



Εθνικό Μετσόβιο Πολυτεχνείο  
Σχολή Αγρονόμων Τοπογράφων Μηχανικών

## Βάσεις Δεδομένων

Βασίλειος Βεσκούκης  
v.vescoukis@cs.ntua.gr

*Περιεχόμενο του μαθήματος, εισαγωγή, διαδικασία*

### Σχετικά με το μάθημα

#### Διδάσκοντες 2005-2006

- **P.Κορακίτης**, **B.Βεσκούκης**, Θ.Καραλόπουλος, Γ.Πανόπουλος

#### Ωρες διαλέξεων

- Τετάρτη 08.45-10.30, Μ. Αμφιθέατρο

#### Εργαστήριο (υποχρεωτικό)

- Τρίτη 17.00-19.00 (τμήμα 2), διατμηματικό εργ. PCLAB, κτίριο ΗΥ
- Πέμπτη 17.00-19.00 (τμήμα 1), διατμηματικό εργ. PCLAB, κτίριο ΗΥ
- Τμήματα ίδια με των μαθημάτων του 1<sup>ου</sup> & 2<sup>ου</sup> Εξαμήνου

#### Υλικό αναφοράς (βιβλίο)

- *Συστήματα Βάσεων Δεδομένων*, Silberschatz, Korth, Sudarshan.  
Υλη: Κεφάλαια 1, 2, 3, 4, 5.1, 5.3, 7, Παράρτημα Α, Β.

## Σχετικά με το μάθημα

### Διεξαγωγή μαθήματος

- Διαλέξεις: υποχρεωτικές μόνο επί της ουσίας
- Εργαστήριο: η ΒΔ "MS Access", ασκήσεις σε SQL
- Εκπόνηση και παράδοση εργασίας

### Διάρκεια

- 13 εβδομάδες...

### Βαθμολογία

- Γραπτή εξέταση: 75%
- Εργαστηριακές ασκήσεις: 25%

Ο βαθμός της εργασίας "κανονικοποιείται" ώστε να μην απέχει περισσότερο από 3 μονάδες από τον βαθμό του διαγωνίσματος.

**Όποιος δεν παραδώσει εργασία γράφει με άριστα το 7,5**

### Επικοινωνία

- E-mail ([v.vescoukis@cs.ntua.gr](mailto:v.vescoukis@cs.ntua.gr))
- Σελίδα web: [www.survey.ntua.gr/main/courses/general/dbintro](http://www.survey.ntua.gr/main/courses/general/dbintro)

Δρ. Βασίλειος Βεσκούκης

## Η θέση του μαθήματος στο πρόγραμμα σπουδών

### Τρία μαθήματα κορμού πληροφορικής σε όλες τις Σχολές του ΕΜΠ

- Εισαγωγή στην πληροφορική (1ο εξ)
- Προγραμματιστικές τεχνικές (2ο εξ)
- **Βάσεις Δεδομένων (3ο εξ)**

### Ουσιώδης απαίτηση για με το μάθημα

«Αρχές Γεωπληροφορικής και Συστήματα Γεωγραφικών Πληροφοριών»



Δρ. Βασίλειος Βεσκούκης

## Αρχικοί προβληματισμοί

Τι είναι βάση δεδομένων;

Για ποιο λόγο οι ΒΔ αποτελούν θέμα προς ενασχόληση;

Ποιες είναι οι πλευρές του θέματος αυτού;

Ποιες από αυτές «ενδιαφέρουν» τους ΑΤΜ;

Τι περιμένουμε από το μάθημα;

Δρ. Βασίλειος Βεσκούκης

## Αρχικοί προβληματισμοί

**Βάση δεδομένων:** Ορισμός της ερώτησης

- Τι είναι δεδομένα;
- Τι σχέση έχουν δεδομένα και πληροφορία;

**Δεδομένα**

- Δεδομένα (DATA) είναι κάθε παράσταση σε μορφή κειμένου, αριθμού, συμβόλων και γενικά κάθε περιγραφή περιστατικού, στην οποία είναι δυνατό να αποδοθεί μια σημασία (έννοια).

