

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Θεολογική Σχολή
Τμήμα: Θεολογίας
Τομέας: Κοινωνιολογία – Πληροφορική

Ορολογία και αναπαράσταση της πληροφορίας στις Βάσεις Δεδομένων

Εργασία του Παυλίδη Σάββα,
μεταπτυχιακού φοιτητή AM 1150
Υπεύθυνος καθηγητής: Μυρίδης Νικόλαος

Θεσσαλονίκη
Σεπτέμβριος 2007

Περιεχόμενα

ΕΙΣΑΓΩΓΗ	3
1. ΙΣΤΟΡΙΑ	8
2. Η ΠΛΗΡΟΦΟΡΙΑ	10
3. ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ	14
3.1. Βασικές αρχές σχεδίασης βάσης δεδομένων	14
3.2. Χαρακτηριστικά της βάσης δεδομένων	17
3.3. Συστήματα Διαχείρισης των βάσεων δεδομένων	18
4. ΜΟΝΤΕΛΑ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ	19
4.1. Μοντέλο οντοτήτων συσχετίσεων	19
4.2. Σχεσιακό μοντέλο	21
5. ΟΡΟΛΟΓΙΑ	26
6. ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΣΤΙΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ	35
6.1. Ιεραρχικό μοντέλο δεδομένων	35
6.1. Δικτυακό μοντέλο δεδομένων	37
6.2. Σχεσιακό μοντέλο δεδομένων	39
7. ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΘΕΟΛΟΓΙΚΗ ΕΠΙΣΤΗΜΗ	36
7.1. Οι Βάσεις Δεδομένων στη θεολογική επιστήμη	40
ΒΙΒΛΙΟΓΡΑΦΙΑ	42

ΕΙΣΑΓΩΓΗ

Ο άνθρωπος κατά τη διάρκεια της ζωής του βρίσκεται σε συνεχή επικοινωνία με το περιβάλλον του. Η επικοινωνία αυτή μπορεί να γίνει με πολλούς τρόπους όπως: ομιλία, γραπτός λόγος, αισθήσεις. Το κυριότερο χαρακτηριστικό αυτής της επικοινωνίας είναι η ανταλλαγή πληροφοριών ανάμεσα στον κάθε άνθρωπο και το χώρο μέσα στον οποίο ζει και κινείται. Ο άνθρωπος είναι τόσο πομπός όσο και δέκτης της πληροφορίας, δηλαδή μπορεί να μεταδώσει και να λάβει πληροφορίες. Η πληροφορία, έχει να κάνει με τα στοιχεία εκείνα που μεταδίδονται από μια πηγή προς κάποιον δέκτη. Έτσι, ανάλογα με τη σκοπιά κάτω από την οποία προσεγγίζει κανείς την πληροφορία, αυτή μπορεί να έχει να κάνει με την ενημέρωση αν η σκοπιά είναι επικοινωνιακή, για τη διάκριση μεταξύ σημαίνοντος και σημαινόμενου αν την εξετάσουμε υπό το πρίσμα της γλωσσολογίας¹ ή ακόμα και για ένα τηλεφωνικό ειδοποιητήριο όπως συχνά συμβαίνει στην καθημερινή ζωή.

Ουσιαστικά, προσεγγίζοντας την έννοια της πληροφορίας διαισθητικά, διαπιστώνουμε πως γενικά παραπέμπει σε καινούργια γνώση για κάτι. Επιπλέον, αν και η «σημαίνουσα» (Η προσέγγιση που επιχειρείται στην έννοια της πληροφορίας, βασίζεται στη θεωρία του Ferdinand de Saussure για τη γλώσσα όπου διακρίνει σε μια λέξη το σημαίνον που αντιπροσωπεύει την εικόνα ή ακουστική εικόνα της λέξης και το σημαινόμενο που αντιπροσωπεύει την ιδέα, το νόημα που κρύβεται πίσω από τη λέξη) αξία της πληροφορίας υφίσταται αναλλοίωτα, το «σημαινόμενο» αυτής είναι δυνατόν να μεταβάλλεται κάθε φορά με τρόπο καθοριστικό για την ίδια την πληροφορία. Ο τρόπος που το «σημαινόμενο» μεταβάλλεται καθορίζεται από τους τρεις παράγοντες που αφενός στοιχειοθετούν και αφετέρου δίνουν υπόσταση στην πληροφορία: την πηγή, το δέκτη και το μέσο διάδοσης. Δηλαδή, αν μια πληροφορία υφίσταται κατ' αρχήν έχει νόημα μόνο αν την γνωρίσει κάποιος, επιπλέον, ενώ η πληροφορία υφίσταται (σημαίνον) το νόημά της (ή περιεχόμενό της ή σημαινόμενο) εξαρτάται από:

- το πώς θα το μεταδώσει η πηγή,
- το πόσο καλά (ή αξιόπιστα) θα το μεταφέρει το μέσο προς το δέκτη και
- το πώς θα το αντιληφθεί ο δέκτης.

¹ F. D. Saussure, Μαθήματα γενικής γλωσσολογίας, εκδ. Παπαζήση (1979). Μετάφραση 3ης γαλλ. εκδ. (Γενεύη 1915)

Η παραπάνω αξίωση θεωρούμε πως έχει καθολική εφαρμογή από την καθημερινή ζωή μέχρι όλων των ειδών τις επιστήμες και δεν είναι δύσκολο να το διαπιστώσει κανείς είτε μιλάμε για περιπτώσεις ανθρώπων (πομπός και δέκτης: άνθρωπος, μέσο: π.χ. φωνή, λέξη κλπ) είτε για περιπτώσεις κυκλωμάτων (πομπός: κύκλωμα ή διάταξη A, δέκτης: κύκλωμα ή διάταξη B, μέσο: π.χ. καλώδια ή διασυνδέσεις ή κύματα).

Οι πληροφορίες αυτές μπορεί να είναι οποιασδήποτε μορφής και να αφορούν σκέψεις, περιγραφές, συναισθήματα. Μπορούν να παραμένουν ίδιες ή να αλλάζουν με το χρόνο. Σε όλες, όμως, τις περιπτώσεις έχουν ένα κοινό χαρακτηριστικό: μεταφέρουν κάποιο περιεχόμενο. Το περιεχόμενο αυτό εξαρτάται άμεσα από το φυσικό σύστημα στο οποίο αναφέρεται: ένα λεκτικό μήνυμα σε φυσική γλώσσα, ένα πίνακας αριθμητικών αποτελεσμάτων ή μετρήσεων, μια μαθηματική έκφραση, μια εικόνα ή ένα κομμάτι μουσικής, η οσμή της φερομόνης ενός εντόμου ή το DNA ενός κυττάρου θα μπορούσαν να αποτελούν οντότητες πληροφορίας. Πέραν του φιλοσοφικού πλαισίου, πρακτικά μπορούμε να «διαχειριστούμε» την πληροφορία μεταδίδοντάς την (transmission), αποθηκευόντάς τη (storage) ή με επεξεργασία της (processing) ώστε να τη μεταβάλλουμε (δηλαδή να παράγουμε καινούργια πληροφορία μεταβάλλοντάς την).

Οι θεωρητικοί C. S. Pierce (1839 – 1914) και C. W. Morris (1901 – 1979)² πραγματοποίησαν την εξής διάκριση της πληροφορίας αναφορικά με την αλληλεπίδρασή της με τα διάφορα δίπολα πομπού – δέκτη:

Συντακτικό επίπεδο (syntactic level): Η πληροφορία σε αυτό το επίπεδο έχει να κάνει με τους επίσημους δεσμούς που υπάρχουν μεταξύ των διαφόρων στοιχείων που συνθέτουν την πληροφορία, τους κανόνες που διέπουν τον κώδικα επικοινωνίας, τη χωρητικότητα των διαύλων επικοινωνίας και το σχεδιασμό συστημάτων και μεθόδων κωδικοποίησης για μετάδοση, επεξεργασία και αποθήκευση της πληροφορίας.

Εννοιολογικό επίπεδο (semantic level): Το επίπεδο αυτό έχει να κάνει με το πώς διαμορφώνεται η έννοια – το νόημα της πληροφορίας. Όσον αφορά τις πληροφορίες που δίνονται σε φυσική γλώσσα, αυτές βασίζονται στις συμφωνημένες, γραπτές ή άγραφες, πολιτισμικές, πολιτιστικές, ηθικές ή απλά συμφωνημένες συμβάσεις που κάνουν μεταξύ τους τα μέλη μιας ομάδας ανθρώπων. Στις διάφορες εννοιολογικές μονάδες (semantic units) που μπορεί να είναι, για παράδειγμα, οι

² <http://el.wikipedia.org/wiki/Πληροφορία>

λέξεις που στοιχειοθετούν μια πρόταση έχει προσδοθεί μια περισσότερη ή λιγότερη ακριβής ή ελεύθερη έννοια. Στην περίπτωση των πιο τεχνικών ή μηχανοποιημένων γλωσσών, όπως είναι οι κώδικες των υπολογιστών, το νόημα των εννοιολογικών μονάδων, είναι αμφιμονοσήμαντα ορισμένο βάσει των ιδιοτήτων που παρουσιάζουν και των λειτουργιών που μπορούν να εφαρμόσουν. Σύμφωνα με τον Shannon, τον πατέρα των σύγχρονων επικοινωνιών και της τεχνολογικής αντιμετώπισης της πληροφορίας, οι έννοιες (semantics), τεχνολογικά τουλάχιστον, δεν είναι προαπαιτούμενο για τη σωστή χρήση επεξεργασία της σύνταξης της πληροφορίας (syntactics of information), παρόλα αυτά μπορούν να φανούν χρήσιμες σε περιπτώσεις όπως η συμπίεση δεδομένων (με στόχο την αύξηση της αποδοτικότητας της συμπίεσης).

Πραγματιστικό επίπεδο (pragmatic level): Εδώ, η πληροφορία σχετίζεται με την αξία της χρησιμότητάς της. Το επίπεδο καθορίζεται σε μεγάλο βαθμό από το υπόβαθρο του λήπτη με αποτέλεσμα να επηρεάζεται σημαντικά από οικονομικούς, πολιτικούς, κοινωνικούς ή/και ψυχολογικούς παράγοντες ενώ πολλές φορές αποτελεί συνάρτηση του χρόνου αφού πολλές φορές οι πληροφορίες που φτάνουν με καθυστέρηση δεν αξίζουν πάντα, ενώ οι πληροφορίες που φτάνουν εγκαίρως μπορεί να είναι εξαιρετικά πολύτιμες.

Τα τρία αυτά επίπεδα αντιπροσωπεύουν μια ιεραρχία όπου επιτρέπεται η διαχείριση της πληροφορίας (μεταφορά, επεξεργασία, αποθήκευση) στο συντακτικό επίπεδο χωρίς να είναι απαραίτητη η γνώση του εννοιολογικού περιεχομένου της πληροφορίας. Επιπλέον, βάσει αυτής της ιεραρχίας έχουμε τη δυνατότητα να διαχειριζόμαστε τις πληροφορίες βάσει του εννοιολογικού της περιεχομένου χωρίς να λαμβάνουμε υπόψη την πρακτική τους χρησιμότητα.

Όσον αφορά τώρα τα αντικείμενα που στοιχειοθετούν την πληροφορία, αυτά είναι συμβολικής εκφράσεις μιας δεδομένης, κάθε φορά, γλώσσας και τα οποία φέρουν το νόημα της πληροφορίας και την πραγματιστική της αξία. Ανάλογα με το είδος της πληροφορίας, οι συμβολικές αυτές εκφράσεις μπορεί να αποτελούνται από λέξεις που σχηματίζουν προτάσεις, από νότες που σχηματίζουν μελωδίες ή από χρωματισμένες περιοχές που σχηματίζουν εικόνες. Από αυτά τα στοιχεία, τα στοιχειωδέστερα, δηλαδή αυτά που δε μερίζονται επιπλέον, καλούνται χαρακτήρες (characters ή tokens).

Οι φυσικές γλώσσες, τα σήματα Μορς κ.λ.π. αποτελούνται από ομάδες διακριτών (discrete) χαρακτήρων κάτι που μοιάζει ασύμβατο για την περίπτωση, ως

πούμε, της κλασικής μουσικής που αποτελεί ουσιαστικά μια «γλώσσα» που αποτελείται από τόνους διαφορετικών και διακριτών συχνοτήτων, πλατών και διαμορφώσεων. Στην περίπτωση αυτή και στον ήχο γενικότερα έχουμε μια συνεχή (continuous) κλίμακα συχνοτήτων και πλατών. Αντίστοιχα, οι εικόνες αποτελούνται από εικονοστοιχεία (pixels), τα οποία χαρακτηρίζονται από το χρώμα και τη φωτεινότητά τους μεγέθη που μπορεί επίσης να τοποθετούνται σε συνεχείς κλίμακες.

Για να μπορεί να γίνει διαχείριση της πληροφορίας, δηλαδή για να μπορεί η πληροφορία να μεταδοθεί, να γίνει προϊόν επεξεργασίας και για να αποθηκευτεί είναι αναγκαία η ύπαρξη ενός φορέα (carrier) της πληροφορίας, καθώς και ένα μέσο (medium) που θα χρησιμοποιήσει ο φορέας. Η πληροφορία, κωδικοποιείται στον

φορέα αναδομώντας ή σχηματοποιώντας τον στο χώρο ή το χρόνο. Οι δομές των φορέων της πληροφορίας που προκύπτουν με αυτό τον τρόπο καλούνται σήματα (signals). Έτσι λοιπόν, σήματα μπορεί να θεωρηθούν οι διάφορες δομές και σχηματισμοί στο χώρο που δημιουργούνται από μελάνι πάνω στο φορέα που είναι το χαρτί και που απεικονίζουν γράμματα ή παραστάσεις. Αν το φως έχει το ρόλο φορέα, τότε η

Τύπος πληροφορίας	Βασικοί χαρακτήρες	Σύνολο βασικών χαρακτήρων
Κώδικας Μοίσε	• □ κενό	3
Γραπτή Αγγλική γλώσσα	Γράμματα: a b c ... z A B C ... Z Ψηφία: 0 1 2 ... 9 Σημεία στίξης: . , ! ? ...	111 τυπώσιμοι χαρακτήρες ASCII
Γραπτή κινεζική, προσαρμοσμένη για μέσα εκτύπωσης	房屋问题谱...	1945
Δεκαδικοί αριθμοί	0 1 2 3 4 5 6 7 8 9 0 1	10
Δωαδικοί αριθμοί	ή αληθώς ψευδώς	2
Δεκαεξάδικοι αριθμοί	0 1 2 3 4 5 6 7 8 9 A B C D E F	16
Γενετικός κώδικας DNA	A C G T Αδενίνη: Κυτοσίνη: Γουανίνη: Θυμίνη:	4
Κλασική μουσική	Συχνότητες: 12 νότες ανά οκτάβα 6 οκτάβες Πλάτη: 5 επίπεδα	12 × 6 5
Ηχος	Διάρκειες: 6 αζίτες	6
Μυρωδιάς	Συχνότητες: Φάσμα από 20 Hz έως 20 kHz Πλάτη: Φάσμα από 0 dB έως 120 dB Χημικές συνθέσεις: ? Πηκτικές ? Διεγερτικές του κέντρου όσφρησης του εγκεφάλου	Συνεχείς > 10 ⁵

Παραδείγματα τύπων πληροφορίας και των χαρακτήρων που τα στοιχειοθετούν

πληροφορία θα κωδικοποιείται μέσω σημάτων που θα

Εικόνα 1: παραδείγματα τύπων πληροφορίας και των χαρακτήρων που τα στοιχειοθετούν.

