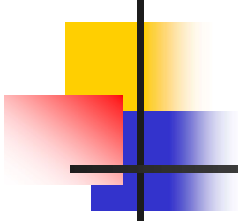


Βάσεις Δεδομένων Ι

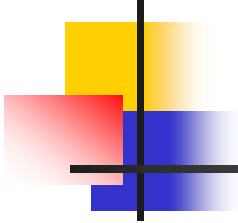
Πληροφοριακά Συστήματα Διοίκησης

Διδάσκων: Δημόκας Νίκος, Ph.D.



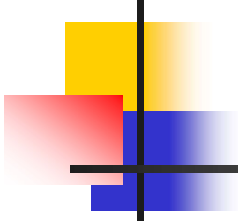
Συστήματα Διαχείρισης Βάσεων Δεδομένων (1/7)

- Ένα Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ) (Database Management System -DBMS) αποτελείται από μία συλλογή σχετιζόμενων δεδομένων και ένα σύνολο προγραμμάτων για την προσπέλαση σε αυτά τα δεδομένα
- Η συλλογή των δεδομένων ονομάζεται βάση δεδομένων (ΒΔ) (database).
- Ο κύριος σκοπός ενός ΣΔΒΔ είναι να παρέχει ένα περιβάλλον για την εύκολη και αποδοτική αποθήκευση και ανάκληση μεγάλου όγκου πληροφορίας.



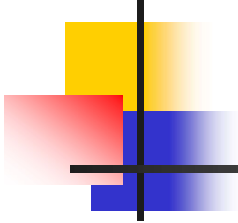
Συστήματα Διαχείρισης Βάσεων Δεδομένων (2/7)

- Η διαχείριση των δεδομένων περιλαμβάνει:
 - Τον ορισμό δομών για την αποθήκευση πληροφορίας
 - Μηχανισμούς για τον χειρισμό της πληροφορίας και την ενημέρωσή της
 - Μηχανισμούς για την ασφάλεια πληροφορίας που είναι προσβάσιμη από πολλούς χρήστες.



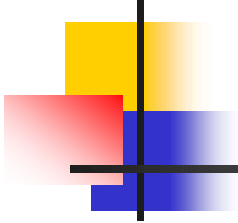
Συστήματα Διαχείρισης Βάσεων Δεδομένων (3/7)

- Παράδειγμα: ένα τραπεζικό σύστημα αποθηκεύει πληροφορία σχετικά με τους πελάτες της τράπεζας και τους λογαριασμούς τους.
 - Η πληροφορία αποθηκεύεται σε αρχεία του συστήματος τα οποία διαχειρίζονται προγράμματα για:
 - Χρέωση ή πίστωση λογαριασμού
 - Προσθήκη νέου λογαριασμού εύρεση υπολοίπου
 - μηνιαίες/ετήσιες καταστάσεις κίνησης κ.λ.π.
 - Καινούργια αρχεία και προγράμματα εφαρμογών προστίθενται ανάλογα με τις ανάγκες



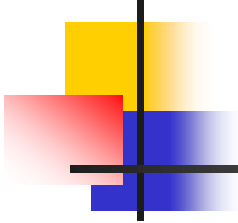
Συστήματα Διαχείρισης Βάσεων Δεδομένων (4/7)

- Η υλοποίηση ενός τέτοιου συστήματος μπορεί να βασιστεί σε ένα «τυπικό» σύστημα διαχείρισης και επεξεργασίας αρχείων (file processing system), πάσχει όμως από διάφορα μειονεκτήματα:
 - Πλεονασμός και Ασυνέπεια δεδομένων (data redundancy and inconsistency)
 - Η ίδια πληροφορία πιθανόν να επαναλαμβάνεται σε διαφορετικά αρχεία με αποτέλεσμα να υπάρχει μεγαλύτερο κόστος αποθήκευσης και ανάκλησης
 - Επιπλέον, ενημέρωση ενός αρχείου πρέπει να συνοδεύεται από την ενημέρωση όλων των αρχείων στα οποία η πληροφορία επαναλαμβάνεται.



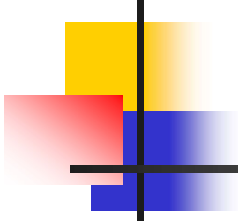
Συστήματα Διαχείρισης Βάσεων Δεδομένων (5/7)

- Δυσκολία στην πρόσβαση στα δεδομένα
 - Τα συστήματα διαχείρισης αρχείων δεν παρέχουν αποδοτικούς μηχανισμούς για αναζήτηση και ανάκληση πληροφορίας με διαφορετικούς τρόπους.
 - Επιπλέον, τα δεδομένα διασκορπίζονται σε αρχεία που χρησιμοποιούν διαφορετικές μορφές.
- Προβλήματα ακεραιότητας δεδομένων (data integrity)
 - Τα δεδομένα υπόκεινται σε περιορισμούς συνέπειας ή ακεραιότητας (consistency or integrity constraints).
 - Οι περιορισμοί αυτοί εφαρμόζονται μέσω των προγραμμάτων εφαρμογών τα οποία πρέπει να αλλαχθούν για την προσθήκη νέων ή τη μεταβολή περιορισμών.



