποίηση των αρχείων εξαρτάται από τη γλώσσα προγραμματισμού.
- Δυσκολίες στη διαλειτουργικότητα μεταξύ εφαρμογών, γλωσσών, λειτουργικών συστημάτων.
- Εντονότερο πρόβλημα όταν τα δεδομένα αποθηκεύονται σε δυαδική μορφή αντί για αλφαριθμητική (text).
 - Δυαδικά αρχεία είναι αρχεία που διαβάζονται από συγκεκριμένες εφαρμογές (πχ word)
 - Χρειάζονται συγκεκριμένη επεξεργασία για το διάβασμα των περιεχομένων τους

Μειονεκτήματα – Ταυτόχρονες πράξεις

- Πολλοί χρήστες ταυτόχρονα:
 - προστασία και ακεραιότητα κατά αναζήτηση/εισαγωγή/διαγραφή είναι περίπλοκη (π.χ. κρατήσεις θέσεων).
- Σε ορισμένες λειτουργίες απαιτείται ατομικότητα εκτέλεσης
 - (π.χ. τραπεζικές συναλλαγές).
 - μια ομάδα από επιμέρους ενέργειες εκτελείται ως μία αδιαίρετη μονάδα (συναλλαγή): ή θα ολοκληρωθούν όλες και θα αποθηκευτούν οριστικά (commit), ή αν κάτι αποτύχει, θα ακυρωθούν όλες (rollback), σαν να μην έγινε καμία.

Μειονεκτήματα – Δικαιώματα χρηστών

- Πολλές φορές πρόσβαση από πολλούς χρήστες
- Κάθε χρήστης έχει διαφορετικά δικαιώματα προσπέλασης.
- Δύσκολη η εφαρμογή κανόνων προστασίας μόνο με εργαλεία λειτουργικού συστήματος.
- Πολύ δύσκολος ο ορισμός διαφορετικών δικαιωμάτων μέσα στο ίδιο αρχείο.
 - Πχ Ο ταμίας στη τράπεζα μπορεί να δει αν έχω αρκετά χρήματα αλλά δεν μπορεί αν δει το pin μου!

Μειονεκτήματα – Δυσκαμψία στα ερωτήματα

- Πως αναζητάμε δεδομένα στο αρχείο;
- Η λογική αναζήτησης υλοποιείται στον κώδικα της εφαρμογής, άρα οι τύποι ερωτημάτων είναι περιορισμένοι και προκαθορισμένοι.
 - Και δύσκολα υλοποιήσιμοι σε κάθε ξεχωριστή εφαρμογή
- Σχεδόν αδύνατο να προβλεφθούν όλα τα ερωτήματα που μπορεί να θέλει ο κάθε χρήστης.

Εναλλακτική προσέγγιση: Βάσεις Δεδομένων (1)