μπορούσαν να είναι μεταβολές στην ένταση του φωτός (διαμόρφωση πλάτους) ή μεταβολές του μήκους κύματος (διαμόρφωση συχνότητας). Οι νευρώνες του εγκεφάλου επίσης, αποτελούν το μέσο διάδοσης των ηλεκτρικών σημάτων που μεταφέρουν τις διάφορες διεγέρσεις που λαμβάνουν χώρα στο νευρικό σύστημα ενός οργανισμού. Είναι προφανές πως υπάρχουν φορείς και μέσα που μπορεί να είναι περισσότερο ή λιγότερο κατάλληλοι να πραγματοποιήσουν διάφορες λειτουργίας

πάνω στην εκάστοτε υπάρχουσα πληροφορία. Έτσι, τα ηλεκτρομαγνητικά κύματα είναι ιδιαιτέρως κατάλληλα να μεταδώσουν πληροφορία αλλά είναι ανεπαρκή για να την αποθηκεύσουν. Τα μαγνητικά υλικά από την άλλη, αποτελούν πολύ καλή λύση για την αποθήκευση της πληροφορίας αλλά διόλου άριστα για να τη μεταδώσουν.

Στο σημείο αυτό έχοντας πλέον μια περισσότερο ολοκληρωμένη εικόνα της έννοιας της πληροφορίας μπορούμε να καταλήξουμε στο ότι μπορούμε να αλλάξουμε τη σύνταξη της πληροφορίας (syntactics) χωρίς να επηρεάζεται το νόημά της (semantics) δηλαδή για παράδειγμα μπορούμε να μετατρέψουμε τους χαρακτήρες του Αγγλικού αλφαβήτου σε κώδικα κατά τον τρόπο που ορίζει το πρότυπο ASCII – American Standard Code for Information Interchange ή να μετατρέψουμε ένα ηχητικό απόσπασμα σε δυαδικό κώδικα κλπ. Οι μετατροπές αυτές μπορεί να είναι ολοκληρωμένες και απόλυτες όπως στην πρώτη περίπτωση ή ατελείς όπως στη δεύτερη περίπτωση και γενικά σε κάθε προσπάθεια κωδικοποίησης ενός συνεχούς σήματος (αναλογικού) όπως είναι ο ήχος. Επιπλέον, αν και η πληροφορία πάντα χρειάζεται κάποιον φορέα, το είδος του φορέα δεν είναι άμεσης σημασίας πάντοτε αφού για παράδειγμα, μια δυαδική πληροφορία μπορεί να μεταφερθεί χωρίς πρόβλημα από ένα μαγνητικό δίσκο σε μια ημιαγωγική μνήμη και αντίστροφα χωρίς να μεταβάλλεται η πληροφορία (ούτε η σύνταξη –syntactics- ούτε το περιεχόμενό της –semantics). Βέβαια, όπως είναι λογικό, ο κάθε φορέας ορίζει το πώς πρέπει να δομείται η πληροφορία που θα τον χρησιμοποιήσει και κατά συνέπεια καθορίζει το είδος των κωδικών για τους οποίους είναι κατάλληλος. Έτσι, κατάλληλοι φορείς για τα γράμματα θα μπορούσαν να είναι χάρτινες ή κατάλληλες πλαστικές κλπ. επιφάνειες ή κάποιος τύπος οθόνης αλλά όχι ένα μαγνητικό μέσο, το οποίο απαιτεί μετατροπή του σε άλλο κώδικα πριν μεταφερθούν σε αυτό.

1. ΙΣΤΟΡΙΑ

Η παλαιότερη γνωστή χρήση του όρου βάση δεδομένων ήταν το Νοέμβριο του 1963, όταν η εταιρία ανάπτυξης System Development Corporation υποστήριξε ένα συμπόσιο με τίτλο «Development and Management of a Computer-centered Data Base». Η βάση δεδομένων ως μεμονωμένη λέξη έγινε κοινή στην Ευρώπη στις αρχές του 1970 και μέχρι το τέλος της δεκαετίας χρησιμοποιούνταν σε σημαντικές αμερικανικές εφημερίδες. (Ο όρος «τράπεζα δεδομένων», χρησιμοποιούνταν από την εφημερίδα Washington Post από το 1966.).

Τα πρώτα συστήματα διαχείρισης βάσεων δεδομένων αναπτύχθηκαν στη δεκαετία του '60. Ένας πρωτοπόρος στον τομέα ήταν ο Charles Bachman. Τα πρώιμα συγγράμματα του Bachman δείχνουν ότι ο στόχος του ήταν να κάνει αποτελεσματικότερη τη χρήση των νέων συσκευών αποθήκευσης άμεσης πρόσβασης: μέχρι τότε, στοιχεία - η επεξεργασία δεδομένων ήταν βασισμένη στις τρυπημένες με διατρητική μηχανή κάρτες και τη μαγνητική ταινία, και η τμηματική επεξεργασία ήταν η κυρίαρχη δραστηριότητα.

Δύο βασικά πρότυπα βάσεων προέκυψαν:

- CODASYL : ανέπτυξε το δικτυακό μοντέλο, βασισμένο στις ιδέες Bachman, και (προφανώς ανεξάρτητα) το ιεραρχικό μοντέλο χρησιμοποιήθηκε σε ένα σύστημα που αναπτύχθηκε από την North American Rockwell, που υιοθετήθηκε αργότερα από την IBM ως ακρογωνιαίος λίθος του προϊόντος IMS τους. Ενώ το IMS μαζί με το CODASYL IDMS ήταν οι μεγάλες βάσεις δεδομένων που αναπτύχθηκαν στη δεκαετία του '60, διάφορες άλλες γεννήθηκαν επίσης σε εκείνη την δεκαετία. Δύο άξιες λόγου είναι οι βάσεις δεδομένων PICK και MUMPS, που αναπτύχθηκαν αρχικά ως λειτουργικά συστήματα με μια ενσωματωμένη βάση δεδομένων και αργότερα ως γλώσσες προγραμματισμού και βάσεις δεδομένων για την ανάπτυξη του data-based λογισμικού.
- Το σχεσιακό μοντέλο προτάθηκε από E. Φ. Codd το 1970. Επέκρινε τα υπάρχοντα μοντέλα για τη σύγκριση της αφηρημένης περιγραφής της δομής πληροφοριών με τις περιγραφές των φυσικών μηχανισμών πρόσβασης. Για αρκετό καιρό, εντούτοις, το σχεσιακό μοντέλο προκαλούσε το ακαδημαϊκό ενδιαφέρον μόνο. Ενώ τα προϊόντα CODASYL (IDMS) και τα προϊόντα του

δικτυακού μοντέλου (IMS) θεωρούνταν ως πρακτικές λύσεις εφαρμοσμένης μηχανικής που λαμβάνουν υπόψη την τεχνολογία όπως υπήρχε τότε. Μεταξύ των πρώτων εφαρμογών, του σχεσιακού μοντέλου, ήταν το Michael Stonebraker's Ingres στο Μπέρκλεϋ, και το πρόγραμμα συστημάτων P στην IBM. Και τα δύο ήταν ερευνητικά πρωτότυπα, που ανακοινώθηκαν κατά τη διάρκεια του 1976. Τα πρώτα εμπορικά προϊόντα, Oracle και DB2, εμφανίστηκαν περίπου το 1980. Το πρώτο επιτυχές προϊόν βάσεων δεδομένων για τους μικροϋπολογιστές ήταν dBASE για τα λειτουργικά συστήματα του MS-dos CP/M και PC-DOS.

Κατά τη διάρκεια της δεκαετίας του '80, η ερευνητική δραστηριότητα εστίασε στα συστήματα διανεμημένων βάσεων δεδομένων και τις μηχανές βάσεων δεδομένων, αλλά αυτές οι εξελίξεις είχαν λίγη επίδραση στην αγορά. Μια άλλη σημαντική θεωρητική ιδέα ήταν το Functional Data Model, αλλά εκτός από μερικές εξειδικευμένες εφαρμογές στη γενετική, η μοριακή βιολογία, και η έρευνα απάτης, δεν είχαν μεγαλύτερη απήχηση.

Στη δεκαετία του '90, η προσοχή μετατοπίστηκε στις αντικειμενοστρεφείς βάσεις δεδομένων. Αυτές είχαν κάποια επιτυχία στους τομείς όπου ήταν απαραίτητο να αντιμετωπιστούν τα πιο σύνθετα δεδομένα από τα σχεσιακά μοντέλα, όπως οι χωρικές βάσεις δεδομένων, τα δεδομένα εφαρμοσμένης μηχανικής (συμπεριλαμβανομένων των αποθηκών τεχνολογίας λογισμικού), και τα δεδομένα πολυμέσων. Μερικές από αυτές τις ιδέες υιοθετήθηκαν από τους σχετικούς προμηθευτές, οι οποίοι ενσωμάτωσαν τα νέα χαρακτηριστικά γνωρίσματα στα προϊόντα τους. Η δεκαετία του '90 είδε επίσης τη διάδοση των Open Source databases, όπως PostgreSQL και MySQL³.

Από το 2000, η μοντέρνα περιοχή για την καινοτομία είναι η βάση δεδομένων XML. Όπως με τις αντικειμενοστρεφείς βάσεις δεδομένων, το πρότυπο αυτό έχει δημιουργήσει μια νέα ομάδα νεοσύστατων εταιρειών, αλλά συγχρόνως οι βασικές ιδέες ενσωματώνονται στα καθιερωμένα σχετικά προϊόντα. Οι βάσεις δεδομένων XML στοχεύουν να αφαιρέσουν την παραδοσιακή διαίρεση μεταξύ των εγγράφων και των στοιχείων, που επιτρέπουν στις πηγές πληροφοριών μιας όλης οργάνωσης για να κρατηθούν σε μια θέση, εάν είναι ιδιαίτερα δομημένες ή όχι.

³ <http://en.wikipedia.org/wiki/Database>

2. Η ΠΛΗΡΟΦΟΡΙΑ

Αποτελεί απαραίτητο συστατικό κάθε φυσικού συστήματος

2.1. Ιδιότητες και ιδιαίτερα χαρακτηριστικά της πληροφορίας

Η πληροφορία αποτελεί αφηρημένη οντότητα με χαρακτηριστικά που δεν συναντώνται στα υλικά αγαθά. Παρακάτω καταγράφονται ορισμένα χαρακτηριστικά που αποτελούν ιδιαιτερότητες της πληροφορίας η οποία εκλαμβάνεται ως άυλο αγαθό έναντι των υλικών αγαθών.

1. **Πληροφορία – επικοινωνία:** Η έννοια πληροφορία δεν είναι αυθύπαρκτη, αλλά είναι στενά συνδεδεμένη με την έννοια επικοινωνία και μετάδοση. Η πληροφορία και η επικοινωνία αποτελούν σημαντικά δομικά στοιχεία της κοινωνικής οργάνωσης.
2. **Μη αναλωσιμότητα:** Σε όρους περιεχομένου νοήματος, η πληροφορία χαρακτηρίζεται από μη αναλωσιμότητα. Το περιεχόμενο νόημα δεν αποτελεί αναλώσιμη οντότητα αλλά αξιοποιήσιμη.
3. **Μη ανταγωνιστικότητα:** Στην οικονομική της διάσταση, η φυσική ιδιότητα της μη αναλωσιμότητας καθιστά την πληροφορία – νόημα μη ανταγωνιστική στη χρήση, αφού είναι δυνατό να χρησιμοποιηθεί από πολλούς χωρίς να αναλωθεί ή να απολεσθεί η κατοχή της.
4. **Πολυχρηστικότητα:** Η πληροφορία εμφανίζει χαρακτηριστικά πολυχρηστικότητας, μπορεί δηλαδή να χρησιμοποιηθεί από περισσότερους του ενός χρήστη. Η χρησιμοποίησή της από ένα χρήστη δεν στερεί τη δυνατότητα να χρησιμοποιηθεί και από άλλους χρήστες ή/και από τον ίδιο επανειλημμένα.
5. **Εξατομίκευση:** Η πρόσληψη της πληροφορίας είναι εξατομικευμένη ακόμη και στις περιπτώσεις μαζικής μετάδοσής της.
6. **Πολυεδρικότητα:** Η πληροφορία χαρακτηρίζεται από πολυεδρικότητα (και πολυλειτουργικότητα), δηλαδή περισσότερες από μια «όψεις» (facets) και συνεπώς από την δυνατότητα να τύχει περισσότερων της μιας προσεγγίσεων που καθορίζονται από τις εκάστοτε ανάγκες.
7. **Πολυχρησιμότητα:** Η εξατομικευμένη πρόσληψη έχει ως συνέπεια να προσδίδει στην πληροφορία έντονα χαρακτηριστικά πολυχρησιμότητας, την καθιστά επιδεκτική διαφορετικών εξατομικευμένων τρόπων, μεθόδων ή

σκοπών αξιοποίησης, είτε από διαφορετικούς είτε από τους ίδιους τους χρήστες.