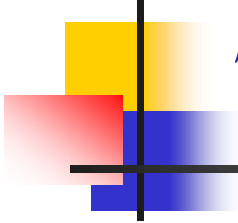
Συστήματα Διαχείρισης Βάσεων Δεδομένων (6/7)

- Προβλήματα ατομικότητας δοσοληψιών (transaction atomicity)
 - το σύστημα πρέπει να εγγυάται ότι σε περίπτωση σφάλματος η πληροφορία πρέπει να μπορεί να αποκατασταθεί.
 - Π.χ. μεταφορά ενός ποσού X από λογαριασμό A σε λογαριασμό B. Αν συμβεί κάποιο σφάλμα κατά τη διάρκεια της μεταφοράς, ο λογαριασμός A μπορεί να χρεωθεί χωρίς να πιστωθεί ο λογαριασμός B.
 - Οι δοσοληψίες πρέπει να χαρακτηρίζονται από ατομικότητα, δηλαδή, ή όλη η δοσοληψία ολοκληρώνεται ή κανένα μέρος αυτής.



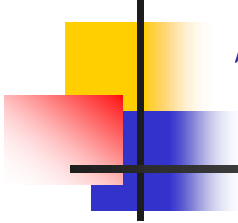
Συστήματα Διαχείρισης Βάσεων Δεδομένων (7/7)

- Ανωμαλίες ταυτόχρονης πρόσβασης (concurrent access)
 - Η ταυτόχρονη πρόσβαση και ενημέρωση της πληροφορίας από πολλούς χρήστες μπορεί να δημιουργήσει προβλήματα συνέπειας.
 - Π.χ. έστω ότι ο λογαριασμός Α έχει υπόλοιπο Χ. Αν δύο χρήστες αφαιρέσουν ποσά Δ1 και Δ2 "ταυτόχρονα" από τον Α, το αποτέλεσμα μπορεί να μην είναι ορθό.
- Προβλήματα ασφάλειας (security)
 - Διαφορετικοί χρήστες θα πρέπει να έχουν διαφορετικά δικαιώματα πρόσβασης σε αρχεία ή σε συγκεκριμένα δεδομένα.



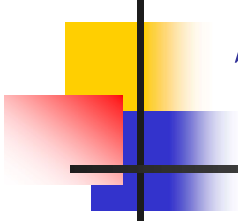
Όψεις Δεδομένων (1/3)

- Ένας από τους κύριους στόχους ενός ΣΔΒΔ είναι να παρέχει μια αφηρημένη όψη των δεδομένων αποκρύπτοντας λεπτομέρειες της αναπαράστασης και της αποθήκεσής τους στο σύστημα.
- Ένα ΣΔΒΔ είναι ένα πρόγραμμα για τη διαχείριση ηλεκτρονικών αρχείων επιχειρησιακών δεδομένων με δομημένο τρόπο.
- Για παράδειγμα
 - ΣΔΒΔ ενός πανεπιστημίου διαχειρίζεται αρχεία για τους φοιτητές, τις εγγραφές τους σε μαθήματα, βαθμολογία, κ.λ.π



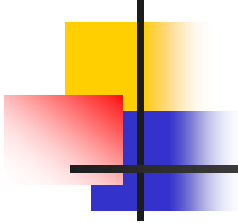
Όψεις Δεδομένων (2/3)

- Αεροπορικές εταιρίες χρησιμοποιούν ΣΔΒΔ για την αποθήκευση στοιχείων σχετικά με πτήσεις, δρομολόγια, διαθεσιμότητα θέσεων, κρατήσεις, προσωπικό, κ.λ.π
- Ένα ΣΔΒΔ μιας βιβλιοθήκης καταγράφει τα διαθέσιμα βιβλία, τα άτομα που τα δανείζονται, και παρέχει θεματική κατηγοριοποίηση.
- Η συλλογή των αρχείων και εγγραφών που κρατούνται για ένα συγκεκριμένο σκοπό ονομάζεται βάση δεδομένων (database).
- Κατά κανόνα, οι βάσεις δεδομένων αποθηκεύονται σε δευτερεύουσα μνήμη.
 - Ένα ΣΔΒΔ μπορεί να διαχειρίζεται περισσότερες από μία βάσεις δεδομένων.



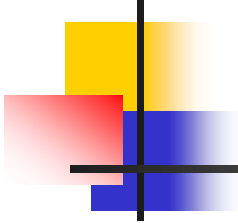
Όψεις Δεδομένων (3/3)

- Οι βάσεις δεδομένων μπορούν να μοιράζονται πληροφορία.
- Η πρόσβαση σε αυτές ελέγχεται από ένα κεντρικό πρόγραμμα το οποίο ονομάζεται διαχειριστής βάσεων δεδομένων (db manager).
- Η δόμηση της πληροφορίας που αποθηκεύεται στις βάσεις δεδομένων καθορίζεται από το μοντέλο δεδομένων (data model).



Μοντέλα Δεδομένων

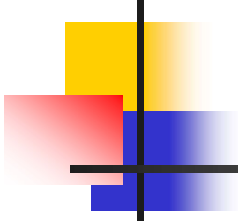
- Αρχικά μοντέλα δεδομένων:
 - Ιεραρχικό (hierarchical data model): διαφορετικοί τύποι εγγραφών δομούνται υπό μορφή ιεραρχίας (δένδρου)
 - Μοντέλο δικτύου (network data model): γενίκευση του ιεραρχικού μοντέλου.
 - Οι εγγραφές οργανώνονται υπό μορφή κατευθυνόμενου άκυκλου γράφου.
- Μειονέκτημα των αρχικών μοντέλων δεδομένων:
 - Η επερώτηση στις βάσεις δεδομένων γίνεται πολύπλοκη.



