

ΚΕΦΑΛΑΙΟ 4

ΔΙΑΧΕΙΡΙΣΗ ΚΑΙ ΟΡΓΑΝΩΣΗ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Δεδομένα σε χάος

«Μάρκο, τις προάλλες ήμουν στο γραφείο και παρατηρούσα τα παιδιά που περνούσαν στο σύστημα κρατήσεις για δωμάτια και διάφορες άλλες υπηρεσίες...»

«Ναι; Και τι παρατήρησες;»

«Παρατήρησα πως σε ό,τι ερχόταν από τηλέφωνο, email ή με επιτόπου κράτηση από πελάτες στο γραφείο καθυστερούσαν πολύ στην εισαγωγή των δεδομένων στον υπολογιστή».

«Ίσως, Στέλλα, να είναι απλώς λίγο αργοί στην πληκτρολόγηση. Δεν είναι δα σημαντικό. Πώς και έδωσες σημασία σε κάτι τέτοιο;»

«Θα σου πω, Μάρκο. Ζήτησα να μου δείξουν τι πεδία έχουν να συμπληρώσουν κάθε φορά που εισάγουν κάτι στον υπολογιστή. Λοιπόν, ήταν ένα σκέτο χάος!»

«Δηλαδή;»

«Πολλά πεδία μου φάνηκαν τελείως άσχετα και, από ό,τι μου είπαν, κανείς δεν συμπλήρωνε ποτέ τίποτα σε αυτά. Μετά περίμενα ότι θα καταχώριζαν κάπου στοιχεία που έχουν μεγάλο ενδιαφέρον, όπως, ας πούμε, από ποιο κανάλι έγινε μια κράτηση, αλλά δεν υπήρχε πεδίο για αυτό».

«Ίσως φταίει το GDPR⁵».

«Μάρκο, αυτό είναι τελείως άσχετο! Συνεχίζω. Ένα άλλο που παρατήρησα είναι ότι σε κάποιες περιπτώσεις υπήρχαν στα έντυπα κάποιες πληροφορίες, αλλά δεν τις περνούσαν στο σύστημα, γιατί, από ό,τι είπαν, δεν υπήρχε ούτε για αυτά τα στοιχεία πεδίο. Δηλαδή συλλέγουμε χρήσιμα δεδομένα, αλλά καταλήγουμε να τα πετάμε...»

«Στέλλα, όπως μου τα περιγράφεις, μάλλον κάτι δεν γίνεται σωστά. Αλλά τι μπορούμε να κάνουμε; Να αλλάξουμε τα έντυπα;»

«Φυσικά και όχι, Μάρκο! Άλλωστε ο στόχος είναι η ψηφιοποίηση των πληροφοριών, όχι να γυρίσουμε στα έντυπα. Δηλαδή χρειαζόμαστε και κάποια έντυπα και ναι, πρέπει να γίνουν αλλαγές σε αυτά. Αλλά αυτό είναι τελείως δευτερεύον. Άλλο είναι αυτό που χρειαζόμαστε».

«Τι δηλαδή;»

«Χρειάζεται επειγόντως να οργανώσουμε τα δεδομένα μας, να καταγράψουμε τις ανάγκες λαμβάνοντας φυσικά υπόψη το GDPR, να δούμε τα μοντέλα που υποστηρίζουν τα συστήματα και να συγχρονίσουμε τη συλλογή των δεδομένων με την αποθήκευση και τη γενικότερη διαχείρισή τους, συνυπολογίζοντας τις ανάγκες μας για ανάλυση των δεδομένων».

«Λίγο θεωρητικά μου ακούγονται όλα αυτά. Όσο για την ανάλυση, μου χτυπάς πάλι αυτό που έγινε με τη διαχείριση του στόλου;»

«Μάρκο, δεν σου χτυπάω τίποτα. Πρέπει να οργανώσουμε καλύτερα τις βάσεις δεδομένων μας. Πεδία, εγγραφές, πίνακες, τα πάντα. Αυτό».



Περιγράφοντας τις ατέλειες του συστήματος.