8. **Ευρύτητα επιπτώσεων:** Η εξατομικευμένη πρόσληψη και η πολυχρησιμότητα έχει ως αποτέλεσμα, σε αρκετές περιπτώσεις, η αξιοποίησή της να έχει ευρύτερες επιπτώσεις και εφαρμογές από τις προθέσεις ή τις προβλέψεις του δημιουργού της.
9. **Μοναδικότητα:** Το περιεχόμενο κάθε συγκεκριμένης πληροφορίας χαρακτηρίζεται από μοναδικότητα, είναι δηλαδή μια μοναδική οντότητα που δεν μπορεί να υποκατασταθεί πλήρως από μια παρόμοια πληροφορία, έστω και αν νοητικά ή λεκτικά τις κατατάσσουμε στην ίδια κατηγορία πληροφοριών. Μια συγκεκριμένη πληροφορία μπορεί να είναι σχετική ή συμπληρωματική μιας άλλης, αλλά όχι πλήρες υποκατάστατο, όπως συμβαίνει με τα υλικά αγαθά.
10. **Αναδυτικότητα:** Η συγκέντρωση πληροφοριών εμφανίζει μια αναδυτική ιδιότητα με τη συστημική έννοια του όρου. Η νοητική διασύνδεση και σύνθεση περισσότερων πληροφοριών μπορεί να μην αποτελεί απλά ένα απόθεμα πληροφοριών ή μόνον το «νοηματικό άθροισμά» τους, αλλά μπορεί να δημιουργεί νέες πληροφορίες.
11. **Κόστος αναπαραγωγής:** Η αναπαραγωγή (αναπαράσταση) της πληροφορίας έχει χαμηλό ή μηδενικό κόστος. Η συνεχής αναπαράστασή της ή η αναπαραγωγή της δεν αντιμετωπίζει σπανιότητα πόρων. Ειδικά όσον αφορά την ψηφιακή πληροφορία μπορεί να αναπαράγεται επ' άπειρο σε αντίθεση με τα υλικά αγαθά, η συνεχής παραγωγή των οποίων πιθανόν να αντιμετωπίσει σπανιότητα πρώτων υλών⁴.

Η πληροφορία αποτελεί απαραίτητο συστατικό κάθε φυσικού συστήματος. Στην επιστήμη της Πληροφορικής όλες αυτές οι πληροφορίες χαρακτηρίζονται ως **δεδομένα (data)**. Το δεδομένο (στα αγγλ. data, από τον πληθ. του λατιν. datum) είναι μια αντικειμενική πληροφορία του πραγματικού ή του φανταστικού κόσμου η οποία μπορεί να επαληθευθεί με παρατήρηση (συμπεριλαμβανομένης της μέτρησης). Επίσης μπορεί να θεωρηθεί ως το σύμβολο μιας συγκεκριμένης έννοιας, ενός αντικειμένου, ενός στοιχειώδους μηνύματος ή γεγονότος, κωδικοποιημένου ή όχι, το οποίο αποτελεί το ακατέργαστο υλικό που βρίσκεται σε μια τυποποιημένη μορφή,

⁴ Καρακαπιλίδης Νικόλαος, Πληροφοριακά Συστήματα Διοίκησης, Πανεπιστημιακές παραδόσεις

που η κατάλληλη επεξεργασία του από τον άνθρωπο ή από αυτόματα μέσα, βοηθάει στη λήψη σωστών αποφάσεων.

Τα δεδομένα, λοιπόν, μπορούν να οριστούν ως **τρόποι αναπαράστασης εννοιών και γεγονότων που μπορούν να υποστούν διαχείριση και επεξεργασία από τον άνθρωπο ή από ένα αυτοματοποιημένο υπολογιστικό σύστημα**. Αυτοί οι τρόποι αναπαράστασης συσχετίζονται άμεσα με τη φύση της πληροφορίας που μεταφέρονται από τα δεδομένα.

Στην επιστήμη της Πληροφορικής, τα δεδομένα αναπαρίστανται και μοντελοποιούνται με εντελώς συγκεκριμένο τρόπο. Το αριθμητικό σύστημα που χρησιμοποιούν οι υπολογιστές είναι το **δυναδικό αριθμητικό σύστημα**. Στο σύστημα αυτό υπάρχουν μόνο δύο αριθμοί, το 0 και το 1. Ο κάθε ένας από αυτούς τους αριθμούς αποτελεί τη στοιχειώδη ποσότητα πληροφορίας που μπορεί να αποθηκευτεί στον υπολογιστή και ονομάζεται **bit**. Η αμέσως μεγαλύτερη μονάδα πληροφορίας είναι το **byte**, το οποίο είναι ίσο με 8 bits.

Πώς όλα αυτά συσχετίζονται με την αναπαράσταση των δεδομένων στον υπολογιστή;

Με το να ορίσουμε ειδικούς τύπους δεδομένων, οι οποίοι περιγράφουν τα δεδομένα με μοναδικό τρόπο. Δηλαδή: αν είναι αριθμός θα πρέπει να χρησιμοποιήσουμε ένα τύπο δεδομένων που περιγράφει τους αριθμούς. Ένας αριθμός μπορεί να είναι **ακέραιος, δεκαδικός ή μιγαδικός**, οπότε θα πρέπει να ορίσουμε αντίστοιχους τύπους δεδομένων. Το όνομα ενός φίλου είναι μια **λέξη** που ορίζεται με αντίστοιχο τύπο δεδομένων.

Διαπιστώνουμε ότι **υπάρχει μια αντιστοιχία ανάμεσα στα δεδομένα του πραγματικού κόσμου και στην αναπαράσταση στον ηλεκτρονικό υπολογιστή**.

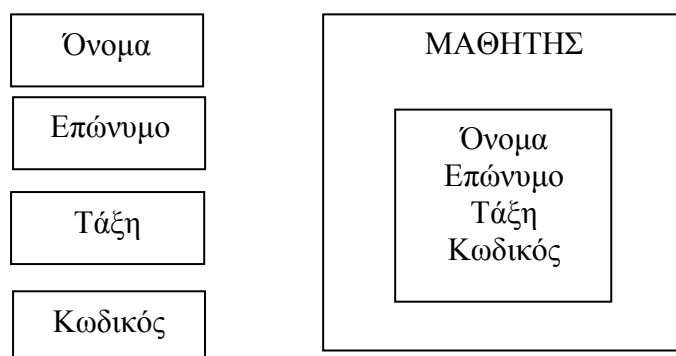
Οι πιο συνηθισμένοι τύποι δεδομένων, που χρησιμοποιούνται στα πληροφορικά συστήματα είναι:

Ονομασία	Περιγραφή	Μέγεθος σε bytes
Integer	Ακέραιοι αριθμοί	2
Real	Πραγματικοί αριθμοί	4
Char	Χαρακτήρες	1
String	Λέξεις και φράσεις	Μεταβλητό
Date	Ημερομηνία	8

Πίνακας 1: Τύποι δεδομένων που χρησιμοποιούνται στους Η/Υ

Σύνθετοι τύποι δεδομένων: Τα όσα έχουν αναφερθεί μέχρι τώρα αφορούν **στοιχειώδεις μορφές δεδομένων**. Τα δεδομένα αυτά συσχετίζονται με μια και μοναδική τιμή και δεν μπορούν να αναλυθούν σε μικρότερες μονάδες με κανένα τρόπο. Τα δεδομένα υφίστανται ευκολότερη και αποτελεσματικότερη διαχείριση αν ομαδοποιηθούν σε μεγαλύτερες οντότητες, τις οποίες διαπραγματευόμαστε όπως ακριβώς και τα απλά δεδομένα.

Παράδειγμα: Η περίπτωση ενός σχολείου. Αν θέλουμε να δημιουργήσουμε ένα πρόγραμμα στο οποίο θα καταχωρούμε τα στοιχεία των μαθητών, τότε τα στοιχεία εκείνα που θα κρατάμε για τον κάθε μαθητή θα είναι το **όνομά** του, το **επώνυμό** του, η **τάξη** του και ο **κωδικός** του. Στην περίπτωση αυτή έχουμε δύο επιλογές: ή να δημιουργήσουμε τέσσερα διαφορετικά πεδία στα οποία καταχωρούμε τα στοιχεία του μαθητή ή ένα σύνθετο τύπο δεδομένων (με τίτλο ΜΑΘΗΤΗΣ) που θα περιέχει τα τέσσερα πεδία:



Παράδειγμα οντότητας με απλούς και σύνθετους τύπους δεδομένων

Σχήμα 1: Περιγραφή οντότητας με απλούς και σύνθετους τύπους δεδομένων

Η δεύτερη προσέγγιση είναι πιο αποτελεσματική σε σχέση με την πρώτη, γιατί τα τέσσερα δεδομένα που σχηματίζουν τον σύνθετο τύπο δεδομένων ΜΑΘΗΤΗΣ υφίστανται διαχείριση σαν να ήταν ένας απλός τύπος δεδομένων. Αυτό είναι πολύ σημαντικό, γιατί μας επιτρέπει να ανακτούμε πληροφορίες από την εφαρμογή με απλό και εύκολο τρόπο.

Στις συμβατικές γλώσσες προγραμματισμού, αυτοί οι σύνθετοι τύποι δεδομένων ονομάζονται **δομές δεδομένων (data structures)**. Στην τεχνολογία των βάσεων δεδομένων ονομάζονται **εγγραφές (records)**.

3. ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

3.1.Βασικές αρχές σχεδίασης βάσης δεδομένων

Μια σωστά σχεδιασμένη βάση δεδομένων⁵ παρέχει πρόσβαση σε ενημερωμένες, ακριβείς πληροφορίες. Επειδή η σωστή σχεδίαση είναι ουσιαστικής σημασίας για την επίτευξη των στόχων που έχουν καθοριστεί, η επένδυση του χρόνου που απαιτείται για την εκμάθηση των αρχών της καλής σχεδίασης είναι σημαντική.

Υπάρχουν ορισμένες αρχές που καθοδηγούν τη διαδικασία της σχεδίασης βάσης δεδομένων. Η **πρώτη αρχή** είναι ότι οι διπλότυπες πληροφορίες (που επίσης ονομάζονται πλεονάζοντα δεδομένα) είναι κακές, διότι σπαταλούν χώρο και αυξάνουν την πιθανότητα σφαλμάτων και ασυνεπειών. Η **δεύτερη αρχή** είναι ότι η ορθότητα και η πληρότητα των πληροφοριών είναι σημαντικές. Εάν η βάση δεδομένων περιέχει λανθασμένες πληροφορίες, οι εκθέσεις που αντλούν πληροφορίες από τη βάση δεδομένων επίσης θα περιέχουν λανθασμένες πληροφορίες. Ως αποτέλεσμα, οι όποιες αποφάσεις βασίζονται σε αυτές τις εκθέσεις θα είναι βασισμένες σε παραπληροφόρηση.

Επομένως, μια καλή σχεδίαση βάσης δεδομένων είναι αυτή που:

- Χωρίζει τις πληροφορίες σε πίνακες που βασίζονται σε θέματα, για να ελαττωθούν τα πλεονάζοντα δεδομένα.
- Παρέχει στην Βάση τις πληροφορίες που απαιτούνται για να συνδεθούν οι πληροφορίες στους πίνακες, σύμφωνα με τις ανάγκες.
- Βοηθά την υποστήριξη και εξασφαλίζει την ακρίβεια και την ακεραιότητα των πληροφοριών.
- Εξυπηρετεί την επεξεργασία των δεδομένων και των αναγκών έκθεσης.

Καθορισμός του σκοπού που θα εξυπηρετεί η βάση δεδομένων: Αυτό το βήμα είναι σημαντικό γιατί βοηθά στην προετοιμασία για τα υπόλοιπα βήματα. Καταγράφουμε σε χαρτί το σκοπό που θα εξυπηρετεί η βάση δεδομένων, τον τρόπο χρήσης, καθώς και το πρόσωπο που θα τη χρησιμοποιεί. Για μια μικρή βάση

⁵ <http://office.microsoft.com/el-gr/access/HA100644501032.aspx>

δεδομένων γράφουμε κάτι απλό, όπως "Η βάση δεδομένων διατηρεί μια λίστα με πληροφορίες μαθητών". Εάν η βάση δεδομένων είναι πιο σύνθετη ή χρησιμοποιείται από πολλά άτομα, η καταγραφή του σκοπού της θα μπορούσε να περιλαμβάνει το πότε και το πώς θα χρησιμοποιεί ο καθένας τη βάση δεδομένων. Ο στόχος είναι να υπάρχει μια καλά ανεπτυγμένη δήλωση αποστολής, στην οποία να γίνεται αναφορά σε όλη τη διάρκεια της διαδικασίας της σχεδίασης. Μια τέτοια δήλωση σας βοηθά να εστιαστείτε στους στόχους σας κατά τη λήψη αποφάσεων.

Εύρεση και οργάνωση των απαιτούμενων πληροφοριών: Συγκέντρωση όλων των τύπων πληροφοριών που θα εγγραφούν στη βάση δεδομένων. Εύρεση και οργάνωση των απαιτούμενων πληροφοριών, από τις ήδη υπάρχουσες. Για παράδειγμα, διατήρηση πληροφοριών των μαθητών σε έγγραφες φόρμες, μέσα σε ένα ντοσιέ αρχείων. Συγκέντρωση αυτών των εγγράφων και καταγραφή κάθε τύπου πληροφοριών που εμφανίζεται (π.χ. κάθε πλαίσιο που συμπληρώνεται σε μια φόρμα). Εάν δεν υπάρχουν διαθέσιμες φόρμες, γίνεται σχεδίαση μιας φόρμας για καταγραφή των πληροφοριών. Ποιες πληροφορίες θα περιλαμβάνει η φόρμα; Ποια πλαίσια προς συμπλήρωση θα δημιουργηθούν; Έπειτα, γίνεται η καταγραφή του τύπου των εκθέσεων που θα παράγει η βάση δεδομένων.

Ένα σημείο-κλειδί είναι ότι πρέπει να διασπαστεί κάθε πληροφορία στα μικρότερα χρήσιμα μέρη της. Στην περίπτωση ενός ονοματεπώνυμου, για να γίνει το επώνυμο άμεσα διαθέσιμο, πρέπει να διασπαστεί σε δύο μέρη — Όνομα και Επώνυμο. Για να γίνει ταξινόμηση μιας έκθεσης κατά επώνυμο, για παράδειγμα, είναι χρήσιμο να αποθηκευτεί το επώνυμο του μαθητή χωριστά. Γενικά, για την ταξινόμηση, την αναζήτηση, τον υπολογισμό και την έκθεση βάσει ενός στοιχείου μιας πληροφορίας, πρέπει να τοποθετείται αυτό το στοιχείο σε χωριστό πεδίο.