**Πληροφορία**

- Πληροφορία είναι η σημασία που δίνει ο άνθρωπος στα δεδομένα

**Θεωρητικός ορισμός της πληροφορίας**

(Claude Shannon, "The Mathematical Theory of Communication", 1948)

**Information is a decrease in uncertainty**

- **Πληροφορία** είναι ό,τι μειώνει την **εντροπία**
  - $h(p) = -\log_2(p)$ , N ανεξάρτητες καταστάσεις  $a_1..a_n$ , με πιθανότητες εμφάνισης  $p_1..p_n$
- **Εντροπία** είναι το μέτρο της αταξίας / αβεβαιότητας
  - $H = p_1 * h(p_1) + p_2 * h(p_2) + p_3 * h(p_3) + \dots + p_{N-1} * h(p_{N-1}) + p_N * h(p_N)$

Δρ. Βασίλειος Βεσκούκης

## Αρχικοί προβληματισμοί

Για ποιο λόγο οι ΒΔ αποτελούν θέμα προς ενασχόληση;

Γιατί τα δεδομένα «σήμερα»

- Είναι πολλά
- Είναι σύνθετα
- Είναι χρήσιμα ή άχρηστα
- Καταγράφονται σχετικά με τα πάντα (αυτό δεν είναι υποχρεωτικά καλό!)
- Περιγράφουν ή χρησιμοποιούνται στην τεκμηρίωση ή/και την παραγωγή νέας γνώσης

Η διαχείριση δεδομένων με τη χρήση Η/Υ οριοθετεί τη γνωστική περιοχή των «Βάσεων Δεδομένων»

Δρ. Βασίλειος Βεσκούκης

## Αρχικοί προβληματισμοί

Πλευρές του θέματος

Η παράσταση των δεδομένων

- Διάκριση σε κατηγορίες (μορφές)
- Απόδοση χαρακτηριστικών ιδιοτήτων σε κάθε κατηγορία

Η οργάνωση των δεδομένων

- Ταξινόμηση, ομαδοποίηση και δόμηση των δεδομένων
- Ο εντοπισμός και η χρήση συσχετίσεων μεταξύ των δεδομένων
- Η θεωρητική θεμελίωση των παραπάνω

Η αποθήκευση των δεδομένων

- Ο τρόπος καταγραφής τους στα μέσα μόνιμης αποθήκευσης των Η/Υ
- Οι διαδικασίες πρόσβασης και ανάκτησης
- Η βελτιστοποίηση των επιδόσεων

Η προτυποποίηση όλων των παραπάνω

Η ανάπτυξη και χρήση εργαλείων για όλα τα παραπάνω

Δρ. Βασίλειος Βεσκούκης

## Αρχικοί προβληματισμοί

«Τοπικά» ενδιαφέρουσες πλευρές του θέματος

Η παράσταση των δεδομένων

- Διάκριση σε κατηγορίες (μορφές)
- Απόδοση χαρακτηριστικών ιδιοτήτων σε κάθε κατηγορία

Η οργάνωση των δεδομένων

- Ταξινόμηση, ομαδοποίηση και δόμηση των δεδομένων
- Ο εντοπισμός και η χρήση συσχετίσεων μεταξύ των δεδομένων
- Η θεωρητική θεμελίωση των παραπάνω

Η αποθήκευση των δεδομένων

- Ο τρόπος καταγραφής τους στα μέσα μόνιμης αποθήκευσης των Η/Υ
- Οι διαδικασίες πρόσβασης και ανάκτησης
- Η βελτιστοποίηση των επιδόσεων

Η προτυποποίηση όλων των παραπάνω

Η ανάπτυξη και χρήση εργαλείων για όλα τα παραπάνω σε εφαρμογές GIS κ.ά.