Πηγή: shutterstock

«Και πώς θα γίνει αυτό; Δηλαδή τι χρειάζεται να γίνει; Γιατί θεωρείς ότι έχουν πρόβλημα οι βάσεις δεδομένων; Δεν έχει φτιαχτεί κάτι σωστά;»

«Μάλλον πρέπει να ανοίξουμε άλλο ένα κεφάλαιο, Μάρκο, και να μιλήσουμε για διαχείριση και οργάνωση πληροφορίας. Λοιπόν, σημειώνεις;»

5. GDPR: General Data Protection Regulation. Ο Ευρωπαϊκός Κανονισμός για την Προστασία Προσωπικών Δεδομένων. Βλ. περισσότερα στο Κεφάλαιο 16 για την ασφάλεια.

Στόχοι κεφαλαίου

- Να γνωρίζετε τις βασικές έννοιες των βάσεων δεδομένων
- Να διακρίνετε τις έννοιες «βάση δεδομένων», «σύστημα διαχείρισης βάσης δεδομένων» και «προγράμματα εφαρμογών βάσεων δεδομένων»
- Να αναγνωρίζετε τον ρόλο των συστημάτων διαχείρισης βάσεων δεδομένων
- Να απαριθμείτε τα πλεονεκτήματα των συστημάτων διαχείρισης βάσεων δεδομένων
- Να περιγράφετε τα βασικά στάδια σχεδιασμού μιας βάσης δεδομένων
- Να γνωρίζετε πώς σχεδιάζεται μια βάση δεδομένων
- Να διακρίνετε τα διάφορα μοντέλα δεδομένων
- Να αναγνωρίζετε μια σχεσιακή βάση δεδομένων
- Να διακρίνετε τα βασικά δομικά στοιχεία του μοντέλου οντοτήτων-συσχετίσεων
- Να αναγνωρίζετε πώς εισάγονται δεδομένα σε μια βάση δεδομένων
- Να αντιλαμβάνεστε πώς να διαχειρίζεστε μια βάση δεδομένων
- Να διακρίνετε τη διαφορά μεταξύ του μοντέλου οντοτήτων-συσχετίσεων και των σχεσιακών βάσεων
- Να διακρίνετε τη διαφορά μεταξύ σχεσιακών και μη σχεσιακών βάσεων δεδομένων
- Να αναγνωρίζετε την αξία και τον ρόλο των βάσεων δεδομένων

Έννοιες που θα παρουσιαστούν

- Βάσεις δεδομένων
- Σύστημα διαχείρισης βάσεων δεδομένων
- Σχεσιακές βάσεις δεδομένων, εγγραφές, πίνακες
- Μοντέλο οντοτήτων-συσχετίσεων
- Οντότητες, γνωρίσματα, συσχετίσεις, πρωτεύοντα κλειδιά
- Ο κύκλος ζωής σχεδιασμού μιας βάσης δεδομένων
- NoSQL βάσεις δεδομένων

4.1 Βάσεις δεδομένων

Στην καθημερινή μας ζωή χρησιμοποιούμε και παράγουμε πληθώρα δεδομένων. Τα δεδομένα είναι ουσιαστικό στοιχείο για την επιβίωση και την εξέλιξη των επιχειρήσεων, καθώς αυτές διαρκώς συλλέγουν, αποθηκεύουν, συγκεντρώνουν, επεξεργάζονται, μοιράζονται και διαχειρίζονται δεδομένα. Για να είναι αποδοτική η διαχείριση των δεδομένων, τα δεδομένα αποθηκεύονται σε μια βάση δεδομένων. Κάθε μέρα αλληλεπιδρούμε με μια βάση δεδομένων μέσα από ένα πρόγραμμα εφαρμογής, όπως όταν βλέπουμε τον λογαριασμό μας σε ένα κοινωνικό μέσο, όταν κάνουμε μια ηλεκτρονική αγορά ή μια κράτηση σε κάποιο ξενοδοχείο.