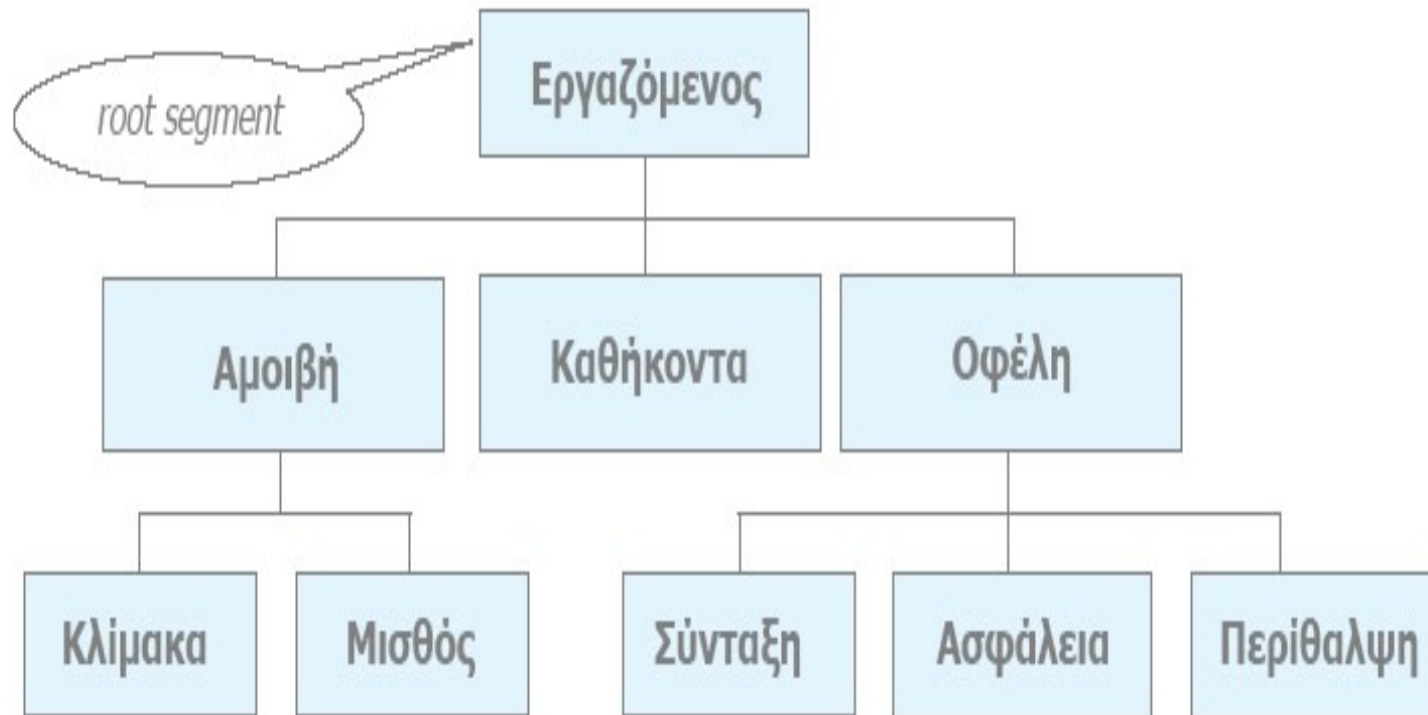
Ιεραρχικό Μοντέλο (1/4)

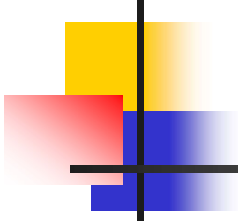
- Γενικά:

- Τα δεδομένα οργανώνονται σε μια δομή δέντρου
- Σε κάθε εγγραφή, τα δεδομένα οργανώνονται σε τμήματα (segments)
- Στον χρήστη, κάθε εγγραφή μοιάζει σαν ένα οργανόγραμμα με ένα βασικό τμήμα (root segment) στην κορυφή του
- Οργάνωση με βάση τη σχέση «πατέρας-παιδί»



Ιεραρχικό Μοντέλο (2/4)