Χωρισμός των πληροφοριών σε πίνακες: Χωρισμός των στοιχείων πληροφοριών σε κύριες οντότητες ή θέματα, όπως "Μαθητές" ή "Τάξεις". Κάθε θέμα γίνεται τότε πίνακας. Για τον χωρισμό των πληροφοριών σε πίνακες, γίνεται επιλογή των κυρίων οντοτήτων. Οι οντότητες μπορεί να είναι, για παράδειγμα οι μαθητές, οι καθηγητές, οι τάξεις. Επομένως, είναι λογικό να δημιουργηθούν τρεις πίνακες: έναν για τα δεδομένα σχετικά με τους μαθητές, έναν για τα δεδομένα σχετικά με τους καθηγητές, και έναν για τα δεδομένα σχετικά με τις τάξεις.

Μετατροπή των στοιχείων πληροφοριών σε στήλες: Αποφασίζουμε για το ποιες πληροφορίες θέλετε θα αποθηκεύσουμε σε κάθε πίνακα. Κάθε στοιχείο μετατρέπεται σε πεδίο και εμφανίζεται στον πίνακα ως στήλη. Για παράδειγμα, ο

πίνακας "Μαθητές" μπορεί να περιλαμβάνει πεδία όπως "Επώνυμο" και "Ημερομηνία γέννησης".

Καθορισμός πρωτεύοντων κλειδιών: Κάθε πίνακας πρέπει να περιλαμβάνει μια στήλη ή σύνολο στηλών, που να προσδιορίζει μοναδικά κάθε γραμμή που είναι αποθηκευμένη στον πίνακα. Αυτός είναι συχνά ένας μοναδικός αριθμός αναγνώρισης. Στην ορολογία βάσης δεδομένων, αυτή η πληροφορία ονομάζεται πρωτεύον κλειδί του πίνακα. Το πρωτεύον κλειδί είναι μια στήλη που χρησιμοποιείται για τον μοναδικό προσδιορισμό κάθε γραμμής. Για παράδειγμα, "Κωδικός μαθητή" ή "Κωδικός καθηγητή".

Δημιουργία σχέσεων πινάκων: Εξέταση κάθε πίνακα και απόφαση με ποιον τρόπο τα δεδομένα ενός πίνακα σχετίζονται με τα δεδομένα των άλλων πινάκων. Προσθέτουμε πεδία σε πίνακες ή δημιουργούμε νέους πινάκων.

Βελτίωση της σχεδίασης: Ανάλυση της σχεδίασης για σφάλματα. Δημιουργία πινάκων και πρόσθεση μερικών εγγραφών ως δείγμα δεδομένων. Εξετάζουμε εάν η βάση επιστρέφει τα επιθυμητά αποτελέσματα από τους πίνακες, απαραίτητες προσαρμογές στη σχεδίαση.

Εφαρμογή των κανόνων κανονικοποίησης: Εφαρμόζουμε τους κανόνες κανονικοποίησης δεδομένων, και εξετάζουμε αν οι πίνακες είναι σωστά δομημένοι. Κάνουμε τις απαραίτητες προσαρμογές στους πίνακες.

Ολοκληρώνοντας την παρούσα ενότητα θα δούμε με ποιο τρόπο μοντελοποιείται ένα σύστημα, χρησιμοποιώντας σύνθετους τύπους δεδομένων. Θα στηριχτούμε ξανά στο παράδειγμα του σχολείου. Ένα σχολείο είναι γενικά ένα πολύπλοκο σύστημα. Περιλαμβάνει καθηγητές και μαθητές. Κάθε καθηγητής έχει μια ειδικότητα και διδάσκει σε μια ή περισσότερες τάξεις του σχολείου. Κάθε μαθητής ανήκει σε μια τάξη και διδάσκεται έναν αριθμό μαθημάτων. Αν περιγράψουμε σε πλήρη έκταση τη λειτουργία του σχολείου θα βρεθούμε αντιμέτωποι με ένα πολύπλοκο σύστημα το οποίο πρέπει να μηχανογραφήσουμε.

Αρχικά ταυτοποιούμε τις οντότητες που συμμετέχουν στο σύστημα, οι οποίες είναι: οι καθηγητές, μαθητές, τάξεις και μαθήματα. Αφού ταυτοποιήσουμε τις οντότητες πρέπει να δούμε ποιες πληροφορίες θα πρέπει να κρατάμε για κάθε οντότητα.

Ταυτοποιώντας και αυτές τις πληροφορίες μπορούμε να δημιουργήσουμε ένα σύνθετο τύπο δεδομένων για κάθε μια από αυτές τις οντότητες:



Παράδειγμα σύνθετου τύπου δεδομένων για τις οντότητες ενός σχολείου

Σχήμα 2: Σύνθετοι τύποι δεδομένων για τις οντότητες ενός σχολείου

Μετά τη δημιουργία αυτών των τύπων δεδομένων αρχίζει η διαδικασία καταχώρισης των στοιχείων που σχετίζονται με τη λειτουργία του σχολείου. Το σύνολο των πληροφοριών που καταχωρούνται στο σύστημα αποτελούν μια βάση δεδομένων.

Η βάση δεδομένων ορίζεται ως ένα σύνολο δεδομένων που χαρακτηρίζονται από κάποια λογική οργάνωση και ομαδοποίηση έτσι ώστε να είναι εύκολη και αποτελεσματική η διαχείρισή τους.

Ορισμός Βάσης Δεδομένων

3.2.Χαρακτηριστικά της Βάσης Δεδομένων

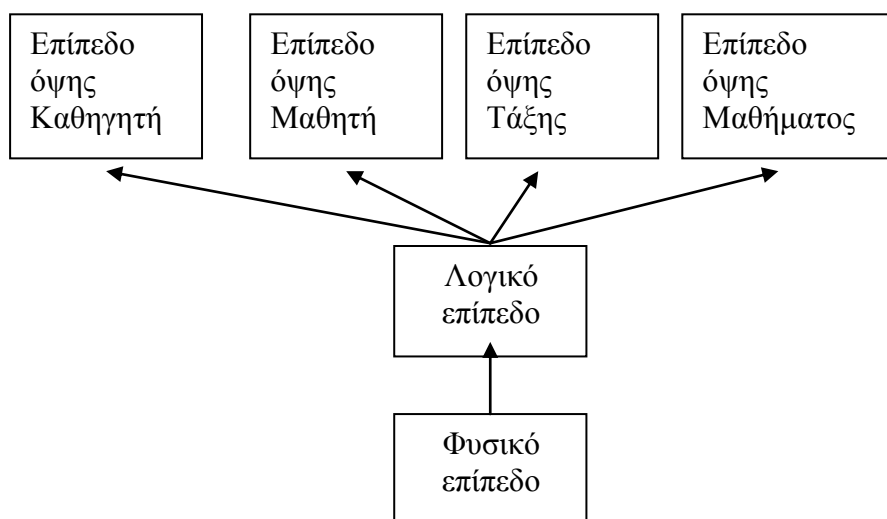
Μια βάση δεδομένων για να θεωρηθεί σωστά σχεδιασμένη και να μπορεί να λειτουργήσει με αποτελεσματικό τρόπο θα πρέπει να διαθέτει ένα πλήθος χαρακτηριστικών εκ των οποίων τα σημαντικότερα είναι:

1. Δεν θα πρέπει να περιλαμβάνει περιττά δεδομένα (redundant data).
2. Θα πρέπει να είναι σχεδιασμένη έτσι ώστε τα δεδομένα που περιλαμβάνει να ανακτώνται εύκολα και γρήγορα.
3. Θα πρέπει να χαρακτηρίζεται από σύστημα ασφαλείας (security system) που να απαγορεύει την πρόσβαση στα δεδομένα μη εξουσιοδοτημένων ατόμων.
4. Θα πρέπει να μπορεί να διαχειρίζεται ταυτόχρονες προσπελάσεις στα ίδια δεδομένα (concurrency control).
5. Θα πρέπει να διαθέτει σύστημα δημιουργίας αντιγράφων ασφαλείας των δεδομένων που είναι καταχωρημένα σε αυτή (backups).

3.3.Συστήματα Διαχείρισης Βάσεων Δεδομένων

Η διαχείριση μιας βάσης δεδομένων πραγματοποιείται με τη βοήθεια ειδικών προγραμμάτων, που ονομάζονται **συστήματα διαχείρισης βάσεων δεδομένων (Database Management Systems, DBMS)**⁶. Ο ρόλος αυτών των συστημάτων είναι η αποτελεσματική επεξεργασία των δεδομένων της βάσης. Μέσω των συστημάτων διαχείρισης μπορούμε να υλοποιήσουμε και να διαχειριστούμε οποιαδήποτε βάση. Διακρίνουμε τρία στάδια:

1. **Σχεδίαση και υλοποίηση της βάσης.** Στο στάδιο αυτό γίνεται ανάλυση των απαιτήσεων που έχουμε από τη βάση δεδομένων. Ακολουθεί ο σχεδιασμός και η υλοποίησή της με τα κατάλληλα εργαλεία ανάπτυξης.



Παράδειγμα
επιπέδων
σχεδιασμού
Βάσεων
Δεδομένων

Σχήμα 3: Επίπεδα σχεδιασμού Βάσεων δεδομένων

2. **Καταχώρηση των δεδομένων στη βάση του συστήματος.** Ακολουθεί η καταχώριση των δεδομένων του συστήματος. Η διαδικασία της καταχώρισης γίνεται με τη βοήθεια κατάλληλα σχεδιασμένων φορμών.
3. **Διαχείριση των δεδομένων της βάσης.** Περιλαμβάνει μια σειρά από δραστηριότητες, όπως η ανάκτηση δεδομένων, δημιουργία αντιγράφων ασφαλείας και τη σωστή συντήρηση της βάσης, εφόσον η δυνατότητα αυτή παρέχεται από το σύστημα διαχείρισης βάσεων δεδομένων.

⁶ <http://dide.flo.sch.gr/Plinet/Tutorials/Tutorials-DatabasesTheory.html>

4. ΜΟΝΤΕΛΑ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

4.1. Μοντέλο οντοτήτων συσχετίσεων⁷

Το μοντέλο οντοτήτων συσχετίσεων (Entity Relationship model, ER model) είναι μια διαγραμματική αναπαράσταση της δομής της βάσης δεδομένων και χρησιμοποιείται κατά το στάδιο του λογικού ή εννοιολογικού σχεδιασμού (conceptual design) της βάσης. Αυτό σημαίνει πως η κεντρική λειτουργία του συνίσταται στην ταυτοποίηση των δεδομένων της εφαρμογής και του τρόπου με τον οποίο αυτά συσχετίζονται μεταξύ τους.

4.1.1. Οι οντότητες και τα χαρακτηριστικά τους

Το πιο χαρακτηριστικό αντικείμενο που εμφανίζεται σε όλα τα μοντέλα οντοτήτων συσχετίσεων είναι η **οντότητα (entity)**. **Μια οντότητα μπορεί να οριστεί ως μια αυτόνομη και αυθύπαρκτη μονάδα του φυσικού κόσμου, η οποία μπορεί να θεωρηθεί τόσο ως πομπός όσο και ως αποδέκτης πληροφορίας.** Μια οντότητα μπορεί να είναι **έμψυχη** (π.χ. καθηγητής) ή **άψυχη** (π.χ. μάθημα, τάξη) αλλά μπορεί να είναι και κάτι πιο γενικό (π.χ. σχολική μονάδα).

Η κάθε οντότητα περιγράφεται από ένα σύνολο ιδιοτήτων οι οποίες στην ορολογία του μοντέλου ονομάζονται **χαρακτηριστικά** ή **πεδία**. Στις πιο πολλές περιπτώσεις, τα πεδία που συσχετίζονται με μια οντότητα παίρνουν απλές τιμές, οι οποίες δεν μπορούν να χωριστούν σε μικρότερες μονάδες. Τέτοιου είδους πεδία ονομάζονται **απλά πεδία (simple attributes)**. Σε άλλες περιπτώσεις ένα πεδίο είναι δυνατό να διασπαστεί σε ένα πλήθος απλών πεδίων. Στην περίπτωση αυτή το πεδίο λέγεται **σύνθετο (complex attribute)**. Αν ένα από τα πεδία κάποιας οντότητας είναι σύνθετο, τότε η τιμή που αποδίδεται σε αυτό προκύπτει από τη συνένωση (concatenation) των τιμών των απλών πεδίων που περιέχει.

Ένας άλλος διαχωρισμός που μπορεί να εφαρμοστεί στα πεδία μιας οντότητας έχει να κάνει με το πλήθος των τιμών που καταχωρούνται σ' αυτά. Στην περίπτωση κατά την οποία ένα πεδίο μπορεί να πάρει περισσότερες από μια τιμές, μπορούμε να θέσουμε περιορισμούς σχετικά με το πλήθος των τιμών που θα καταχωρηθούν σε αυτό.

⁷ Μάργαρης Αθανάσιος, Θεωρία και εφαρμογές στις Βάσεις Δεδομένων, Πανεπιστημιακές παραδόσεις

Σε ορισμένες περιπτώσεις δύο ή περισσότερα πεδία μπορεί να **συσχετίζονται** μεταξύ τους (related attributes), κάτι που επιτρέπει τον υπολογισμό της τιμής του ενός, όταν γνωρίζουμε την τιμή του άλλου.

4.1.2. Τύποι οντοτήτων

Ένα βασικό χαρακτηριστικό που εμφανίζεται σε όλες σχεδόν τις βάσεις δεδομένων, είναι η **ομαδοποίηση των οντοτήτων που περιέχουν με κριτήριο τα κοινά τους χαρακτηριστικά. Το σύνολο των οντοτήτων οι οποίες περιγράφονται από τα ίδια χαρακτηριστικά, ορίζει έναν τύπο οντότητας.** Μπορούμε να ορίσουμε έναν τύπο οντότητας τους καθηγητές, ένα τύπο οντότητας τους μαθητές, ένα τύπο οντότητας τις τάξεις, έναν τύπο οντότητας τα μαθήματα. **Κάθε ένας από αυτούς τους τύπους, περιγράφεται από ένα όνομα και από το σύνολο των πεδίων που περιέχει.** Το σύνολο αυτής της πληροφορίας ορίζει το **σχήμα (schema)** για το συγκεκριμένο τύπο οντότητας. Αντίθετα οι μαθητές που καταχωρούνται στη βάση δεδομένων και για τους οποίους τα πεδία τύπου οντότητας παίρνουν εντελώς συγκεκριμένες τιμές, αποτελούν **στιγμιότυπα (instances)** αυτού του σχήματος.