4.1.1 Χρήσιμοι ορισμοί

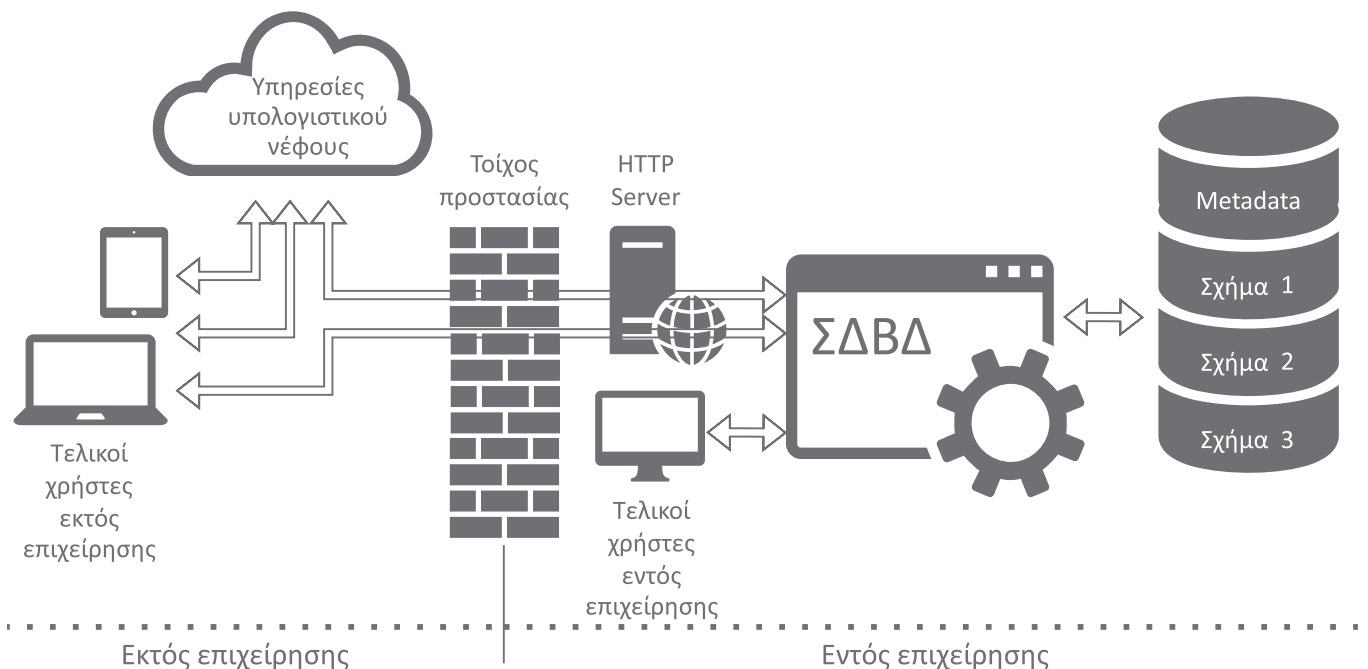
Υπάρχουν διάφοροι ορισμοί που περιγράφουν τι είναι μια βάση δεδομένων. Σύμφωνα με τους Elmasri & Navathe (2011), «βάση δεδομένων είναι μια συλλογή από σχετιζόμενα δεδομένα». Επίσης μια βάση δεδομένων μπορεί να οριστεί ως «μια διαμοιραζόμενη συλλογή λογικά σχετιζόμενων δεδομένων και η περιγραφή της, σχεδιασμένη έτσι ώστε να καλύπτει τις ανάγκες πληροφοριών ενός οργανισμού» (Connolly & Begg, 2015). Βάση δεδομένων είναι μια ενιαία, ενδεχομένως μεγάλη συλλογή από:

- **δεδομένα:** ακατέργαστα γεγονότα που αφορούν τον χρήστη (data)
- **μεταδεδομένα:** δεδομένα που περιγράφουν τα δεδομένα (metadata)

Τα μεταδεδομένα περιγράφουν τα χαρακτηριστικά των δεδομένων και το σύνολο των σχέσεων μεταξύ των δεδομένων που αποθηκεύονται σε μια βάση δεδομένων. Για παράδειγμα, ο αριθμός μητρώου κοινωνικής ασφάλισης είναι ένα δεδομένο, ενώ ο τύπος τιμών του –π.χ. ότι πρόκειται για αριθμό–, καθώς και το πλήθος των ψηφίων του –π.χ. ότι πρόκειται για έναν 11ψήφιο αριθμό– αποτελούν τα μεταδεδομένα. Τα μεταδεδομένα παρέχουν μια ολοκληρωμένη εικόνα των δεδομένων που αποθηκεύονται στη βάση δεδομένων και επιτρέπουν τον χαρακτηρισμό μιας βάσης δεδομένων ως μιας «συλλογής αυτοπεριγραφόμενων δεδομένων».