Δρ. Βασίλειος Βεσκούκης

## Αρχικοί προβληματισμοί

Τι περιμένουμε από το μάθημα;

Τι ΔΕΝ περιμένουμε

- Να γίνουμε ειδικοί σε ΒΔ
- Να βρούμε άμεσα δουλειά προγραμματιστή σε SQL/Oracle κλπ
- Να προάγουμε την επιστήμη στον τομέα των ΒΔ

Τι θα μπορούσαμε να περιμένουμε

- Να καταλάβουμε τις έννοιες περί ΒΔ
- Να εξοικειωθούμε σε ικανοποιητικό βαθμό με το Σχεσιακό Μοντέλο Δεδομένων
- Να μάθουμε βασικά στοιχεία της γλώσσας SQL
- Να μπορούμε να χρησιμοποιήσουμε την MS Access

Δρ. Βασίλειος Βεσκούκης

## Αντικείμενα του μαθήματος

### Θεωρία

- Εισαγωγή στις ΒΔ. Μοντελοποίηση δεδομένων. Το μοντέλο E-R. Αναφορά στα κλασσικά μοντέλα Βάσεων Δεδομένων (Ιεραρχικό, Δικτυωτό). Το Σχεσιακό Μοντέλο.
- Συστήματα Διοίκησης Βάσεων Δεδομένων (Database Management Systems) και η αρχιτεκτονική τους.
- Γλώσσες για Βάσεις Δεδομένων - Η γλώσσα SQL.
- Φυσικός Σχεδιασμός Βάσεων Δεδομένων.
- Εισαγωγή στη γλώσσα συμβολισμών UML. Λογικός Σχεδιασμός και Κανονικοποίηση (normalization).
- Θέματα Διαχείρισης και Λειτουργίας (ακεραιότητα, βελτιστοποίηση, αναδιοργάνωση, ασφάλεια, λειτουργικότητα).
- Διατάξεις ΒΔ (client-server, κατανεμημένες ΒΔ) και προσπέλαση δεδομένων από ετερογενή συστήματα (xDBC, middleware)

### Εργαστήριο

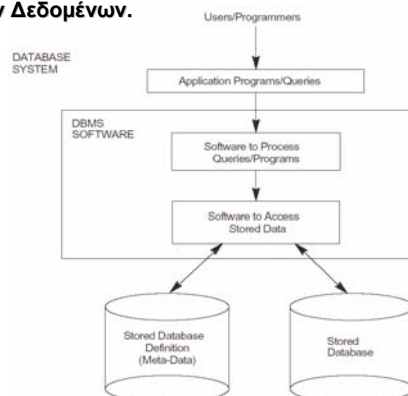
- Το λογισμικό MS Access
- Εκπόνηση εξαμηνιαίας εργασίας

Δρ. Βασίλειος Βεσκούκης

## Βάση δεδομένων

**Βάση δεδομένων** είναι μια συλλογή από δεδομένα που συσχετίζονται, η οποία αποθηκεύεται σε δευτερεύουσα μνήμη

**Σύστημα Διαχείρισης Βάσεων Δεδομένων (DBMS: Data Base Management System)** είναι ένα σύνολο από προγράμματα για τη δημιουργία, επεξεργασία, συντήρηση και αξιοποίηση Βάσεων Δεδομένων.



Δρ. Βασίλειος Βεσκούκης

## Παράδειγμα

- Μια Βάση Δεδομένων για την αποθήκευση στοιχείων εγγραφών και βαθμολογίας (παράδειγμα βιβλίου)

STUDENT	Name	StudentNumber	Class	Major
	Smith	17	1	CS
	Brown	8	2	CS

COURSE	CourseName	CourseNumber	CreditHours	Department
	Intro to Computer Science	CS1310	4	CS
	Data Structures	CS3320	4	CS
	Discrete Mathematics	MATH2410	3	MATH
	Database	CS3380	3	CS