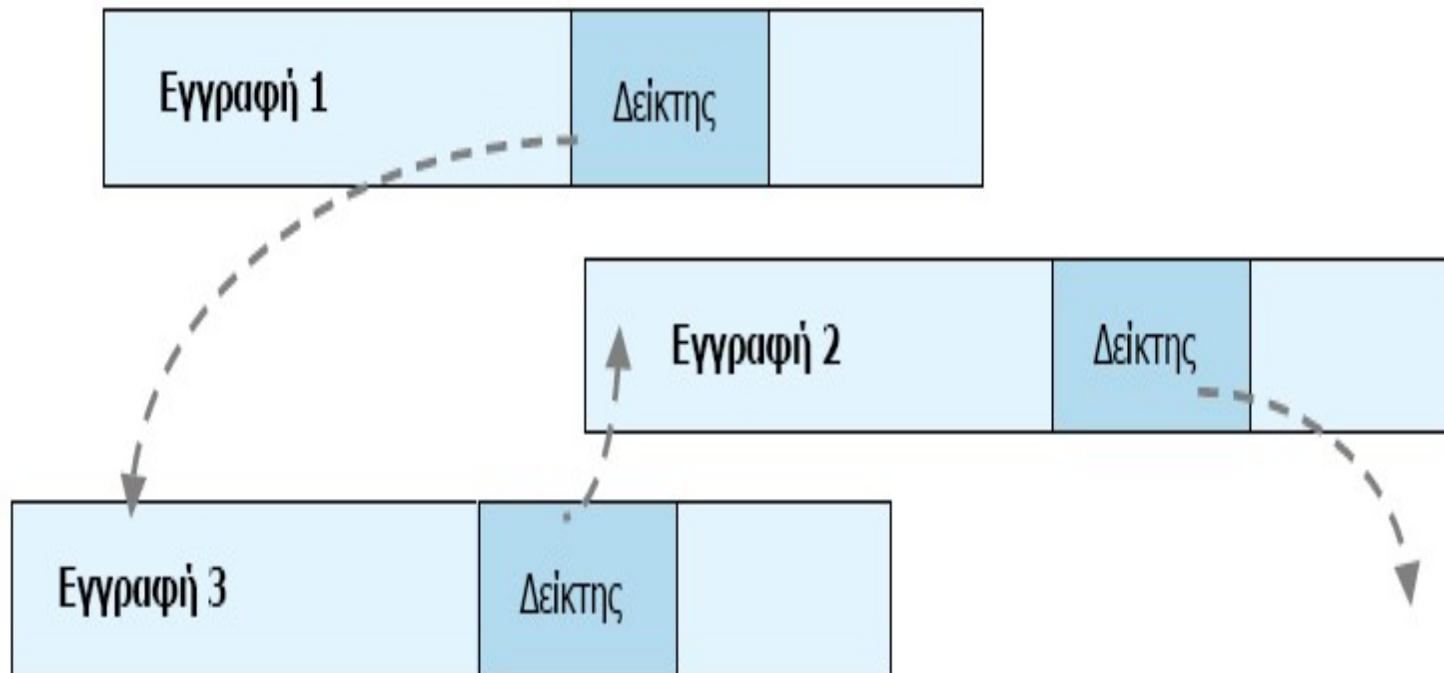


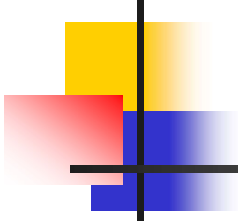


Ιεραρχικό Μοντέλο (3/4)

- Δείκτες (pointers):
 - Σε ένα ιεραρχικό ΣΔΒΔ, τα δεδομένα συνδέονται φυσικά μεταξύ τους μέσω μιας σειράς δεικτών που ορίζουν σύνολα από σχετικά τμήματα δεδομένων
 - Οι δείκτες αυτοί είναι δεδομένα που επισυνάπτονται στο τέλος ενός τμήματος δεδομένων (του «πατέρα») και δίνουν τη φυσική διεύθυνση των σχετικών τμημάτων δεδομένων (των «παιδιών»)

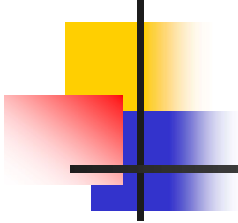
Ιεραρχικό Μοντέλο (4/4)



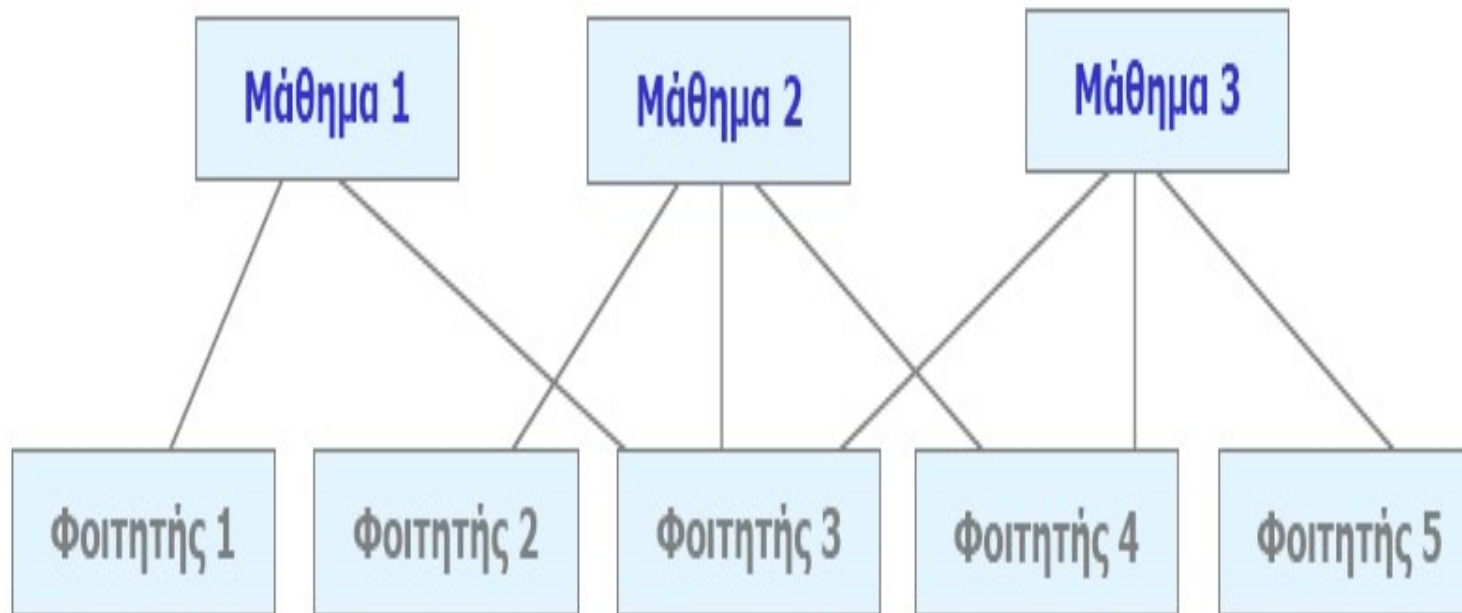


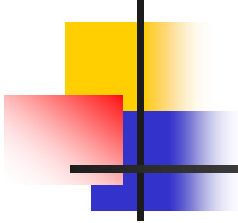
Δικτυακό Μοντέλο (1/2)

- Παραλλαγή του ιεραρχικού μοντέλου
- Ιδιαίτερα χρήσιμο για “many-to-many” σχέσεις
 - Ένα «παιδί» μπορεί να έχει πάνω από έναν «πατέρα»
- Πλεονεκτήματα και μειονεκτήματα
 - Ελαχιστοποίηση πλεονασμού δεδομένων και, συχνά, μείωση του χρόνου απόκρισης.
 - Ραγδαία αύξηση του αριθμού των δεικτών, με αποτέλεσμα η συντήρηση και λειτουργία του συστήματος να γίνεται συχνά πολύπλοκη.