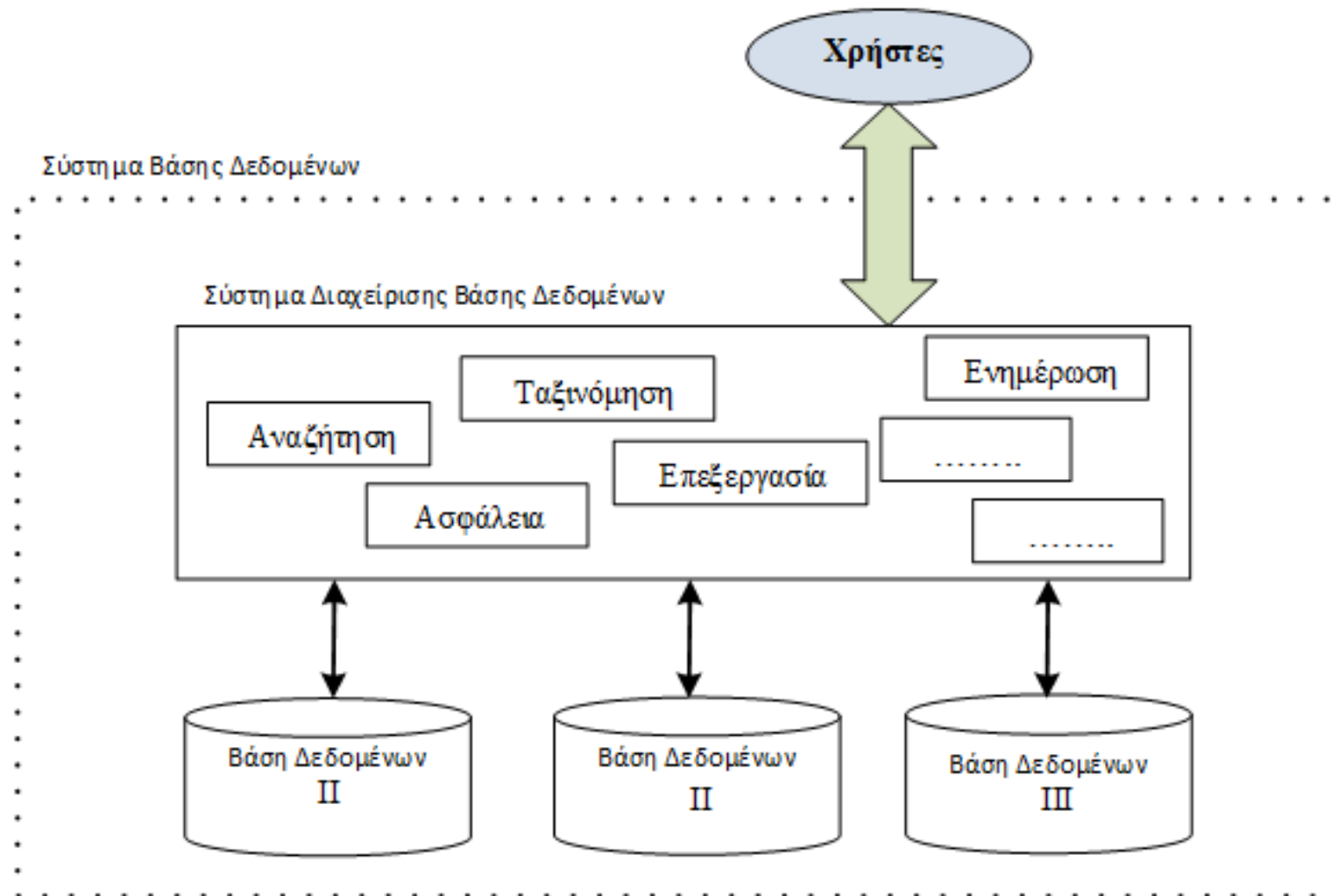
- Βάση Δεδομένων (Database): συλλογή σχετικών, δομημένων και κατάλληλα καταχωρισμένων στοιχείων.
 - Ορισμός: Βάση δεδομένων είναι μια οργανωμένη δομή μέσα στην οποία συλλέγονται δεδομένα και κατηγοριοποιούνται κατάλληλα, με σκοπό την καλύτερη αξιοποίηση τους.
- Σύστημα Διαχείρισης Βάσεων Δεδομένων (DBMS): Λογισμικό για αναζήτηση, εισαγωγή, διαγραφή, συγχρονισμό προσπελάσεων, προστασία κ.ά.
 - Έτοιμα λογισμικά που διευκολύνουν την χρήση βάσεων δεδομένων όπου χρειαστούμε
- Σύστημα Βάσης Δεδομένων (Database System): ΒΔ μαζί με το αντίστοιχο λογισμικό ΣΔΒΔ.

Εναλλακτική προσέγγιση: Βάσεις Δεδομένων (2)

Μια **Βάση Δεδομένων (ΒΔ)** έχει τις παρακάτω **ιδιότητες**:

- Μια ΒΔ αναπαριστά κάποια άποψη του πραγματικού κόσμου, η οποία μερικές φορές λέγεται *μικρόκοσμος*.
 - Η ΒΔ εμπεριέχει αφαίρεση/απόκρυψη λεπτομερειών της πραγματικότητας που δεν ενδιαφέρουν την εφαρμογή
 - Μόνο οι αλλαγές στον μικρόκοσμο αντανακλώνται στη βάση δεδομένων
- Μια βάση δεδομένων είναι μια λογικά συνεκτική συλλογή δεδομένων που έχει κάποια εγγενή σημασία.
 - Μια τυχαία διευθέτηση δεδομένων δεν είναι σωστό να αναφέρεται ως βάση δεδομένων
- Μια βάση δεδομένων σχεδιάζεται, χτίζεται και γεμίζει με δεδομένα για κάποιο συγκεκριμένο σκοπό.
 - Προορίζεται για μια συγκεκριμένη ομάδα χρηστών και για κάποιες προκαθορισμένες εφαρμογές για τις οποίες οι χρήστες αυτοί ενδιαφέρονται