SECTION	SectionIdentifier	CourseNumber	Semester	Year	Instructor
	85	MATH2410	Fall	98	King
	92	CS1310	Fall	98	Anderson
	102	CS3320	Spring	99	Knuth
	112	MATH2410	Fall	99	Chang
	119	CS1310	Fall	99	Anderson
	135	CS3380	Fall	99	Stone

GRADE_REPORT	StudentNumber	SectionIdentifier	Grade
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

PREREQUISITE	CourseNumber	PrerequisiteNumber
	CS3380	CS3320
	CS3380	MATH2410
	CS3320	CS1310

Δρ. Βασίλειος Βεσκούκης

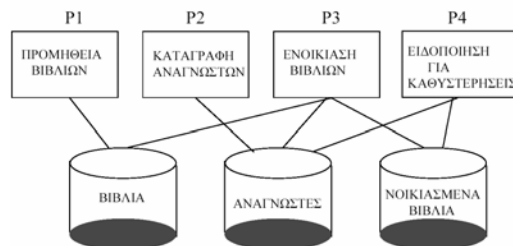
## Λίγο ιστορία

### Πρώτη Γενιά (1950)

- Σειριακή επεξεργασία
- Ταινίες και κάρτες
- Σειριακή επεξεργασία

### Δεύτερη γενιά (1960)

- Δίσκοι, συστήματα αρχείων
- Γλώσσες προγραμματισμού με ικανότητα χειρισμού δεδομένων (COBOL)
- Αποθήκευση δεδομένων σε Αρχεία
- Μερικές πρώτες εφαρμογές

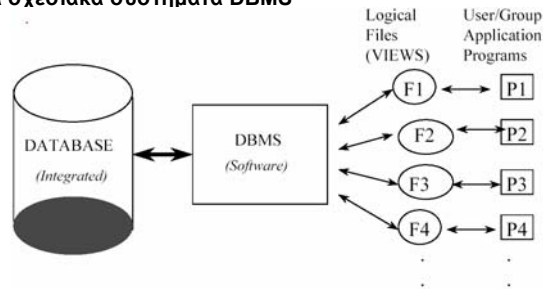


Δρ. Βασίλειος Βεσκούκης

## Λίγο ιστορία

### Τρίτη γενιά (1970)

- Ανάγκη και τάση δημιουργίας προτύπων
- Αποθήκευση δεδομένων μαζί με τον ορισμό τους
- Νέες δυνατότητες προσπέλασης δεδομένων σε γλώσσες όπως η COBOL
- Δικτυακό και Ιεραρχικό μοντέλο δεδομένων
- Ορισμός του Σχεσιακού μοντέλου δεδομένων
- Αναγνώριση της ανάγκης για μοντελοποίηση δεδομένων
- Τα πρώτα σχεσιακά συστήματα DBMS



Δρ. Βασίλειος Βεσκούκης

## Λίγο ιστορία

### Τέταρτη γενιά (1980, 1990)

- ΣΧΕΣΙΑΚΑ ΣΥΣΤΗΜΑΤΑ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ (Relational Database Systems)
- Ισχυρές γλώσσες επεξεργασίας δεδομένων
- Αρχιτεκτονικές πολλών επιπέδων
- Ενσωμάτωση των DBMS σε μεγάλα συστήματα και ανάδειξη του ρόλου τους (πχ διαχείριση δισκοληψιών)
- Γέννηση της SQL (Structured Query Language)
- Θεωρητική θεμελίωση των ΒΔ

### Πέμπτη γενιά (1990, 2000)

- Νέες μορφές δεδομένων (multimedia) σε ΒΔ
- Σύνθετες οντότητες (software objects)
- Κατανεμημένες αρχιτεκτονικές (πελάτες, εξυπηρετητές)
- Ισχυρά DBMS σε προσωπικούς υπολογιστές
- Σύνδεση ΒΔ και Internet για πολλές εφαρμογές
- Εξόρυξη γνώσης από ΒΔ (data mining)
- Ανάδειξη σε αυτοτελές αντικείμενο (data warehousing)