Μια βάση δεδομένων σχεδιάζεται για να αναπαραστήσει μια άποψη του πραγματικού κόσμου, έναν μικρόκοσμο, δημιουργείται για να εξυπηρετήσει έναν συγκεκριμένο σκοπό και προορίζεται για συγκεκριμένες ομάδες χρηστών και προκαθορισμένη χρήση. Το μέγεθος και η πολυπλοκότητα των βάσεων δεδομένων ποικίλλουν. Για παράδειγμα, η βάση δεδομένων ενός μικρού ξενοδοχείου μπορεί να αποτελείται από κάποιες χιλιάδες εγγραφές και να έχει μια απλή δομή, ενώ η βάση δεδομένων ενός μεγάλου διαδικτυακού ταξιδιωτικού γραφείου, όπως το Booking.com, περιλαμβάνει εκατομμύρια εγγραφές καταλυμάτων και χρηστών και αποθηκεύει πληροφορίες για τις κρατήσεις, τα σχόλια και τις βαθμολογίες των χρηστών.

Ένα **Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ)** (Database Management System, DBMS) είναι μια συλλογή από προγράμματα που επιτρέπουν στους χρήστες να ορίσουν, να δημιουργήσουν, να διαχειριστούν, να συντηρήσουν και να ελέγξουν την πρόσβαση σε μια βάση δεδομένων. Το ΣΔΒΔ είναι το λογισμικό το οποίο αλληλεπιδρά τόσο με τα προγράμματα εφαρμογών των χρηστών όσο και με την ίδια τη βάση δεδομένων, επιτρέποντας τη μεταξύ τους επικοινωνία (βλ. Σχήμα 4.1).



ΣΧΗΜΑ 4.1

Ένα ΣΔΒΔ διαμεσολαβεί στην αλληλεπίδραση των χρηστών της βάσης και των δεδομένων.

4.1.2 Ο ρόλος των Συστημάτων Διαχείρισης Βάσεων Δεδομένων

Ο ρόλος ενός ΣΔΒΔ είναι ιδιαίτερα σημαντικός, καθώς λαμβάνει όλα τα αιτήματα των προγραμμάτων εφαρμογής και τα υλοποιεί κρύβοντας την πολυπλοκότητα της βάσης δεδομένων. Η χρήση των ΣΔΒΔ προσφέρει σημαντικά πλεονεκτήματα (Coronel & Morris, 2018· Connolly & Begg, 2015· Elmasri & Navathe, 2011):