Δικτυακό Μοντέλο (2/2)

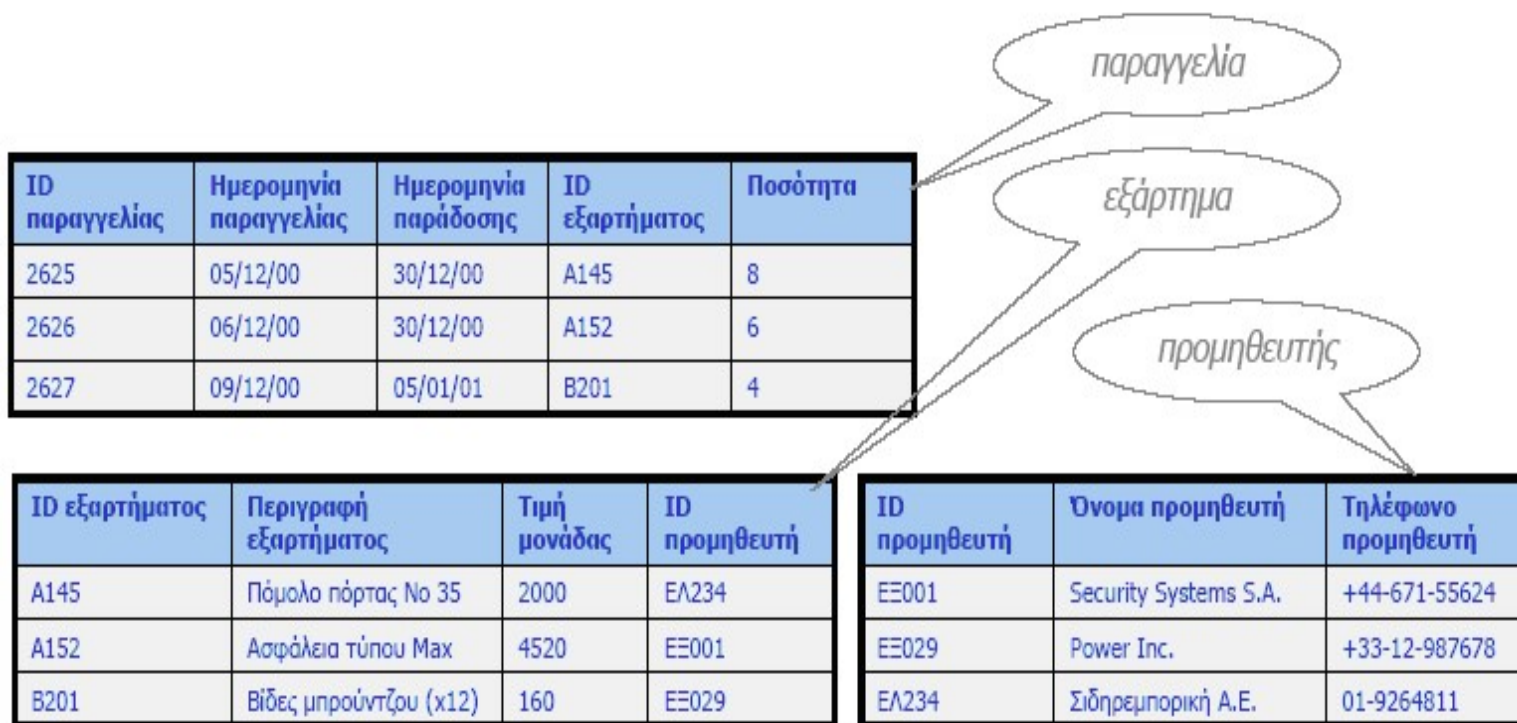


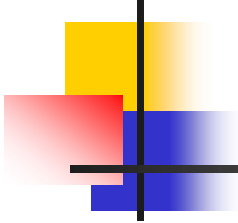


Σχεσιακό Μοντέλο (1/3)

- Το σχεσιακό μοντέλο (relational model) είναι το πλέον επιτυχημένο μοντέλο δεδομένων από αυτά έχουν προταθεί.
 - Τα δεδομένα οργανώνονται σε μορφή πινάκων εγγραφών ή σχέσεων (tables of records or relations).
 - Οι σχέσεις είναι συλλογές γνωρισμάτων (attributes), όπου κάθε γνώρισμα (μια στήλη του πίνακα) αναπαριστά μια ιδιότητα της εγγραφής.
 - Παρέχει εύκολο τρόπο επερώτησης των δεδομένων.
 - Εμπορικά σχεσιακά ΣΔΒΔ: Ingres, Oracle, DB2, Informix, Sybase, κλπ.