Δρ. Βασίλειος Βεσκούκης





Εθνικό Μετσόβιο Πολυτεχνείο  
Σχολή Αγρονόμων Τοπογράφων Μηχανικών

## Βάσεις Δεδομένων

**Βασίλειος Βεσκούκης**

*v.vescoukis@cs.ntua.gr*

**ΜΟΝΤΕΛΑ ΔΕΔΟΜΕΝΩΝ**

### Βάσεις Δεδομένων

#### Βάση δεδομένων (ΒΔ)

- μια συλλογή από συσχετιζόμενα δεδομένα

#### Δεδομένα

- γνωστά γεγονότα που μπορούν να καταγραφούν και έχουν κάποια υπονοούμενη σημασία

#### Παράδειγμα βάσης δεδομένων:

- τα ονόματα, οι αριθμοί τηλεφώνων και οι διευθύνσεις μιας ατζέντας

#### Μια βάση δεδομένων έχει τις ακόλουθες ιδιότητες

- αναπαριστά κάποια άποψη του πραγματικού κόσμου (μικρόκοσμος)
- αποτελεί μια λογικά συνεκτική συλλογή δεδομένων που έχει κάποια εγγενή σημασία (όχι τυχαία διευθέτηση δεδομένων)
- σχεδιάζεται, χτίζεται και φιλοξενεί δεδομένα για συγκεκριμένο σκοπό και προορίζεται για συγκεκριμένους χρήστες και εφαρμογές

## Βάσεις Δεδομένων

Μια βάση δεδομένων δημιουργείται και συντηρείται

- χειρόγραφα (π.χ., τα βιβλία των υποθηκοφυλακείων)
- με χρήση Η/Υ (π.χ., η βάση δεδομένων του κτηματολογίου)

Μια βάση δεδομένων σε Η/Υ δημιουργείται και συντηρείται

- από μια ομάδα προγραμμάτων εφαρμογών, που έχουν συνταχθεί ειδικά για το σκοπό αυτό
- από ένα σύστημα διαχείρισης βάσεων δεδομένων (ΣΔΒΔ)

Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ)

- συλλογή από προγράμματα που επιτρέπουν στους χρήστες να δημιουργήσουν και να συντηρήσουν μια βάση δεδομένων

Σύστημα βάσης δεδομένων (ΣΒΔ)

- καλείται μια βάση δεδομένων μαζί με το λογισμικό χειρισμού της

Δρ. Βασίλειος Βεσκούκης

## Βάσεις Δεδομένων

Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ)

- γενικής χρήσης σύστημα λογισμικού που διευκολύνει τις διαδικασίες ορισμού, κατασκευής και χειρισμού βάσεων δεδομένων για διάφορες εφαρμογές

Ορισμός μιας βάσης δεδομένων

- προδιαγραφή των τύπων, δομών, και περιορισμών των δεδομένων που θα αποθηκευτούν στη βάση

Κατασκευή μιας βάσης δεδομένων

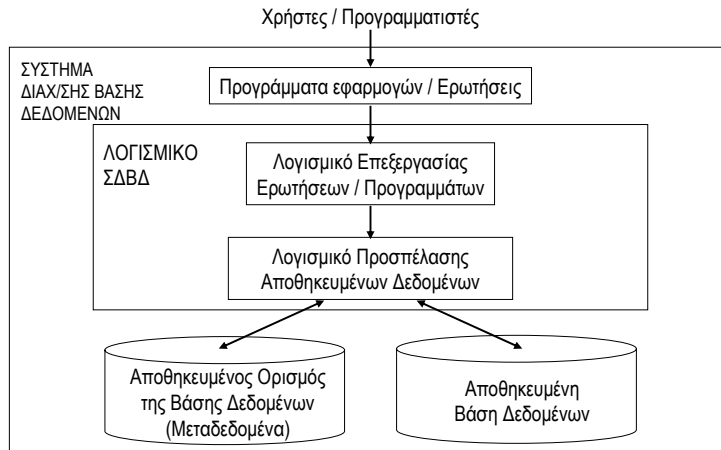
- αποθήκευση των ίδιων των δεδομένων σε ένα μέσο αποθήκευσης που ελέγχεται από το ΣΔΒΔ