4.1.3. Πεδίο - κλειδί (key attribute)

Για κάθε τύπο οντότητας μπορεί να υπάρχουν καταχωρημένα στη βάση δεδομένων δεκάδες στιγμιότυπα για αυτή την οντότητα. Επειδή όλα τα στιγμιότυπα χαρακτηρίζονται από την ίδια δομή και οργάνωση της πληροφορίας που περιέχουν πρέπει να διακριθούν μεταξύ τους. Αυτό γίνεται με τη χρήση ενός από τα πεδία, το οποίο λέγεται **πεδίο - κλειδί (key attribute)** το οποίο και περιέχει μοναδική τιμή για κάθε στιγμιότυπο. Το πεδίο που θα χρησιμοποιηθεί ως κλειδί πρέπει να ικανοποιεί τη **συνθήκη μοναδικότητας.** Σε ορισμένες περιπτώσεις το πεδίο κλειδί ενός τύπου οντότητας μπορεί να είναι σύνθετο, να αποτελείται από πολλά πεδία. Στην περίπτωση αυτή η συνθήκη της μοναδικότητας για την τιμή του πρωτεύοντος κλειδιού εφαρμόζεται στο συνδυασμό αυτών των πεδίων.

4.1.4. Η έννοια της συσχέτισης

Δεύτερο χαρακτηριστικό στη δομή ενός μοντέλου οντοτήτων συσχετίσεων, είναι η **συσχέτιση (relationship)**, η οποία καθορίζει με ποιο τρόπο **οι τύποι οντότητας μιας βάσης δεδομένων συσχετίζονται μεταξύ τους**.

Κάθε τύπος συσχέτισης που συμμετέχει σε κάποιο τύπο συσχέτισης παίζει και ένα συγκεκριμένο ρόλο σε αυτή τη συσχέτιση. Το όνομα αυτού του ρόλου (role name) καθορίζει σε πολύ μεγάλο βαθμό το πως οι διάφοροι τύποι οντότητας αλληλεπιδρούν μεταξύ τους.

4.1.5. Χαρακτηριστικά τύπων συσχέτισης

Κάθε τύπος συσχέτισης που ορίζεται ανάμεσα σε δύο τύπους οντότητας χαρακτηρίζεται από δύο ιδιότητες που καθορίζουν το πλήθος και το είδος της συμμετοχής των δύο οντοτήτων στα στιγμιότυπα αυτού του τύπου συσχέτισης.

1. Η **πολλαπλότητα (cardinality)** ενός τύπου συσχέτισης, καθορίζει το πλήθος των στιγμιότυπων αυτού του τύπου συσχέτισης, στον οποίο η οντότητα μπορεί να συμμετέχει.
2. Η δεύτερη σημαντική ιδιότητα έχει να κάνει με το αν η συμμετοχή των δύο τύπων οντότητας σε αυτή τη συσχέτιση είναι **ολική (total participation)** ή **μερική (partial participation)**.

4.2. Σχεσιακό μοντέλο

Το **σχεσιακό μοντέλο αναπαράστασης** των δεδομένων μιας εφαρμογής (**relational data model**) καθιερώθηκε το 1970 από τον Codd και αποτέλεσε ένα από τα πιο απλά και ευέλικτα μοντέλα αυτού του είδους⁸.

Σε αυτό το μοντέλο, τα δεδομένα μιας εφαρμογής αναπαρίστανται ως ένα σύνολο από **σχέσεις (relations)** οι οποίες μπορεί να είναι **πίνακες** ή **αρχεία**. Στις πιο πολλές περιπτώσεις υιοθετείται η χρήση **πινάκων (tables)** που περιέχουν ένα πλήθος **γραμμών (rows)** και **στηλών (columns)**. Η κάθε μια από αυτές τις γραμμές, οι οποίες στην ορολογία του μοντέλου ονομάζονται **πλειάδες (tuples)**, περιέχει ένα

⁸ Βεσκούκης Βασίλειος, Βάσεις Δεδομένων, Πανεπιστημιακές παραδόσεις

σύνολο **απλών πεδίων (attributes)** τα οποία συσχετίζονται μεταξύ τους. Επειδή οι πίνακες χρησιμοποιούνται για την αναπαράσταση των τύπων οντοτήτων καθώς και των τύπων συσχετίσεων μπορεί να θεωρηθεί κάθε μια από τις γραμμές ενός πίνακα σαν ένα **στιγμιότυπο οντότητας ή συσχέτισης** ανάλογα με το αντικείμενο στο οποίο αναφέρεται. Ένα παράδειγμα πίνακα, όπως αυτός ορίζεται από το σχεσιακό μοντέλο, παρουσιάζεται παρακάτω:

ΜΑΘΗΤΗΣ								
Κωδ.	Επίθετο	Όνομα	Πατρώνυμο	Τάξη	A Τρίμηνο	B Τρίμηνο	Γ Τρίμηνο	Γραπτά
1	Αλεξίου	Γιώργος	Νικόλαος	α Γυμνασίου	16	18	20	17
2	Αντωνάρας	Χρήστος	Ανδρέας	α Γυμνασίου	12	17	17	14
3	Ανωνιάδου	Κωνσταντίνα	Δημήτριος	α Γυμνασίου	19	20	20	18
4	Αρμπατζόγλου	Βασιλική	Ηρακλής	α Γυμνασίου	15	14	17	16
5	Βάγιου	Όλγα	Στυλιανός	α Γυμνασίου	19	19	18	19
6	Βακρατσά	Νικολέτα	Σάββας	α Γυμνασίου	14	16	17	15
7	Γκόνη	Αικατερίνη	Μάρκος	α Γυμνασίου	18	18	20	16
8	Εσπιαλίδου	Άννα-Μαρία	Ιωάννης	α Γυμνασίου	14	14	16	12

Παράδειγμα
πίνακα
στο
σχεσιακό
μοντέλο

Πίνακας 2: Παράδειγμα πίνακα στο σχεσιακό μοντέλο αναπαράστασης δεδομένων

Στον παραπάνω πίνακα η κάθε **πλειάδα ή εγγραφή (record)** αναφέρεται και σε ένα ξεχωριστό στιγμιότυπο του τύπου οντότητας ΜΑΘΗΤΗΣ. Η κάθε μια από αυτές τις εγγραφές αποτελείται από ένα σύνολο χαρακτηριστικών ή πεδίων τα οποία λαμβάνουν συγκεκριμένες τιμές, διαφορετικές για κάθε εγγραφή. Έτσι ο πίνακας ΜΑΘΗΤΗΣ περιέχει τα πεδία Κωδικός, Επίθετο, Όνομα, Πατρώνυμο, Τάξη, Α΄ Τρίμηνο, Β΄ Τρίμηνο, Γ΄ Τρίμηνο και Γραπτά. Επειδή οι τιμές της κάθε στήλης αναφέρονται ουσιαστικά στο ίδιο πεδίο του πίνακα, είναι προφανές πως θα πρέπει να ανήκουν στον ίδιο τύπο δεδομένων.

4.2.1. Θεμελιώδεις ορισμοί του σχεσιακού μοντέλου

Κάθε **πεδίο ή στήλη** ενός πίνακα δέχεται τιμές οι οποίες ανήκουν σε ένα συγκεκριμένο και καθορισμένο εκ των προτέρων **σύνολο τιμών (domain)**.

Σε ορισμένες περιπτώσεις είναι δυνατό να καθορίσουμε όχι μόνο το εύρος των τιμών για ένα πεδία, αλλά και τη **μορφή που θα έχουν οι τιμές που καταχωρούνται σε αυτό**.

Ορισμός 1: Ως **σχήμα μιας σχέσης (relation schema)** ορίζεται ένα σύνολο **πεδίων** κάθε ένα εκ των οποίων παίρνει τιμές από ένα **σύνολο τιμών (domain)**. Το

πλήθος των πεδίων του συνόλου π.χ. ΜΑΘΗΤΗΣ ονομάζεται **βαθμός (degree)** της σχέσης, ενώ η συμβολοσειρά ΜΑΘΗΤΗΣ είναι το όνομα της σχέσης. Ένα παράδειγμα μιας σχέσης είναι ο πίνακας ΜΑΘΗΤΗΣ (σχήμα 2) που μπορεί να γραφεί με τη μορφή:

ΜΑΘΗΤΗΣ (Κωδικός, Επίθετο, Όνομα, Πατρώνυμο, Τάξη, Α΄ Τρίμηνο, Β΄ Τρίμηνο, Γ΄ Τρίμηνο, Γραπτά).

Ορισμός 2: Ως **σχέση ή στιγμιότυπο σχέσης (relation instance)**, του **σχήματος σχέσης ΜΑΘΗΤΗΣ**, ορίζεται ένα σύνολο από πλειάδες. Κάθε πλειάδα είναι μια **διατεταγμένη λίστα (ordered list)** από n τιμές, με κάθε τιμή να παίρνει τιμές από το πεδίο του συνόλου των τιμών, ενώ εναλλακτικά μπορεί να πάρει και την τιμή **NULL**.

ΠΕΔΙΑ - ΚΛΕΙΔΙΑ: Ένα από τα βασικά χαρακτηριστικά μιας σχέσης είναι η μοναδικότητα κάθε μιας από τις πλειάδες που περιέχει. Αυτό σημαίνει **πως όλες οι πλειάδες είναι διαφορετικές μεταξύ τους** και σε καμιά περίπτωση δεν πρόκειται να καταχωρήσουμε δύο φορές την ίδια πλειάδα. Επειδή όμως μια πλειάδα είναι ένα σύνολο πεδίων, μπορούμε να θεωρήσουμε πως σε κάθε μια από τις πλειάδες ενός πίνακα, υπάρχει ένας συνδυασμός πεδίων με την ακόλουθη ιδιότητα: **ο συνδυασμός των τιμών αυτών των πεδίων είναι μοναδικός για κάθε εγγραφή**. Το σύνολο αυτών των πεδίων ονομάζεται **superkey** του σχεσιακού σχήματος στο οποίο αναφερόμαστε.

4.2.2. Κανόνες ακεραιότητας (Integrity constraints)

1. **Key constraint:** επιβάλλει την απόδοση **μοναδικών τιμών** σε όλα τα πεδία του πίνακα που έχουν χαρακτηριστεί υποψήφια κλειδιά. Δεν θα πρέπει να υπάρχουν στον ίδιο πίνακα δύο εγγραφές που να έχουν την ίδια τιμή στα υποψήφια κλειδιά τους.
2. **Entity integrity constraint:** απαγορεύει την απόδοση της τιμής **NULL**, στο πρωτεύον κλειδί της εγγραφής του πίνακα. Το πρωτεύον κλειδί μας επιτρέπει να ξεχωρίσουμε τις εγγραφές του πίνακα, και επομένως αν δύο εγγραφές έχουν τιμή **NULL** στο πρωτεύον κλειδί τους, δεν είναι δυνατή η διάκριση ανάμεσά τους.
3. **Referential integrity constraint:** αναφέρεται στις συσχετίσεις που υφίστανται ανάμεσα στους πίνακες. **Αν η τιμή κάποιου πεδίου μιας από τις**

εγγραφές αναφέρεται σε εγγραφή κάποιου άλλου πίνακα, αυτή θα πρέπει να είναι μια υπάρχουσα εγγραφή.

4.2.3. Λογικός σχεδιασμός

Ένα από τα βασικά ερωτήματα που προκύπτουν κατά την πραγματοποίηση αυτής της διαδικασίας (ομαδοποίηση των πεδίων των τύπων οντότητας και δημιουργία πινάκων για τους τύπους οντότητας και συσχέτισης), έχει να κάνει με την επιλογή των πεδίων που θα ομαδοποιηθούν και θα σχηματίσουν ένα πίνακα.

Η διαδικασία της σωστής σχεδίασης της δομής των πινάκων της βάσης, έτσι ώστε να ικανοποιεί τις προδιαγραφές που έχουν τεθεί και ταυτόχρονα να είναι ευέλικτη και αποδοτική στη χρήση της ονομάζεται **λογικός ή εννοιολογικός σχεδιασμός (conceptual or logical design)**.

Βασικές αρχές σχεδίασης της δομής του Σχεσιακού Μοντέλου

1. Οι πίνακες πρέπει να είναι οργανωμένοι έτσι ώστε μια απλή ανάγνωση του σχεσιακού σχήματος να είναι αρκετή για να καταλάβουμε το είδος της πληροφορίας που καταχωρούμε σε κάθε περίπτωση. Αυτό σημαίνει να δημιουργούμε ένα ξεχωριστό πίνακα για κάθε τύπο οντότητας και για κάθε τύπο συσχέτισης και σε καμιά περίπτωση να μη συνδυάζουμε πεδία διαφορετικών τύπων οντότητας στον ίδιο πίνακα, για να αποφύγουμε την εμφάνιση **περιττών δεδομένων (redundant data)**.
2. Η οργάνωση της δομής των πινάκων πρέπει να γίνεται με τέτοιο τρόπο ώστε να μην εμφανίζονται στις εγγραφές των πινάκων επαναλαμβανόμενα πεδία.

4.2.4. Κανονικοποίηση

Η διαδικασία της **κανονικοποίησης (normalization process)** προτάθηκε από τον Codd το 1970, και η βασική της λειτουργία είναι να δέχεται ως είσοδο το σχεσιακό σχήμα μιας βάσης δεδομένων και να εφαρμόζει πάνω του μια σειρά από ελέγχους προκειμένου να διαπιστώσει αν ανήκει ή όχι σε κάποια κανονική μορφή. Ο Codd εισήγαγε τρεις κανονικές μορφές, την πρώτη, την δεύτερη και τρίτη κανονική μορφή οι οποίες σχετίζονται με τις συναρτησιακές εξαρτήσεις που υφίστανται ανάμεσα στα πεδία των πινάκων βάσης, ενώ αργότερα προτάθηκαν

άλλες δύο κανονικές μορφές, η τέταρτη και η πέμπτη, οι οποίες στηρίζονται σε άλλου είδους εξαρτήσεις (multivalued dependencies και join dependencies).

1. Πρώτη κανονική μορφή: απομάκρυνση των επαναλαμβανόμενων πεδίων.
2. Δεύτερη κανονική μορφή: απομάκρυνση όλων των μερικών συσχετίσεων.
3. Τρίτη κανονική μορφή: απομάκρυνση όλων των μεταβατικών συσχετίσεων
4. Τέταρτη κανονική μορφή: απομάκρυνση όλων των συσχετίσεων με πεδία πολλαπλών τιμών.
5. Πέμπτη κανονική μορφή: απομάκρυνση όλων των υπόλοιπων ανωμαλιών.