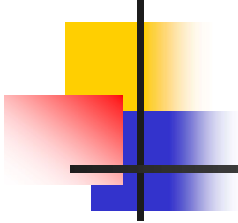
Σχεσιακό Μοντέλο (2/3)





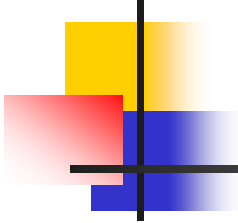
Σχεσιακό Μοντέλο (3/3)

Τύπος ΒΔ	Απόδοση επεξεργασίας δεδομένων	Ευελιξία	Φιλικότητα στον τελικό χρήστη	Πολυπλοκότητα στον προγραμματισμό
Ιεραρχική	Υψηλή	Χαμηλή	Χαμηλή	Υψηλή
Δικτυακή	Μέτρια προς υψηλή	Χαμηλή προς μέτρια	Χαμηλή προς μέτρια	Υψηλή
Σχεσιακή	Χαμηλότερη (βελτιώνεται)	Υψηλή	Υψηλή	Χαμηλή



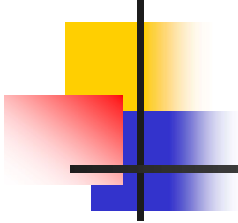
Χρήστες ΣΔΒΔ (1/2)

- Ένα ΣΔΒΔ πρέπει να παρέχει εύκολη πρόσβαση στα δεδομένα σε σχετικά μη έμπειρους χρήστες οι οποίοι λέγονται τελικοί χρήστες (end users).
- Η φιλικότητα προς το χρήστη και η χρήση μιας γλώσσας επερωτήσεων υψηλού επιπέδου είναι σημαντικά χαρακτηριστικά που πρέπει να διαθέτει ένα ΣΔΒΔ.
- Διαφορετικοί χρήστες έχουν διαφορετικές απαιτήσεις:
 - Τελικοί χρήστες
 - Προγραμματιστές εφαρμογών (application programmers)
 - Διαχειριστές ΒΔ (database administrators)



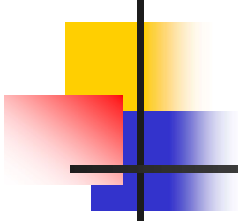
Χρήστες ΣΔΒΔ (2/2)

- Τελικοί χρήστες:
 - Συνήθεις ή περιστασιακοί χρήστες (casual users): χρησιμοποιούν μια δομημένη γλώσσα (π.χ. SQL) για την διατύπωση ερωτήσεων στη ΒΔ
 - Αδαείς χρήστες (naive users): διατυπώνουν ερωτήσεις χρησιμοποιώντας φόρμες ή μενού
- Προγραμματιστές εφαρμογών:
 - Γράφουν προγράμματα εφαρμογών για να εξυπηρετήσουν τις ανάγκες των τελικών χρηστών. Έχουν την ευθύνη της υλοποίησης ενός τρόπου επερώτησης με αποδοτικό τρόπο.
- Διαχειριστής ΒΔ: είναι υπεύθυνος για τη σχεδίαση και συντήρηση των ΒΔ



Επίπεδα Αφαίρεσης ΣΔΒΔ (1/5)

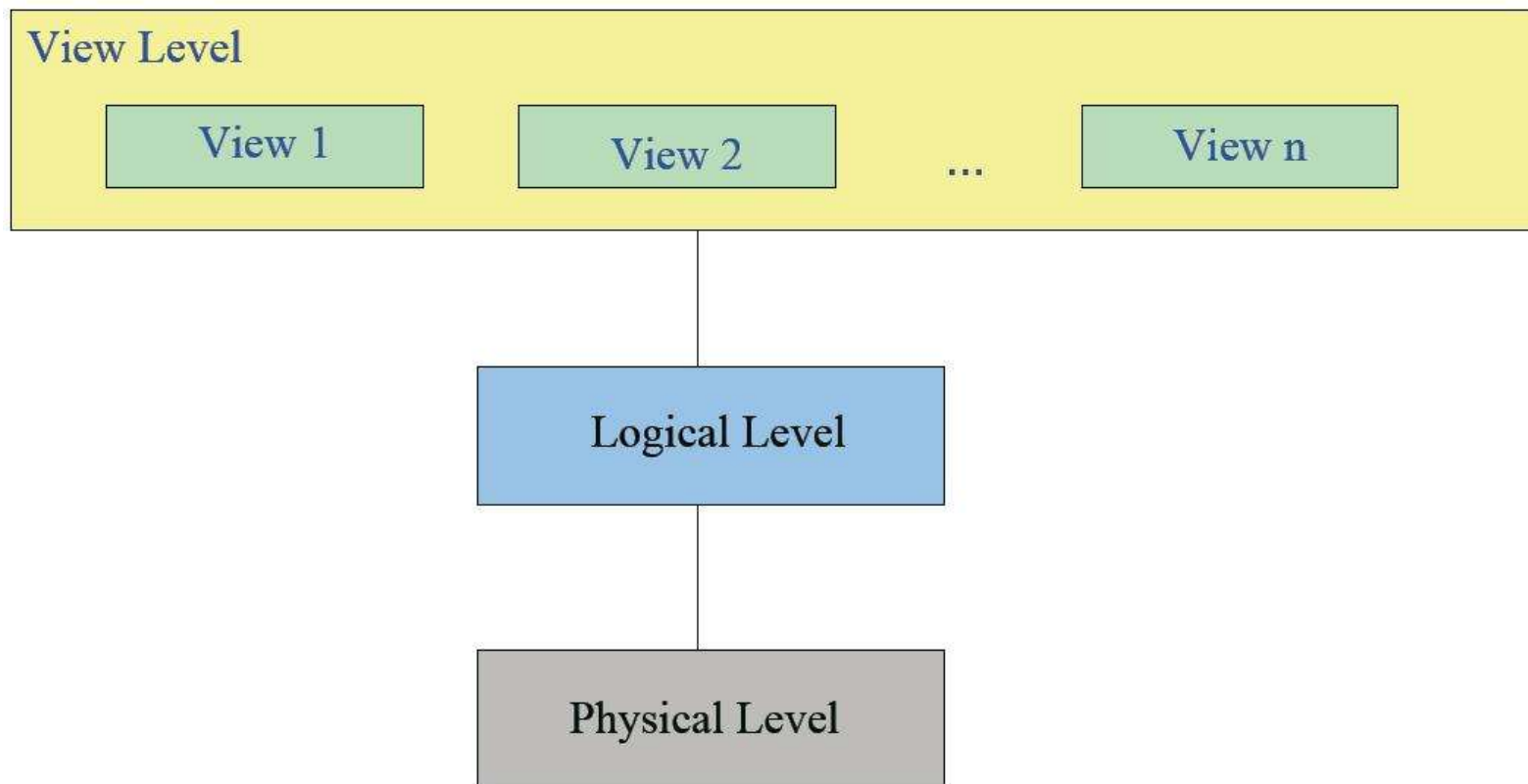
- Ένας από τους κύριους στόχους ενός ΣΔΒΔ είναι να παρέχει μια αφηρημένη όψη των δεδομένων, αποκρύπτοντας από το χρήστη λεπτομέρειες σχετικά με την αναπαράσταση και την αποθήκευσή τους.
- Σε ένα ΣΔΒΔ, τα δεδομένα αναπαρίστανται σε 3 επίπεδα αφαίρεσης (abstraction levels):
 - Φυσικό επίπεδο (physical level):
 - Το χαμηλότερο επίπεδο αφαίρεσης
 - Περιλαμβάνει σύνθετες δομές και λεπτομέρειες αναπαράστασης και αποθήκευσης