Χειρισμός μιας βάσης δεδομένων

- υποβολή ερωτήσεων προς τη βάση για ανάκτηση συγκεκριμένων δεδομένων, ενημέρωση της βάσης με τις αλλαγές του μικρόκοσμου, και παραγωγή αναφορών από τα δεδομένα

Δρ. Βασίλειος Βεσκούκης

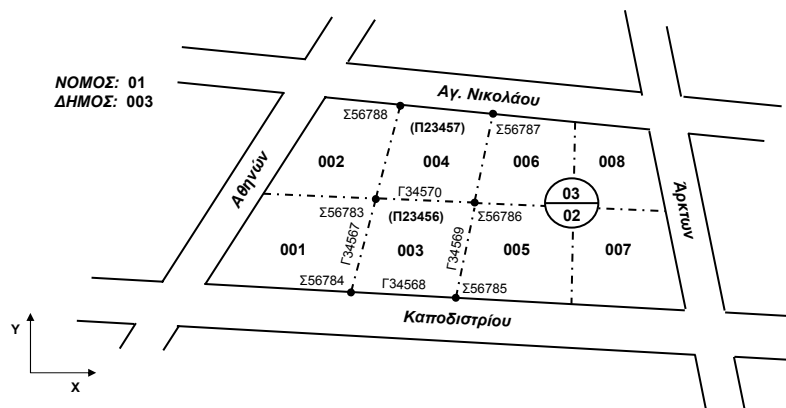
## Συστήματα Διαχείρισης Βάσεων Δεδομένων



Δρ. Βασίλειος Βεσκούκης

## Βάσεις δεδομένων - παράδειγμα

Βάση δεδομένων για εφαρμογές κτηματολογίου



Δρ. Βασίλειος Βεσκούκης

## Βάσεις δεδομένων - παράδειγμα

### Βάση δεδομένων για εφαρμογές κτηματολογίου

#### ΓΕΩΤΕΜΑΧΙΑ

ΚΑΕΚ	ΕΜΒΛΑΘΟΝ	ΧΡΗΣΗ	ΔΙΕΥΘΥΝΣΗ	ΠΟΛΥΓΩΝΟ
01.003.03.02.003	1.562	ΟΙΚΙΣΤΙΚΗ	ΚΑΠΟΔΙΣΤΡΕΙΟΥ 32	Π23456
01.003.03.02.004	2.578	ΚΑΛΛΙΕΡΓΕΙΑ	ΑΓ. ΝΙΚΟΛΑΟΥ 141	Π23457
...	...	...	...	...

#### ΔΙΚΑΙΟΥΧΟΙ

ΚΩΔΙΚΟΣ	ΕΠΩΝΥΜΟ	ΟΝΟΜΑ	ΔΙΕΥΘΥΝΣΗ
46419735	ΔΗΜΗΤΡΙΟΥ	ΝΙΚΟΛΑΟΣ	ΑΘΗΝΩΝ 45, 11562 ΑΘΗΝΑ
56712945	ΠΑΡΑΣΧΟΥ	ΜΑΡΙΑ	ΝΙΚΗΣ 22, 74100 ΡΕΘΥΜΝΟ
...	...	...	...

#### ΔΙΚΑΙΩΜΑΤΑ

ΚΑΕΚ	ΔΙΚΑΙΟΥΧΟΣ	ΠΟΣΟΣΤΟ	ΗΜΕΡΟΜΗΝΙΑ
01.003.03.02.003	46419735	60%	28-7-1954
01.003.03.02.003	56712945	40%	28-7-1954
01.003.03.02.004	56712945	100%	12-3-1987
...	...	...	...