4.2.5. Απεικόνιση του μοντέλου οντοτήτων συσχετίσεων στο σχεσιακό μοντέλο

Η διαδικασία με την οποία είναι δυνατή η δημιουργία του σχεσιακού μοντέλου μιας βάσης δεδομένων, από το μοντέλο οντοτήτων συσχετίσεων, είναι γνωστή ως **απεικόνιση (mapping)** του ενός μοντέλου στο άλλο και είναι ένα από τα πιο σημαντικά στάδια του λογικού σχεδιασμού μιας βάσης δεδομένων.

5. ΟΡΟΛΟΓΙΑ⁹

A

Ασφάλεια Δεδομένων: Με τον όρο ασφάλεια εννοούμε την ασφάλεια των δεδομένων σε ότι αφορά τη χρήση τους από μη εξουσιοδοτημένα άτομα. Οι βάσεις δεδομένων διαθέτουν ασφάλεια πολλαπλών επιπέδων.

1. σε επίπεδο Βάσης Δεδομένων. Για να έρθει σε επαφή ο χρήστης πρέπει να έχει σχετική εξουσιοδότηση.
2. σε επίπεδο αρχείου. Μια βάση έχει πολλά αρχεία τα οποία δεν πρέπει να τα βλέπουν όλοι οι χρήστες. Για τα αρχεία που υπάρχουν περιορισμοί έχουμε ειδικά κλειδιά (κωδικούς) χωρίς τα οποία είναι αδύνατη η χρήση τους.
3. σε επίπεδο record (εγγραφής). Ένα αρχείο αποτελείται από πολλές εγγραφές. Μερικές από αυτές μπορεί να έχουν πληροφορίες που δε πρέπει να εμφανίζονται σε όλους τους χρήστες. Γι αυτό το λόγο μπορούμε να επιλέξουμε και να ασφαλίσουμε τις συγκεκριμένες εγγραφές.
4. σε επίπεδο πεδίου. Κάθε εγγραφή αποτελείται από πολλά πεδία. Ασφαλίζουμε τα πεδία τα οποία δεν θέλουμε να είναι προσπελάσιμα από όλους τους χρήστες.

B

Βάση δεδομένων: Μία βάση Δεδομένων είναι μια ενοποιημένη συλλογή δεδομένων η οποία είναι φτιαγμένη σε φυσικές σχέσεις δεδομένων έτσι που να δίνει όλους τους απαραίτητους δρόμους προσπέλασης σε κάθε μονάδα δεδομένων με αντικειμενικό σκοπό να ικανοποιεί τις διάφορες ανάγκες όλων των χρηστών.

Γ

Γραμμή (row): βλέπε εγγραφή

Γλώσσα ορισμού δεδομένων (Data Definition Language, DDL): Δομική μονάδα της SQL. Περιλαμβάνει εντολές που μας επιτρέπουν να υλοποιήσουμε πίνακες, σχέσεις ανάμεσα σε πίνακες, και γενικά όλη τη δομή μιας βάσης δεδομένων.

⁹ <http://databases.about.com/od/administration/a/glossary.htm>

Γλώσσα χειρισμού δεδομένων (Data Manipulation Language, DML): Δομική μονάδα της SQL. Επιτρέπει τη διαχείριση των δεδομένων της εφαρμογής, όπως την εισαγωγή, διαγραφή, ανάκτηση και τροποποίηση δεδομένων.

Δ

Δεδομένα: Στην επιστήμη της Πληροφορικής κάθε πληροφορία χαρακτηρίζεται ως δεδομένο (data). Είναι τρόποι αναπαραστάσεις εννοιών και γεγονότων που μπορούν να υποστούν διαχείριση και επεξεργασία από τον άνθρωπο ή ένα αυτοματοποιημένο υπολογιστικό σύστημα.

Διαθέσιμα δεδομένα (operational data): είναι τα δεδομένα μιας ΒΔ τα οποία είναι διαθέσιμα για επεξεργασία.

Διάγραμμα οντοτήτων □ συσχετίσεων (Entity Relationship Diagram, E-R Diagram): Το διάγραμμα οντοτήτων - συσχετίσεων είναι ένα εξειδικευμένο γραφικό που επεξηγεί τις αλληλεξαρτήσεις μεταξύ των οντοτήτων σε μια βάση δεδομένων. Τα διαγράμματα του E-R χρησιμοποιούν ειδικά σύμβολα που αντιπροσωπεύουν τρεις διαφορετικούς τύπους πληροφοριών. Τα παραλληλόγραμμα χρησιμοποιούνται συνήθως για να αντιπροσωπεύσουν τις οντότητες. Οι ρόμβοι χρησιμοποιούνται για να αντιπροσωπεύσουν τις σχέσεις και ελλείψεις χρησιμοποιούνται για να αντιπροσωπεύσουν τις ιδιότητες.

Διαχείριση ακεραιότητας (Integrity): Δομική μονάδα της SQL. Επιτρέπει λεπτομερή έλεγχο των δεδομένων που καταχωρούνται στη βάση, έτσι ώστε να μη παραβιάζονται οι κανόνες ακεραιότητας που έχουμε ορίσει και οι οποίοι όταν τηρούνται, απομακρύνουν τον κίνδυνο καταχώρησης ασυνεπών δεδομένων.

Δομημένη γλώσσα ερωταποκρίσεων (Structured Query Language, SQL): Η SQL αποτελεί σήμερα την πιο διαδεδομένη γλώσσα διαχείρισης σχεσιακών βάσεων δεδομένων. Η SQL παρέχει δυνατότητες για:

1. τον ορισμό, τη διαγραφή και τη μεταβολή πινάκων και κλειδιών,
2. τη σύνταξη ερωτήσεων (queries),
3. την εισαγωγή, διαγραφή και μεταβολή στοιχείων,
4. τον ορισμό όψεων (views) πάνω στα δεδομένα,
5. τον ορισμό δικαιωμάτων πρόσβασης,
6. τον έλεγχο της ακεραιότητας των στοιχείων,
7. τον έλεγχο συναλλαγών (transaction)

Διαδικό αριθμητικό σύστημα: Το αριθμητικό σύστημα που χρησιμοποιούν οι υπολογιστές. Στο σύστημα αυτό υπάρχουν μόνο δύο αριθμοί. Το 0 και το 1. Ο κάθε ένας από αυτούς αποτελεί τη στοιχειώδη ποσότητα πληροφορίας που μπορεί να αποθηκευτεί στον υπολογιστή και ονομάζεται bit. Η αμέσως μεγαλύτερη μονάδα πληροφορίας είναι το byte, ίσο με 8 bits.

Ε

Εγγραφή (Record): Κάθε γραμμή σε έναν πίνακα μιας ΒΔ ονομάζεται εγγραφή και περιέχει όλα τα δεδομένα που περιγράφουν μια συγκεκριμένη καταχώριση του πίνακα.

Επίπεδο όψης: Είναι ο καθορισμός του τμήματος της ΒΔ που θα είναι ορατό σε κάθε χρήστη. Αυτό βρίσκει εφαρμογή σε μεγάλες βάσεις δεδομένων με πολύπλοκη σχεδίαση και μεγάλο αριθμό χρηστών. Ο λόγος για τον οποίο γίνεται ο διαχωρισμός των χρηστών έχει σχέση με ζητήματα ασφαλείας της ΒΔ.

Κ

Κανονικοποίηση (normalization process): Είναι η διαδικασία εκείνη η οποία βοηθάει στη μετατροπή μιας σχέσης με προβλήματα, σε δύο ή περισσότερες σχέσεις απαλλαγμένες αυτών των προβλημάτων.

Κανόνες ακεραιότητας (integrity constraints): Οι κανόνες ακεραιότητας ενός σχεσιακού σχήματος μιας ΒΔ διασφαλίζουν τη συνέπεια των δεδομένων της βάσης και είναι:

1. απόδοση μοναδικών τιμών σε όλα τα πεδία που έχουν χαρακτηριστεί υποψήφια κλειδιά.
2. απαγόρευση της απόδοσης της τιμής NULL στο πρωτεύον κλειδί κάθε εγγραφής του πίνακα.
3. εάν η τιμή κάποιου πεδίου μιας από τις εγγραφές ενός πίνακα αναφέρεται σε εγγραφή κάποιου άλλου πίνακα, αυτή θα πρέπει να είναι μια υπάρχουσα εγγραφή.

Λ

Λειτουργικές μονάδες: Για την αυτοματοποίηση σύνθετων εργασιών χρησιμοποιείται η γλώσσα προγραμματισμού VBA (Visual Basic for Applications).

Με τη γλώσσα προγραμματισμού VBA δημιουργούνται διάφορα «προγράμματα» τα οποία θα εκτελούν συγκεκριμένες εργασίες χρησιμοποιώντας τα περιεχόμενα της ΒΔ. Ο κώδικας αυτών των προγραμμάτων αποθηκεύεται σε στοιχεία της ΒΔ που ονομάζονται λειτουργικές μονάδες.

Λευκές τιμές (NULL values): Μια τιμή ενός πίνακα , μπορεί να είναι συγκεκριμένη ή όχι. Στη δεύτερη περίπτωση λέμε ότι η τιμή είναι ΛΕΥΚΗ (NULL). Όταν εγγράφουμε τιμές σε έναν πίνακα και μια τιμή μας είναι άγνωστη, τότε στη θέση της δεν εγγράφουμε τίποτε. Οι λευκές τιμές προκύπτουν όταν σε μια θέση του πίνακα δεν έχουμε εγγράψει κάποια τιμή, η θέση δηλ. είναι άδεια. Οι τιμές αυτές δεν πρέπει να συγχέονται με τα κενά τα οποία είναι συγκεκριμένες τιμές.

Λογισμικό (Software): Ανάμεσα στο χρήστη και τη Βάση Δεδομένων όπως αυτή είναι τοποθετημένη στον Η/Υ υπάρχει μια σειρά από προγράμματα τα οποία βοηθούν τους χρήστες στη χρήση της Βάσης Δεδομένων. Τα προγράμματα αυτά αποτελούν το Σύστημα Διοίκησης των Βάσεων Δεδομένων ΣΔΒΔ (Data Base Management System DBMS). Οτιδήποτε θέλει να κάνει ο χρήστης με τη βάση δεδομένων γίνεται μέσω του ΣΔΒΔ.

Λογικά χαρακτηριστικά ΒΔ: Κάθε πεδίο δεδομένων παίρνει μια συγκεκριμένη τιμή από ένα χώρο δυνατών τιμών. Η τιμή ενός πεδίου δεδομένων μπορεί να είναι:

1. Αριθμητική
2. Αλφαβητική
3. Αλφαριθμητική

Λογικό επίπεδο σχεδιασμού ΒΔ: Στο επίπεδο λογικού σχεδιασμού καθορίζουμε ποια δεδομένα θα αποθηκευτούν στη βάση του συστήματος και ποιες είναι οι σχέσεις που υφίστανται μεταξύ τους. Η διαφορά ανάμεσα στο φυσικό και λογικό επίπεδο έχει να κάνει με τον τρόπο διαχείρισης των δεδομένων: στο φυσικό επίπεδο καθορίζουμε πως τα δεδομένα θα αποθηκευτούν στη βάση του συστήματος, ενώ στο λογικό επίπεδο καθορίζουμε ποια είναι αυτά τα δεδομένα.

Ξ

Ξένο κλειδί (foreign key): Το ξένο κλειδί είναι ένα πεδίο από σχετικό πίνακα που ταιριάζει με το πεδίο του πρωτεύοντος κλειδιού ενός άλλου σχετικού πίνακα. Μπορεί να χρησιμοποιηθεί ως σημείο αναφοράς του σχετικού πίνακα.

O

Οντότητα (entity): Μια οντότητα μπορεί να οριστεί ως μια αυτόνομη και αυθύπαρκτη μονάδα του φυσικού κόσμου η οποία μπορεί να θεωρηθεί ως πομπός και ως αποδέκτης της πληροφορίας. Μπορεί να είναι έμψυχη (μαθητής), άψυχη (μάθημα), αλλά και πιο γενική (σχολική μονάδα).

Ορισμός όψεων της βάσης (view definition): Δομική μονάδα της SQL. Επιτρέπει τη δημιουργία όψεων της βάσης δεδομένων οι οποίοι ορίζονται ως εικονικοί πίνακες οι οποίοι περιέχουν δεδομένα από έναν ή περισσότερους πίνακες της βάσης.

Ορισμός εξουσιοδοτήσεων (authorization): Δομική μονάδα της SQL. Επιτρέπει τη δημιουργία ομάδων χρηστών και την απόδοση διαφορετικών δικαιωμάτων πρόσβασης σε κάθε έναν από αυτούς, προκειμένου η κάθε ομάδα χρηστών, να διαχειρίζεται μόνο τα δικά της δεδομένα.

Π

Πεδίο □ κλειδί (key attribute): Για κάθε τύπο οντότητας υπάρχουν καταχωρημένα στη ΒΔ δεκάδες ή εκατοντάδες στιγμιότυπα για αυτή την οντότητα. Επειδή όλα τα στιγμιότυπα χαρακτηρίζονται από την ίδια δομή και οργάνωση της πληροφορίας που περιέχουν, θα πρέπει με κάποιο τρόπο να διακριθούν μεταξύ τους. Αυτό γίνεται με τη χρήση ενός εκ των πεδίων που περιέχονται στο σχήμα αυτής της οντότητας που λέγεται πεδίο □ κλειδί, το οποίο περιέχει μοναδική τιμή για κάθε στιγμιότυπο.

Πίνακας (table): Όλα τα δεδομένα που καταχωρούνται σε μια ΒΔ αποθηκεύονται σε έναν ή περισσότερους πίνακες. Ένας πίνακας είναι μια συλλογή δεδομένων που περιγράφουν ομοειδή αντικείμενα.

Πλειάδα (tuples): Κάθε γραμμή του πίνακα της ΒΔ στην ορολογία του Σχεσιακού Μοντέλου ονομάζεται πλειάδα.

Πλεονασματικά ή περιττά δεδομένα (redundant data): με τον όρο πλεονασματικά δεδομένα εννοούμε τις διπλές ή πολλαπλές εγγραφές σε μια ΒΔ. Τα ίδια δεδομένα δεν πρέπει να καταχωρούνται στη ΒΔ διότι ανακύπτουν δύο σοβαρά προβλήματα. 1) άσκοπη σπατάλη αποθηκευτικού χώρου στο σκληρό δίσκο του Η/Υ και 2) κίνδυνος δημιουργίας ασυνεπών δεδομένων (inconsistent data).