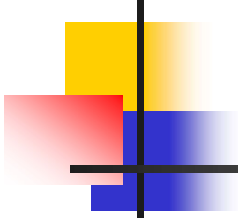


Επίπεδα Αφαίρεσης ΣΔΒΔ (2/5)

- Λογικό επίπεδο (logical level): περιγράφει τα δεδομένα και τις μεταξύ τους σχέσεις χρησιμοποιώντας ένα σχετικά μικρό αριθμό απλών δομών και εννοιών
 - Οι χρήστες οι οποίοι βλέπουν το σύστημα στο λογικό επίπεδο δεν χρειάζεται να γνωρίζουν τις λεπτομέρειες του φυσικού επιπέδου
- Επίπεδο όψεων (view level):
 - Το υψηλότερο επίπεδο αφαίρεσης
 - Περιγράφει κάποια μέρη της αποθηκευμένης πληροφορίας, καθώς κάποιοι χρήστες χρειάζονται πρόσβαση μόνο σε μέρος των ΒΔ
 - Αποτελεί συγχρόνως και μηχανισμό ασφάλειας

Επίπεδα Αφαίρεσης ΣΔΒΔ (3/5)

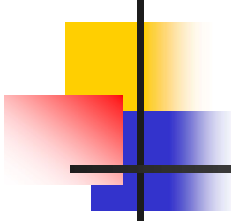




Επίπεδα Αφαίρεσης ΣΔΒΔ (4/5)

- Η διάκριση μεταξύ των διαφορετικών επιπέδων αφαίρεσης είναι ανάλογη με τη διάκριση μεταξύ επιπέδων αφαίρεσης στις γλώσσες προγραμματισμού
- Για παράδειγμα, σε μια γλώσσα προγραμματισμού, οι δηλώσεις:

```
type customer = record
  customer-name : string;
  customer-id : string;
  customer-address: string;
end;
type account=record
  account-number: integer;
end;
```



Επίπεδα Αφαίρεσης ΣΔΒΔ (5/5)

```
type employee=record  
    employee-name: string;  
    salary: integer;  
end;
```

- Αντιστοιχούν στο λογικό επίπεδο.
- Στο φυσικό επίπεδο, οι τύποι αυτοί αναπαριστώνται σαν συνεχόμενες θέσεις αποθήκευσης.
- Η λεπτομέρεια αυτή αποκρύπτεται από τον προγραμματιστή



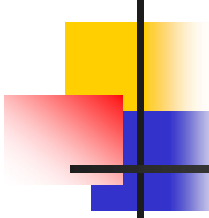
Σχήμα & Στιγμιότυπα

- Οι ΒΔ μεταβάλλονται με την πάροδο του χρόνου καθώς πληροφορία προστίθεται, αφαιρείται ή τροποποιείται.
- Το σύνολο της πληροφορίας το οποίο βρίσκεται αποθηκευμένο σε μια ΒΔ σε μια συγκεκριμένη χρονική στιγμή λέγεται στιγμιότυπο (instance or snapshot) της ΒΔ.
- Το σχήμα (schema) της ΒΔ είναι το σύνολο των εννοιών που περιγράφουν τη δομή της και είναι – σχεδόν πάντα – σταθερό.
- Αναλογία: record types $\leftrightarrow$ schema, variables $\leftrightarrow$ instances
- Μια ΒΔ έχει ένα (ή περισσότερα) σχήματα σε κάθε επίπεδο αφαίρεσης:
 - Φυσικό σχήμα, Λογικό σχήμα, Σχήματα όψεων ή υποσχήματα



Ανεξαρτησία Δεδομένων (1/2)