Εναλλακτική προσέγγιση: Βάσεις Δεδομένων (3)



Παράδειγμα μεγάλης εμπορικής βάσης δεδομένων

Μια μεγάλη εμπορική βάση δεδομένων είναι η Amazon.com

- Περιέχει δεδομένα για εκατομμύρια ενεργούς χρήστες
- Περιέχει εκατομμύρια προϊόντα διαφορετικού τύπου
 - βιβλία, CD, Video, DVD, παιχνίδια, ηλεκτρονικά, συσκευές κ.α.
- Η ΒΔ ενημερώνεται συνεχώς για να αντανακλά τον μικρόκοσμο που την ενδιαφέρει
 - Η ΒΔ της ξέρει το ιστορικό αγορών μου
 - Αλλά (μάλλον) δεν ξέρει το χρώμα ματιών μου (δεν ενδιαφέρει τον μικρόκοσμό της)
- ΠΡΟΣΟΧΗ!
 - **Η ΒΔ δεν είναι απλά μια συλλογή από πίνακες.**
 - Αλλά εμπεριέχει και σχέσεις μεταξύ των πινάκων
 - Π.χ. Ο πίνακας για το ιστορικό παραγγελιών είναι συνδεδεμένος με τον πίνακα χρηστών, προϊόντων κλπ (ποιος χρήστης παρήγγειλε ποιο προϊόν, για παράδοση σε ποια πόλη, ...)

Πλεονεκτήματα ΒΔ

- Περιγραφή δεδομένων (μεταδεδομένα, κατάλογος συστήματος) ξεχωριστά από τα ίδια τα δεδομένα.
- Ανεξαρτησία *Δεδομένων* και *Λειτουργιών* (δεν αλλάζουν τα προγράμματα εφαρμογής όταν αλλάζουν τα δεδομένα).
- Αποδοτικότερη Διαχείριση Δεδομένων (ο προγραμματιστής επικεντρώνεται στην εφαρμογή, όχι σε αλγορίθμους επεξεργασίας).
- Προστασία δεδομένων και δικαιώματα χρηστών (για κάθε σύνολο δεδομένων ξεχωριστά).
- Μηχανισμοί ταυτόχρονης προσπάθειας (μηχανισμοί κλειδαριών από το σύστημα).
- Εύκολη επεκτασιμότητα (σε μεθόδους, χρήστες, δεδομένα, εφαρμογές, υπολογιστές, δίσκους κλπ).

Παραδοσιακές εφαρμογές

- Μισθοδοσία
- Τηλεφωνικός κατάλογος
- Κατάλογος μελών/χρηστών
- Κρατήσεις θέσεων
- Τραπεζικές συναλλαγές
- Διαχείριση εταιρικών δεδομένων

Σύγχρονες εφαρμογές

- Κοινωνικά δίκτυα (Social media)
- Πολυμεσικές εφαρμογές (Multimedia DBs)
- Διαχείριση γεωγραφικών πληροφοριών (GIS)
- Εξόρυξη δεδομένων (Data mining)
- Ανάκτηση πληροφορίας (Information retrieval)
- Μεγάλα δεδομένα (Big data)
- Τεχνητή Νοημοσύνη (Artificial intelligence)
- ...

Απλό παράδειγμα

Το πρόβλημα

Ζητείται ένα ηλεκτρονικό αρχείο καταγραφής των επαφών/φίλων μας, με στοιχεία όπως όνομα, τηλέφωνο, γενέθλια κ.λπ.



Κάθε φίλος μας έχει:

- 1 Όνομα
- 2 Διεύθυνση
- 3 Τηλέφωνο
- 4 Γενέθλια

Κάθε αναπαράσταση σε μια βάση δεδομένων κρύβει πολλά:

- 1 Το όνομα της μητέρας του
- 2 Το νούμερο παπούτσι του
- 3 Το ύψος του
- 4 Ποια πίτσα προτιμάει
- 5 Και άλλα πολλά . . .

Ορισμός

Φίλος είναι «κάτι» που έχει όνομα, διεύθυνση, τηλέφωνο και γενέθλια.