Πληροφορία: Είναι δεδομένα που ανακτούμε από μια ΒΔ, τα οποία είναι οργανωμένα με τέτοιο τρόπο ώστε να έχουν νόημα για το άτομο που τα εξετάζει. Σε μια ΒΔ αποθηκεύουμε δεδομένα και ανακτούμε πληροφορίες.

Πολλαπλότητα (cardinality): Καθορίζει το πλήθος των στιγμιότυπων ενός τύπου συσχέτισης στον οποίο μια οντότητα μπορεί να συμμετάσχει.

Πεδίο: Ένας πίνακας χωρίζεται σε γραμμές και στήλες . Κάθε γραμμή αντιστοιχεί σε μια εγγραφή. Κάθε στήλη αντιστοιχεί σε ένα πεδίο. Πεδία επομένως είναι οι στήλες ενός πίνακα ΒΔ στις οποίες καταχωρούνται οι τιμές για ένα συγκεκριμένο χαρακτηριστικό των εγγραφών του πίνακα.

Πρωτεύον κλειδί (primary key): Η επιλογή ενός πεδίου, από το σύνολο των πεδίων που μπορούν να χρησιμοποιηθούν ως υπονήφια κλειδιά, ονομάζεται πρωτεύον κλειδί του σχήματος της σχέσης.

Σ

Στήλη (column): Βλέπε πεδίο

Συσχέτιση (relationship): Είναι η διαδικασία εκείνη η οποία καθορίζει με ποιον τρόπο οι τύποι οντότητας μιας ΒΔ σχετίζονται μεταξύ τους.

Σχήμα (schema): Κάθε τύπος οντότητας περιγράφεται από ένα όνομα και από το σύνολο των πεδίων που περιέχει. Το σύνολο αυτής της πληροφορίας ορίζει το σχήμα για το συγκεκριμένο τύπο πληροφορίας.

Στιγμιότυπα (instances): Τα δεδομένα που καταχωρούνται στη ΒΔ και για τα οποία τα πεδία τύπου οντότητας παίρνουν εντελώς συγκεκριμένες τιμές, αποτελούν στιγμιότυπα αυτού του σχήματος.

Συναρτησιακή εξάρτηση (functional dependency): Ορίζεται ως μια συσχέτιση ανάμεσα στα πεδία ενός πίνακα. Σε έναν οποιοδήποτε πίνακα μιας σχεσιακής βάσης δεδομένων όλα τα πεδία του πίνακα που δεν ανήκουν στο πρωτεύον κλειδί του, εξαρτώνται συναρτησιακά από τα πεδία του πρωτεύοντος κλειδιού.

Συστήματα Βάσεων δεδομένων (μορφές οργάνωσης):

α) Σύστημα τηλεπεξεργασίας (teleprocessing system): Έχει σκοπό την εξυπηρέτηση πολλών χρηστών. Στο σύστημα αυτό έχουμε έναν κεντρικό υπολογιστή στον οποίο βρίσκεται το Σύστημα Διοίκησης Βάσεων Δεδομένων (ΣΔΒΔ) καθώς και η ίδια η βάση (τα δεδομένα δηλαδή). Ο κάθε χρήστης έχει ένα τερματικό άμεσα συνδεδεμένο με τον κεντρικό υπολογιστή. Όλες οι επεξεργασίες των δεδομένων γίνονται από τον κεντρικό Η/Υ. Τοπικά στο τερματικό δεν γίνεται καμιά επεξεργασία, όλα μεταβιβάζονται στο ΣΔΒΔ.

β) **Το σύστημα πελάτης εξυπηρετητής (the client server system):** Στο εν λόγω σύστημα το τερματικό του κάθε χρήστη πελάτη (client) και ο κεντρικός υπολογιστής είναι συνδεδεμένα μεταξύ τους σε ένα τοπικό δίκτυο.

Η ΒΔ καθώς και το ΣΔΒΔ βρίσκονται στον κεντρικό υπολογιστή και είναι κάτω από το απόλυτο έλεγχό του. Όλες οι ενέργειες που πρέπει να γίνουν στη βάση πχ εισαγωγή νέων δεδομένων, αλλαγές δεδομένων κλπ γίνονται αποκλειστικά από τον κεντρικό Η/Υ. Ο υπολογιστής του χρήστη πελάτη χρησιμοποιείται για τοπική επεξεργασία στοιχείων. Ο κεντρικός υπολογιστής εξυπηρετητής δέχεται τις απαιτήσεις του χρήστη □ πελάτη τις επεξεργάζεται και επιστρέφει τα αποτελέσματα.

γ) **Σύστημα κατανομής πόρων (resource sharing system):** Στο σύστημα αυτό γίνεται αξιοποίηση των δυνατοτήτων του δικτύου. Κάθε Η/Υ πρέπει να διαθέτει πρόγραμμα εφαρμογής, Σύστημα Διοίκησης Βάσης Δεδομένων (ΣΔΒΔ) και λειτουργικό σύστημα του δικτύου.

Στην περίπτωση αυτή ο εξυπηρετητής (server) δεν είναι ο διαχειριστής της βάσης, αλλά τα καθήκοντά του περιορίζονται στη διακίνηση αρχείων της βάσης προς τους Η/Υ πελάτες.

δ) **Σύστημα κατανεμημένης ΒΔ (distributed data base system):** Στο εν λόγω σύστημα η ΒΔ είναι κατανεμημένη (γεωγραφικά) σε πολλούς υπολογιστές. Κάθε υπολογιστής επεξεργάζεται τόσο το πρόγραμμα εφαρμογής του πελάτη, όσο και τις αναφορές στη βάση

Σχεσιακή άλγεβρα (relational algebra): Ορίζεται ως ένα σύνολο πράξεων με τις οποίες μπορούμε να διαχειριστούμε τα δεδομένα των πινάκων μιας σχέσης ΒΔ. Αυτή η διαχείριση περιλαμβάνει την επιλογή γραμμών και στηλών οι οποίες πληρούν κάποια συγκεκριμένα κριτήρια, τόσο από ένα απλό πίνακα όσο και από το συνδυασμό δύο ή περισσότερων πινάκων. Εφόσον ένα πίνακας διαθέτει όλες τις ιδιότητες ενός συνόλου, μπορούμε να εφαρμόσουμε πάνω στις ΒΔ όλες τις γνωστές πράξεις από τη θεωρία συνόλων και οι οποίες είναι:

1. η ένωση (union)
2. η τομή (intersection)
3. η διαφορά (difference) και
4. το καρτεσιανό γινόμενο (Cartesian product)

Η άλλη κατηγορία πράξεων που αποτελούν τμήμα της σχεσιακής άλγεβρας περιλαμβάνει διαδικασίες οι οποίες έχουν οριστεί ειδικά για το σχεσιακό μοντέλο ΒΔ και είναι:

1. η επιλογή (select)
2. η προβολή (project)
3. η σύζευξη (join) και
4. η διαίρεση (division)

T

Τύποι δεδομένων: Είναι το είδος των δεδομένων που μπορεί ο χρήστης να καταχωρίσει σε ένα πεδίο. Οι πιο συνηθισμένοι τύποι δεδομένων είναι:

1. Ακέραιοι αριθμοί (integer)
2. Πραγματικοί αριθμοί (real)
3. Χαρακτήρες (char)
4. Λέξεις και φράσεις (string)
5. Ημερομηνία (date)

Υ

Υλικό (hardware): Με τον όρο αυτό εννοούμε τις δευτερεύουσες μνήμες οι οποίες λειτουργούν σαν αποθήκες για τα δεδομένα της ΒΔ και τέτοιες είναι οι σκληροί δίσκοι, οπτικοί δίσκοι κλπ. Ακόμη στο Υλικό μπορούμε να περιλάβουμε και τις αντίστοιχες μονάδες ελέγχου των μονάδων αυτών, τα κανάλια εισόδου/εξόδου κλπ. Γενικά θεωρούμε ότι μια ΒΔ είναι πολλή μεγάλη ώστε να χωράει ολόκληρη στην κεντρική μνήμη του Η/Υ.

Υποψήφια κλειδιά (candidate keys): Σε έναν πίνακα μιας ΒΔ υπάρχουν περισσότερα από ένα πεδία που μπορούν να χρησιμοποιηθούν ως πεδία κλειδιά μιας σχέσης. Τα πεδία αυτά ονομάζονται υποψήφια κλειδιά.

Φ

Φόρμα (Form): Για την διευκόλυνση της διαδικασίας καταχώρησης και επεξεργασίας των δεδομένων μιας βάσης, σχεδιάζεται από τον προγραμματιστή της ΒΔ, ένα πρότυπο στο οποίο καθορίζουμε τα δεδομένα που θέλουμε να καταχωρίσουμε, να αναζητήσουμε ή να επεξεργαστούμε. Μια φόρμα είναι άμεσα συνδεδεμένη με έναν ή περισσότερους πίνακες και εμφανίζει συγκεντρωμένα τα στοιχεία που αφορούν κάθε εγγραφή.

Φυσικό επίπεδο σχεδίασης ΒΔ: είναι ο καθορισμός του τρόπου με τον οποίο τα δεδομένα της βάσης θα αποθηκεύονται στο σύστημα. Στο επίπεδο αυτό καθορίζουμε τους τύπους των δεδομένων της βάσης και τους αλγόριθμους διαχείρισης δεδομένων.

X

Χρήστες: Χωρίζονται σε τρεις κατηγορίες: α) τον προγραμματιστή εφαρμογών, β) τους απλούς χρήστες και γ) τον υπεύθυνο της Β.Δ.

6. ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΣΤΙΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Στην ενότητα αυτή γίνεται μια σύντομη παρουσίαση των Μοντέλων Βάσεων Δεδομένων που βρίσκονται συχνότερα σε χρήση.

6.1. Ιεραρχικό μοντέλο δεδομένων¹⁰

Γενικά: Το ιεραρχικό μοντέλο (hierarchical) έχει μια ιεραρχική δομή που θυμίζει δένδρο. Οι οντότητες μοιάζουν με απολήξεις από κλαδιά δένδρων και τοποθετούνται σε επίπεδα ιεραρχίας. Τα κλαδιά παριστάνουν τις συσχετίσεις ανάμεσα στις οντότητες.

Από μια οντότητα που βρίσκεται σ' ένα ανώτερο επίπεδο εκκινούν πολλά κλαδιά, καθένα από τα οποία καταλήγει σε μια οντότητα που βρίσκεται σ' ένα χαμηλότερο επίπεδο. Αλλά, σε κάθε οντότητα που βρίσκεται σ' ένα χαμηλότερο επίπεδο αντιστοιχεί μία και μόνο οντότητα που βρίσκεται σ' ένα ανώτερο επίπεδο. Το μοντέλο αυτό ήταν το πρώτο που εμφανίστηκε αλλά σήμερα θεωρείται δύσχρηστο και ξεπερασμένο.

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. Η οργάνωση των δεδομένων γίνεται με βάση τη σχέση «πατέρας – παιδί».

Δείκτες (pointers): Σε ένα ιεραρχικό μοντέλο 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 σχετικών τμημάτων δεδομένων (των «παιδιών»).

Ορισμοί:

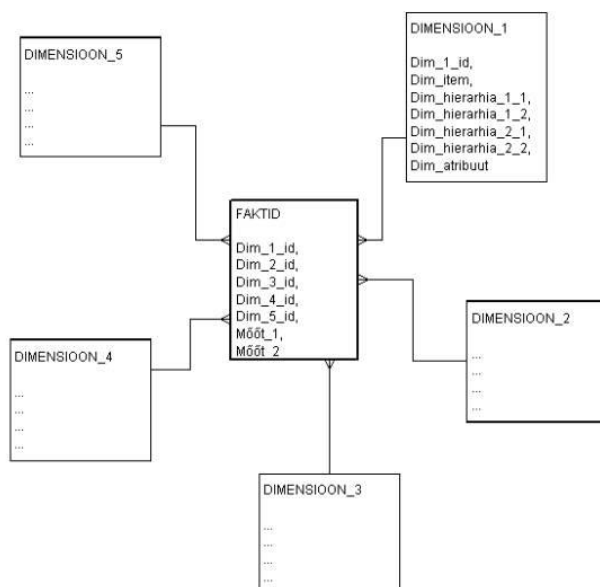
Εγγραφή: είναι μια συλλογή από τιμές πεδίων που περιγράφουν μια οντότητα ή συσχέτιση.

Τύπος εγγραφών: είναι η δομή ενός συνόλου εγγραφών.

Στιγμιότυπο τύπου συσχετίσεων γονέα – παιδιού: μια (1) εγγραφή τύπου – γονέα και 0-N εγγραφές τύπου- παιδιού.

¹⁰ Λάζος Κ. Ε. Μαθήματα Βάσεων Δεδομένων, Θεσσαλονίκη 1999

Ιεραρχικό Σχήμα ή Ιεραρχία: είναι ένα πλήθος τύπων εγγραφών και συσχετίσεων γονέα – παιδιού.



Ιεραρχικό
Μοντέλο
Βάσης
Δεδομένων

Σχήμα 4: Ιεραρχικό μοντέλο ΒΔ

Ιδιότητες ενός ιεραρχικού σχήματος:

1. Υπάρχει ένας τύπος εγγραφών που λέγεται ρίζα και είναι αποκλειστικά γονέας.
2. Κάθε τύπος εγγραφών μπορεί να είναι «παιδί» σε ακριβώς ένα γονέα.
3. Κάθε τύπος εγγραφών γονέα μπορεί να έχει περισσότερα από 1 παιδιά.
4. Αν ένας γονέας έχει περισσότερα από 1 παιδιά, τότε αυτά είναι σε διάταξη.
5. Ένας τύπος εγγραφών που δεν έχει παιδιά, λέγεται φύλλο του σχήματος.