#### ΠΟΛΥΓΩΝΑ

ΚΩΔΙΚΟΣ	ΓΡΑΜΜΗ
Π23456	Γ34567
Π23456	Γ34568
Π23456	Γ34569
Π23456	Γ34570
...	...

#### ΓΡΑΜΜΕΣ

ΚΩΔΙΚΟΣ	ΑΠΟ ΣΗΜΕΙΟ	ΠΡΟΣ ΣΗΜΕΙΟ
Γ34567	Σ56783	Σ56784
Γ34568	Σ56784	Σ56785
...	...	...

#### ΣΗΜΕΙΑ

ΚΩΔΙΚΟΣ	X	Y
Σ56784	45678.34	8938.89
Σ56785	45685.56	8979.67
...	...	...

Δρ. Βασίλειος Βεσκούκης

## Βάσεις δεδομένων - παράδειγμα

### Αρχεία της βάσης δεδομένων

- η βάση δεδομένων είναι οργανωμένη σε έξι αρχεία
- το κάθε αρχείο αποθηκεύει εγγραφές δεδομένων του ίδιου τύπου
  - π.χ., στο αρχείο ΔΙΚΑΙΟΥΧΟΙ αποθηκεύονται δεδομένα για κάθε δικαιούχο

### Ορισμός της βάσης δεδομένων

- προσδιορισμός της δομής των εγγραφών κάθε αρχείου
  - π.χ., κάθε εγγραφή του αρχείου ΔΙΚΑΙΩΜΑΤΑ περιλαμβάνει δεδομένα για την αναπαράσταση του κωδικού της ιδιοκτησίας (ΚΑΕΚ), του κωδικού του δικαιούχου (ΔΙΚΑΙΟΥΧΟΣ), του ποσοστού της ιδιοκτησίας (ΠΟΣΟΣΤΟ) και της ημερομηνίας κτήσης (ΗΜΕΡΟΜΗΝΙΑ)
- ορισμός του τύπου κάθε στοιχειώδους δεδομένου στην εγγραφή
  - π.χ., το στοιχειώδες δεδομένο ΕΠΩΝΥΜΟ είναι αλφαριθμητικό 10 χαρακτήρων, ενώ το ΠΟΣΟΣΤΟ αριθμητικό (0-100) (περιορισμός)

Δρ. Βασίλειος Βεσκούκης

## Βάσεις δεδομένων - παράδειγμα

### Κατασκευή της βάσης δεδομένων

- αποθήκευση των δεδομένων στα αντίστοιχα αρχεία
  - π.χ., τα δεδομένα του δικαιούχου αποθηκεύονται στο αρχείο ΔΙΚΑΙΟΥΧΟΙ, ενώ τα δεδομένα του αγροτεμαχίου αποθηκεύονται στο αρχείο ΙΔΙΟΚΤΗΣΙΕΣ
- οι εγγραφές στα διάφορα αρχεία μπορεί να συσχετίζονται μεταξύ τους
  - π.χ., η εγγραφή για τον Δημητρίου στο αρχείο ΔΙΚΑΙΟΥΧΟΙ σχετίζεται με δύο εγγραφές στο αρχείο ΙΔΙΟΚΤΗΣΙΕΣ που προσδιορίζουν τις ιδιοκτησίες του Δημητρίου

### Χειρισμός της βάσης δεδομένων

- περιλαμβάνει ερωτήσεις και ενημερώσεις
  - π.χ., «ποιες είναι οι ιδιοκτησίες του Δημητρίου;» ή «τροποποίησε τη διεύθυνση του Δημητρίου σε Αθηνών 69, 16675 Γλυφάδα»
- οι ερωτήσεις και ενημερώσεις εκφράζονται στη γλώσσα του συστήματος