- Ο όρος ανεξαρτησία δεδομένων (data independence) χαρακτηρίζει την ικανότητα μεταβολής του σχήματος σε κάποιο επίπεδο αφαίρεσης, χωρίς να επηρεάζεται ο ορισμός του σχήματος στο αμέσως υψηλότερο επίπεδο
- Φυσική ανεξαρτησία δεδομένων:
 - Το φυσικό σχήμα μπορεί να μεταβάλλεται χωρίς να χρειάζεται να γραφούν ξανά τα προγράμματα εφαρμογών που έχουν πρόσβαση στη ΒΔ.
 - Μεταβολές στο φυσικό σχήμα είναι συχνά απαραίτητες για λόγους βελτιστοποίησης της απόδοσης του συστήματος



Ανεξαρτησία Δεδομένων (2/2)

- Λογική ανεξαρτησία δεδομένων:
 - Είναι η ικανότητα μεταβολής του λογικού σχήματος χωρίς την ανάγκη μεταβολής των προγραμμάτων εφαρμογών
 - Τέτοιες μεταβολές είναι απαραίτητες όταν μεταβάλλεται η εννοιολογική δομή μιας ΒΔ
- Η λογική ανεξαρτησία είναι δυσκολότερο να επιτευχθεί από τη φυσική ανεξαρτησία, καθώς τα προγράμματα εφαρμογών εξαρτώνται σε μεγάλο βαθμό από τη λογική δομή μιας ΒΔ.
- Η έννοια της ανεξαρτησίας είναι συγγενής με αυτή των αφηρημένων τύπων δεδομένων (abstract data types).



Μοντέλα Δεδομένων (1/3)

- Ένα μοντέλο δεδομένων (data model) είναι ένα σύνολο εννοιών για την περιγραφή των δεδομένων, των σχέσεων μεταξύ αυτών, τη σημασιολογία τους και τους περιορισμούς στους οποίους υπόκεινται.
- Διακρίνονται σε 3 κατηγορίες:
 - Οντοκεντρικά λογικά μοντέλα (object-oriented)
 - Πλειαδικά λογικά μοντέλα (record-based)
 - Φυσικά μοντέλα
- Τα οντοκεντρικά μοντέλα χρησιμοποιούνται για την περιγραφή δεδομένων στο λογικό επίπεδο και το επίπεδο όψεων.
- Τέτοια μοντέλα είναι τα:
 - Μοντελο οντοτήτων-σχέσεων(entity-relationship model), σημασιολογικό μοντέλο (semantic data model), συναρτησιακό μοντέλο (functional model)



Μοντέλα Δεδομένων (2/3)

- Τέτοια μοντέλα είναι τα:
 - Μοντέλο οντοτήτων-σχέσεων(entity-relationship model), σημασιολογικό μοντέλο (semantic data model), συναρτησιακό μοντέλο (functional model)
- Τα πλειαδικά λογικά μοντέλα αναπαριστούν τη λογική δομή μιας ΒΔ μέσω συγκεκριμένων τύπων που αναπαριστώνται ως πλειάδες(tuple s) ή εγγραφές(records).
- Τέτοια μοντέλα είναι τα:
 - Σχεσιακό μοντέλο (relational model), ιεραρχικό μοντέλο (hierarchical model), μοντέλο δικτύου(network model)



Μοντέλα Δεδομένων (3/3)

- Τα φυσικά μοντέλα χρησιμοποιούνται για την περιγραφή των δεδομένων στο φυσικόεπίπεδο.
- Τέτοια μοντέλα είναι τα:
 - Ενοποιημένο μοντέλο (unifying data model)
 - Μοντέλο πλαισίων μνήμης (frame-memory model)



Γλώσσες Βάσεων Δεδομένων (1/2)

- Γλώσσες για τον ορισμό δεδομένων (data definition languages)
 - Χρησιμοποιούνται για τον ορισμό του σχήματος
 - Εντολές της γλώσσας ορισμού δεδομένων μεταφράζονται σε ένα σύνολο δομών που αποθηκεύονται στο λεξικό ή ευρετήριο δεδομένων (data dictionary or directory)
 - Το ευρετήριο δεδομένων περιέχει μεταδεδομένα (metadata), δηλαδή δεδομένα για τα δεδομένα. Το ΣΔΒΔ συμβουλεύεται το ευρετήριο πριν από οποιαδήποτε πρόσβαση στη ΒΔ.
 - Μέθοδοι αποθήκευσης και πρόσβασης καθορίζονται σε μια ειδική γλώσσα ορισμού δεδομένων η οποία λέγεται γλώσσα ορισμού αποθήκευσης δεδομένων (data storage definition language)



Γλώσσες Βάσεων Δεδομένων (2/2)

- Γλώσσες για το χειρισμό δεδομένων (data manipulation languages)
 - Υποστηρίζουν λειτουργίες για εισαγωγή, ανάκληση, μεταβολή και διαγραφή πληροφορίας
 - Διακρίνονται σε διεργασιακές (procedural) και δηλωτικές (declarative)
 - Διεργασιακές γλώσσες καθορίζουν το «πως» γίνεται ο χειρισμός των δεδομένων, ενώ οι δηλωτικές καθορίζουν «ποιο» είναι το επιθυμητό αποτέλεσμα
 - Οι δηλωτικές γλώσσες είναι ευκολότερες στη χρήση τους αλλά όχι τόσο αποδοτικές
 - Η γλώσσα επρωτήσεων (query language) είναι υποσύνολο της γλώσσας χειρισμού δεδομένων

Αρχιτεκτονική ΣΔΒΔ