Ιεραρχικό σχήμα δεδομένων: είναι μια δενδρική δομή δεδομένων όπου:

- Ο κόμβος είναι ένας τύπος εγγραφών
- Η ακμή είναι μια σχέση γονέα παιδιού

Περιορισμοί ακεραιότητας στο ιεραρχικό ΜΔ

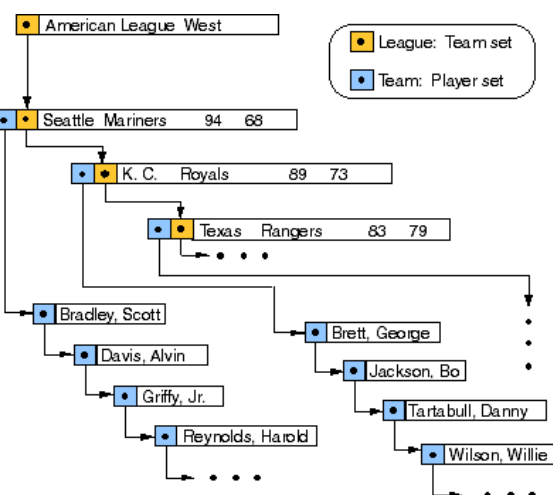
1. Μόνο οι «ρίζες» μπορούν να μην έχουν εγγραφή – γονέα
2. Μια εγγραφή παιδί εισάγεται μόνο κάτω από το γονέα της
3. Η διαγραφή του γονέα συνεπάγεται διαγραφή όλων των παιδιών
4. Αν μια εγγραφή-παιδί έχει δύο ή περισσότερες εγγραφές-γονέα (του ίδιου τύπου), τότε αυτή τηρείται σε όσα αντίγραφα, όσα και οι γονείς
5. Επιτρέπονται «εικονικοί» γονείς (εγγραφές που παρεμβάλλονται μεταξύ σημασιολογικού γονέα και παιδιού)

6. Κάθε τύπος εγγραφής μπορεί να είναι εικονικός γονέας μόνο για έναν (διαφορετικό) τύπο εγγραφής

6.2. Δικτυακό μοντέλο δεδομένων

Είναι προϊόν μιας ομάδας εργασίας DBTG (Database Task Group) του οργανισμού CODASYL (1971), το οποίο αναθεωρήθηκε τα έτη 1978 & 1981. Ως γλώσσα προγραμματισμού χρησιμοποιήθηκε η COBOL.

Γενικά: Είναι μια παραλλαγή του ιεραρχικού μοντέλου. Στο δικτυακό (network) μοντέλο, τα στοιχεία τοποθετούνται σ' ένα επίπεδο ιεραρχίας, αλλά κάθε στοιχείο μπορεί να συσχετιστεί με πολλά στοιχεία είτε σ' ένα κατώτερο ή σ' ένα ανώτερο επίπεδο. Ιδιαίτερα χρήσιμο για πολλαπλές σχέσεις. Ένα «παιδί» μπορεί να έχει πάνω από έναν «πατέρα».



Δικτυακό
μοντέλο
Βάσης
Δεδομένων

Σχήμα 5: Δικτυακό μοντέλο ΒΔ

Πλεονεκτήματα και μειονεκτήματα

- Ελαχιστοποίηση πλεονασμού δεδομένων και μείωση του χρόνου απόκρισης.
- Ραγδαία αύξηση του αριθμού των δεικτών, με αποτέλεσμα η συντήρηση και λειτουργία του συστήματος να γίνεται συχνά πολύπλοκη.

Σύντομη αναφορά στο Δικτυακό ΜΔ

Δύο δομές δεδομένων: εγγραφές και σύνολα.

Εγγραφή (record) είναι μια ομάδα συσχετιζόμενων τιμών δεδομένων. Μια περιγραφή μιας ομάδας ονομάζεται τύπος εγγραφής (record type). Τα επιμέρους στοιχεία ενός τύπου εγγραφής λέγονται στοιχεία δεδομένων (data items) ή γνωρίσματα (attributes).

Παράδειγμα:

Δελτίο Ταυτότητας (γενικά) -> τύπος εγγραφής «Ταυτότητα»

Δελτίο Ταυτότητας του κ. Παπαδόπουλου -> εγγραφή

Στοιχεία δεδομένων: όνομα, επώνυμο, πατρώνυμο, μητρώνυμο, ημ. γέννησης κλπ

Σύνολα

Τύπος συνόλου (set type): η περιγραφή μιας συσχέτισης 1:N μεταξύ δύο τύπων εγγραφών: {όνομα, τύπος 1: «ιδιοκτήτης», τύπος 2: «μέλος»}

Οι συσχετίσεις μεταξύ δεδομένων είναι διαφορετικές ως προς την πολλαπλότητα:

Δελτίο ταυτότητας – φυσικό πρόσωπο

- Ένα φυσικό πρόσωπο έχει μόνο ένα δελτίο ταυτότητας
- Ένα δελτίο ταυτότητας αντιστοιχεί μόνο σε ένα φυσικό πρόσωπο

Σπουδαστής – Σχολή

- Ένας σπουδαστής σπουδάζει σε μια σχολή
- Σε μια σχολή σπουδάζουν πολλοί σπουδαστές

Οι συσχετίσεις μεταξύ των δεδομένων είναι διαφορετικές ως προς τους προορισμούς:

- Δεν μπορεί να υπάρχει Δελτίο Ταυτότητας που δεν αντιστοιχεί σε φυσικό πρόσωπο
- Δεν μπορεί να υπάρχει φυσικό πρόσωπο που δεν έχει Δελτίο Ταυτότητας
- Σε μια σχολή μπορεί να μη φοιτούν σπουδαστές
- Σε μια σχολή δεν μπορεί να φοιτούν πάνω από 1000 σπουδαστές

Τα σύνολα του δικτυακού μοντέλου δεδομένων έχουν ένα διακεκριμένο στοιχείο, είναι διατεταγμένα και διαφέρουν από τα σύνολα με τη μαθηματική έννοια. Υπάρχει ένα ιδιαίτερο σύνολο, το Σύστημα.

- Σύνολα με ιδιοκτήτη το Σύστημα
- Σημεία εισόδου στη Βάση Δεδομένων
- Μέσο διάταξης εγγραφών

Η αναπαράσταση συνόλων γίνεται με:

- Συνήθως κυκλικές κινήσεις
- Διπλά συνδεδεμένες κυκλικές λίστες
- Δείκτης προς ιδιοκτήτη
- Συνεχόμενες εγγραφές μετά την εγγραφή – ιδιοκτήτη
- Πίνακες δεικτών που αποθηκεύονται με την εγγραφή ιδιοκτήτη
- Ευρετήρια

6.3.Σχεσιακό μοντέλο δεδομένων

Προτάθηκε από τον Ted Codd (IBM) το 1970

Γενικά: Η αναπαράσταση όλων των δεδομένων γίνεται μέσω δισδιάστατων πινάκων που ονομάζονται «σχέσεις» (relations). Μπορεί να συσχετίσει δεδομένα τέτοιων πινάκων, αρκεί οι πίνακες αυτοί να έχουν κοινό στοιχείο δεδομένων.

Το σχεσιακό (relational) μοντέλο έχει επικρατήσει σήμερα στην αναπαράσταση των δεδομένων καθώς διαθέτει σημαντικά πλεονεκτήματα ως προς τα άλλα δύο και οι βάσεις δεδομένων που σχεδιάζονται σύμφωνα με αυτό αποκαλούνται σχεσιακές (relational databases). Με τις σχεσιακές βάσεις δεδομένων διαθέτουμε έναν σαφή, απλό και εύκολα κατανοητό τρόπο για να μπορέσουμε να αναπαραστήσουμε και να διαχειριστούμε τα δεδομένα μας. Υστερούν μόνο σε ταχύτητα υπολογισμών και σε χώρο αποθήκευσης, αλλά μόνο όταν έχουμε να κάνουμε με πολύ μεγάλες βάσεις δεδομένων.

Στο μοντέλο αυτό οι βάσεις δεδομένων περιγράφονται με αυστηρές μαθηματικές έννοιες και ο χρήστης βλέπει τις οντότητες και τις συσχετίσεις με τη μορφή πινάκων (tables) και σχέσεων (relations) αντίστοιχα.

Ένας πίνακας (table) αποτελείται από γραμμές (rows) και στήλες (columns), όπου τοποθετούμε τα στοιχεία σε οριζόντια και κάθετη μορφή. Η κάθε στήλη του πίνακα χαρακτηρίζει κάποια ιδιότητα της οντότητας και αποκαλείται χαρακτηριστικό (attribute) ή πεδίο (field), ενώ η κάθε γραμμή του πίνακα περιέχει όλες τις πληροφορίες (στήλες) που αφορούν ένα στοιχείο της οντότητας και αποκαλείται πλειάδα (tuple) ή εγγραφή (record).

Κάθε πεδίο του πίνακα μπορεί να πάρει ορισμένες μόνο τιμές, οι οποίες μπορεί να καθορίζονται από τον τύπο δεδομένων της ιδιότητας, όπως ονόματα ή

αριθμοί για παράδειγμα, ή και από αυτό που εκφράζει, όπως το ότι δεν μπορούμε να έχουμε αρνητικό βάρος ή αρνητικό ΑΦΜ, για παράδειγμα. Το σύνολο των αποδεκτών τιμών μιας οντότητας αποκαλείται πεδίο ορισμού (domain).

7. ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΘΕΟΛΟΓΙΚΗ ΕΠΙΣΤΗΜΗ

Με την αλματώδη πρόοδο της επιστήμης της Πληροφορικής και των δυνατοτήτων της άνοιξαν νέοι ορίζοντες συνεργασίας και στις περιοχές του θεωρητικού λόγου. Ειδικότερα πρόσφερε άμεσα θεωρητικά και πρακτικά μοντέλα στην προσέγγιση της τεχνολογίας και των θεωρητικών επιστημών, άρα και στη Θεολογία.

Με τη βοήθεια της Πληροφορικής διευκολύνεται η προσέγγιση και ανάλυση των θεολογικών κειμένων, με ένα σύνολο συστηματικών γνώσεων και μεθόδων έρευνας¹¹.

Οι δυνατότητες που παρέχουν, πλέον, οι εφαρμογές της Πληροφορικής είναι:

- Η τεκμηρίωση θεολογικών συμπερασμάτων σε ένα προτεινόμενο μοντέλο σημασιολογικού ευρητηριασμού θεολογικών δημοσιεύσεων
- Η χρηστική και μεθοδική παράσταση θεολογικών γνώσεων με ένα συστηματικό τρόπο μέσα από διαγράμματα
- Η στατιστική προσέγγιση του προσωπικού γλωσσικού ύφους για τον προσδιορισμό της πατρότητας ανώνυμων κειμένων της πατερικής γραμματείας
- Η ψηφιοποίηση ανέκδοτων κωδίκων καθώς και σπανίων εκδόσεων
- Η ψηφιοποίηση μεγάλων σε όγκο λεξικών, ταμείων, ευρητηρίων

Οι δυνατότητες αυτές διευκολύνουν τη θεολογική επιστήμη στη προσέγγιση και επεξεργασία των πηγών της, μειώνουν το χρόνο επεξεργασίας τους, με αποτέλεσμα και τη μείωση του κόστους της ερευνητικής εργασίας.

7.1. Οι Βάσεις Δεδομένων στη θεολογική επιστήμη

Οι δυνατότητες των βάσεων δεδομένων προσφέρουν λύσεις για την αξιοποίηση μεγάλου όγκου γραμματειακού υλικού της παγκόσμιας θεολογικής παραγωγής. Από

¹¹ Βασίλειος Γιούλτσης, Εκκλησιαστικές και θεολογικές βάσεις δεδομένων, σελ. 229-239

τις λύσεις αυτές η πιο λειτουργική είναι η καταγραφή και οργάνωση του υλικού σε βάσεις δεδομένων, δηλαδή η ταξινόμηση των πληροφοριών κατά τρόπο που προσφέρει αποτέλεσμα σε όλες τις πιθανές απαιτήσεις της ευρετηριακής έρευνας.

Βασικό πρόβλημα στην καταγραφή των έργων της Πατερικής γραμματείας αποτελεί ο μεγάλος όγκος των πληροφοριών και ο χρόνος που απαιτείται για επεξεργασία.

Είναι σημαντικό να σημειώσουμε ότι η ύπαρξη διγραφικού τρόπου αποτύπωσης των ελληνικών κειμένων (μονοτονικό και πολυτονικό) προκαλεί σοβαρές δυσχέρειες στην καταγραφή τους. Η κατάσταση αυτή αποτελεί σοβαρό εμπόδιο για τη δημιουργία μιας ελληνικής βάσης δεδομένων, με περιεχόμενό της την ελληνική γραμματεία από τότε που γράφτηκε το πρώτο ελληνικό κείμενο μέχρι και σήμερα.

Η δημιουργία ενός προτύπου θα εξυπηρετούσε τον διεθνή χρήστη στη χρήση υλικού μιας Τράπεζας Πληροφοριών με αυθεντικά ελληνικά κείμενα.

Η μεγάλη ανάπτυξη και χρήση του Διαδικτύου (Internet) προσφέρει όλο και περισσότερες δυνατότητες αξιοποίησης του υλικού των βιβλιοθηκών που οργανώνονται σε βάσεις δεδομένων.

Εν κατακλείδι η συνεργασία ειδικών επιστημόνων από τους τομείς της Πληροφορικής και της Θεολογίας έχει πολλά να προσφέρει στην περαιτέρω ανάπτυξη του θεολογικού λόγου.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Βεσκούκης Βασίλειος, *Βάσεις Δεδομένων*, Πανεπιστημιακές παραδόσεις
2. Καρακαπιλίδης Νικόλαος, *Πληροφοριακά Συστήματα Διοίκησης*, Πανεπιστημιακές παραδόσεις
3. Λάζος Κ. Ε. *Μαθήματα Βάσεων Δεδομένων*, Θεσσαλονίκη 1999
4. Μάργαρης Αθανάσιος, *Θεωρία και εφαρμογές στις Βάσεις Δεδομένων*, Πανεπιστημιακές παραδόσεις
5. Βασίλειος Γιούλτσης, *Εκκλησιαστικές και θεολογικές βάσεις δεδομένων*, σελ. 229-239
6. F. D. Saussure, *Μαθήματα γενικής γλωσσολογίας*, εκδ. Παπαζήση (1979). Μετάφραση 3ης γαλλ. εκδ. (Γενεύη 1915)

Διαδίκτυο

1. <http://databases.about.com/od/administration/a/glossary.htm> προσπελάστηκε 12-2-2007
2. <http://dide.flo.sch.gr/Plinet/Tutorials/Tutorials-DatabasesTheory.html> προσπελάστηκε 12-2 -2007
3. <http://en.wikipedia.org/wiki/Database> προσπελάστηκε 20-9-2007
4. <http://office.microsoft.com/el-gr/access/HA100644501032.aspx> προσπελάστηκε 20-9-2007
5. <http://el.wikipedia.org/wiki/Πληροφορία> προσπελάστηκε 20-9-2007