- *Βελτιωμένο διαμοιρασμό δεδομένων:* Το ΣΔΒΔ συμβάλλει στη δημιουργία ενός περιβάλλοντος μέσω του οποίου οι χρήστες έχουν καλύτερη πρόσβαση στα δεδομένα, επιτρέποντας την κοινή χρήση των δεδομένων στη βάση δεδομένων μεταξύ πολλών εφαρμογών και πολλών χρηστών.
- *Βελτιωμένη ασφάλεια δεδομένων:* Το ΣΔΒΔ παρέχει ένα ολοκληρωμένο πλαίσιο για την προστασία της ιδιωτικότητας των δεδομένων και την επιβολή πολιτικών ασφάλειας, καθώς όσο μεγαλώνει ο αριθμός των χρηστών που έχουν πρόσβαση στα δεδομένα τόσο αυξάνει και ο κίνδυνος παραβίασης της ασφάλειας των δεδομένων.
- *Καλύτερη ενοποίηση δεδομένων:* Η πρόσβαση σε καλά διαχειριζόμενα δεδομένα προωθεί μια ενοποιημένη και ολοκληρωμένη όψη των λειτουργιών μιας επιχείρησης, με αποτέλεσμα να είναι πιο εύκολο να περιγραφεί ο τρόπος με τον οποίο οι ενέργειες σε ένα τμήμα μιας επιχείρησης επηρεάζουν τα άλλα τμήματα.
- *Έλεγχος των πλεονασμών δεδομένων:* Η αποθήκευση πολλές φορές των ίδιων δεδομένων, όπως η αποθήκευση των δεδομένων πελατών από το τμήμα κρατήσεων και το τμήμα πωλήσεων ενός ξενοδοχείου, οδηγεί στον πλεονασμό δεδομένων (data redundancy), με συνέπεια τη σπατάλη αποθηκευτικού χώρου και την ασυνέπεια των δεδομένων. Το ΣΔΒΔ έχει στόχο να ελέγξει τον πλεονασμό δεδομένων, έτσι ώστε να μην αποθηκεύονται αντίγραφα των ίδιων δεδομένων. Παρ' όλα αυτά, δεν είναι στόχος η πλήρης εξάλειψη του πλεονασμού δεδομένων, καθώς ενδέχεται να επιλεγθεί η διατήρηση πολλαπλών αντιγράφων με στόχο τη βελτίωση της απόδοσης.
- *Ελαχιστοποίηση ασυνέπειας δεδομένων:* Η εμφάνιση διαφορετικών εκδόσεων των ίδιων δεδομένων σε διαφορετικά σημεία, όπως η διαφορετική μορφή του ονόματος ενός εργαζομένου στο τμήμα προσωπικού και στο τμήμα μισθοδοσίας, οδηγεί στην ασυνέπεια δεδομένων. Μια καλά σχεδιασμένη βάση δεδομένων μειώνει σημαντικά την πιθανότητα ασυνέπειας δεδομένων.
- *Βελτιωμένη πρόσβαση δεδομένων:* Το ΣΔΒΔ καθιστά δυνατή την παραγωγή γρήγορων απαντήσεων σε τυχαία ερωτήματα. Ένα ερώτημα από την οπτική των βάσεων δεδομένων είναι ένα συγκεκριμένο αίτημα που προωθείται στο ΣΔΒΔ για τον χειρισμό των δεδομένων, π.χ. για την ανάγνωση ή την ενημέρωση δεδομένων. Ένα παράδειγμα ερωτήματος είναι «Πόσοι πελάτες μας έχουν ταξιδέψει με την εταιρεία μας περισσότερες από δύο φορές τον τελευταίο χρόνο;».
- *Βελτιωμένη λήψη αποφάσεων:* Η καλύτερη διαχείριση δεδομένων και η βελτιωμένη πρόσβαση σε δεδομένα καθιστούν δυνατή την παραγωγή πληροφοριών καλύτερης ποιότητας, στις οποίες βασίζονται καλύτερες αποφάσεις. Η ποιότητα των παραγόμενων πληροφοριών εξαρτάται από την ποιότητα των υποκείμενων δεδομένων. Η ποιότητα των δεδομένων αφορά την ακρίβεια, την εγκυρότητα και την επικαιρότητα των δεδομένων. Αν και το ΣΔΒΔ δεν εξασφαλίζει την ποιότητα των δεδομένων, παρέχει ένα πλαίσιο που διευκολύνει τη διασφάλισή της.
- *Αυξημένη παραγωγικότητα τελικού χρήστη:* Η διαθεσιμότητα των δεδομένων σε συνδυασμό με τα εργαλεία που μετασχηματίζουν τα δεδομένα σε χρήσιμες πληροφορίες δίνει τη δυνατότητα στους τελικούς χρήστες να λαμβάνουν γρήγορες, ενημερωμένες αποφάσεις, οι οποίες μπορούν να κάνουν τη διαφορά, οδηγώντας σε επιτυχημένες ή αποτυχημένες ενέργειες.
- *Βελτιωμένοι μηχανισμοί τήρησης εφεδρικών αντιγράφων και ανάκαμψης:* Το ΣΔΒΔ παρέχει υπηρεσίες τήρησης εφεδρικών αντιγράφων και μηχανισμούς ανάκαμψης μετά από βλάβες ή αποτυχίες του υλικού ή του λογισμικού, με στόχο να διασφαλίσει την απρόσκοπτη λειτουργία της βάσης δεδομένων και την επαναφορά της στην κατάσταση πριν από την αποτυχία του συστήματος.