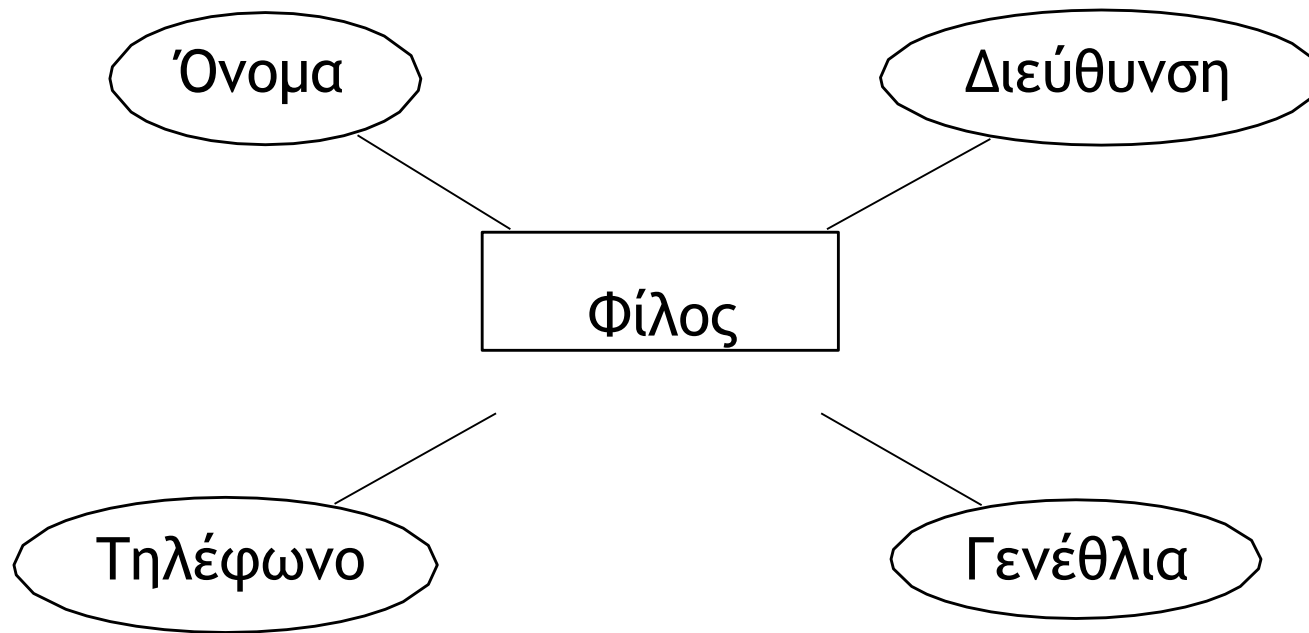
Παράδειγμα

Νίκος Αλεξίου, Λάρισα, 2410*****, 5/Φεβ.

Αφαίρεση

- 1 Οι βάσεις δεδομένων περιέχουν δεδομένα που προέρχονται από **αφαίρεση** της πραγματικότητας.
- 2 Η διαδικασία με την οποία γίνεται η αφαίρεση λέγεται **μοντελοποίηση** δεδομένων.

Τα πολλά λόγια είναι φτώχεια



- 1 Μια οντότητα, ο Φίλος, με 4 ιδιότητες.
- 2 Δεν είναι δυνατό και οι 4 ιδιότητες να έχουν τις ίδιες τιμές.
- 3 Περιγράφουμε τις έννοιες της βάσης δεδομένων με γεωμετρικά σχήματα:
 - Ορθογώνιο: οντότητα της βάσης δεδομένων
 - Έλλειψη: ιδιότητα της οντότητας

Τύποι δεδομένων, προσοχή!

- ❶ Ο τύπος δεδομένων «κείμενο» δεν είναι πάντοτε επαρκής.
- ❷ Σε περιπτώσεις όπου απαιτούνται αλγεβρικές πράξεις, προτιμούμε αριθμητικούς τύπους δεδομένων.
- ❸ Προσοχή! Τα παρακάτω δεν είναι αριθμοί:
 - Αριθμός διαβατηρίου
 - Αριθμός φορολογικού μητρώου
 - Αριθμός τηλεφώνου
 - Αριθμός κυκλοφορίας αυτοκινήτου
 - Αριθμός τραπεζικού λογαριασμού.
- ❹ Υπάρχουν ειδικοί τύποι δεδομένων για το χειρισμό χρονολογικών δεδομένων, αποθηκεύονται ως αριθμοί, παρουσιάζονται ως κείμενο.
- ❺ Πολλοί άλλοι τύποι δεδομένων για ειδικές περιπτώσεις.

Επόμενο μάθημα: Εισαγωγή στην σχεδίαση βάσεων δεδομένων