Η βάση δεδομένων του Booking.com

Η βάση δεδομένων του Booking.com περιλαμβάνει πάνω από 29 εκατομμύρια καταχωρίσεις καταλυμάτων και πάνω από 198 εκατομμύρια κριτικές χρηστών, ενώ υποστηρίζει την πραγματοποίηση 1,5 εκατομμυρίου κρατήσεων για διανυκτερεύσεις κάθε ημέρα, δηλαδή περίπου 1.000 κρατήσεις κάθε λεπτό. Το ΣΔΒΔ που χρησιμοποιεί το Booking.com για πάνω από 15 χρόνια είναι το MySQL (MySQL, 2018) για τους ακόλουθους λόγους:

- *Κλιμάκωση*: Υποστηρίζει την κλιμάκωση της εταιρείας και τις απαιτητικές ανάγκες της επιχείρησης
- *Ταχύτητα*: Παρέχει διαρκώς βελτιωμένη απόδοση και ταχύτατες συναλλαγές
- *Αξιοπιστία*: Είναι ιδιαίτερα αξιόπιστη και δεν εμφανίζονται προβλήματα
- *Ευκολία*: Έχει ένα εύκολο και ευέλικτο περιβάλλον διαχείρισης

Το Booking.com χρησιμοποιεί το MySQL ως το σύστημα υποστήριξης (backend) για την ανάκτηση και την αποθήκευση δεδομένων υποστηρίζοντας όλες τις λειτουργίες και εφαρμογές, όπως:

- Την περιήγηση σε καταλύματα μέσω του παγκόσμιου ιστού
- Την πρόσβαση στον ιστότοπο μέσω κινητού τηλεφώνου
- Την ανάγνωση κριτικών για τα καταλύματα
- Την πραγματοποίηση των κρατήσεων
- Τη διαχείριση των πιστωτικών καρτών
- Την ολοκλήρωση των πληρωμών
- Την ανάλυση των δεδομένων

4.1.3 Προγράμματα εφαρμογών βάσεων δεδομένων

Οι χρήστες αλληλεπιδρούν με μια βάση δεδομένων μέσα από πληθώρα προγραμμάτων εφαρμογών. Πρόγραμμα εφαρμογών είναι ένα υπολογιστικό πρόγραμμα το οποίο αλληλεπιδρά με μια βάση δεδομένων στέλνοντας κατάλληλα ερωτήματα μέσω ενός ΣΔΒΔ. Στον κλάδο του τουρισμού και της φιλοξενίας μπορούμε να εντοπίσουμε προγράμματα εφαρμογών βάσεων δεδομένων για όλες τις προσφερόμενες υπηρεσίες και επιχειρήσεις, που πραγματοποιούνται συνήθως είτε μέσω του παγκόσμιου ιστού είτε μέσω κινητών εφαρμογών.

Αναζήτηση πακέτου διακοπών

Η αναζήτηση διάφορων πακέτων διακοπών σε ταξιδιωτικά γραφεία, καταλυμάτων και πτήσεων γίνεται μέσα από ποικίλα προγράμματα εφαρμογών και απαιτεί την πρόσβαση σε διάφορες βάσεις δεδομένων. Για παράδειγμα, ο χρήστης αναζητά ένα πακέτο διακοπών θέτοντας συγκεκριμένες προτιμήσεις, όπως τη χώρα προορισμού, το ανώτατο όριο χρημάτων που μπορεί να διαθέσει, καθώς και το χρονικό διάστημα που επιθυμεί. Το πρόγραμμα εφαρμογών στέλνει αυτές τις προτιμήσεις με τη μορφή ενός ερωτήματος στη βάση δεδομένων μέσω του ΣΔΒΔ. Από την πληθώρα των αποθηκευμένων δεδομένων στη βάση δεδομένων το ΣΔΒΔ αναζητά τα πακέτα διακοπών που ικανοποιούν τις προτιμήσεις του χρήστη και επιστρέφει τις σχετικές πληροφορίες στο πρόγραμμα εφαρμογών, όπου προβάλλονται στον χρήστη.

Κράτηση αεροπορικού εισιτηρίου

Η κράτηση ενός αεροπορικού εισιτηρίου ή ενός δωματίου είναι μια πιο πολύπλοκη διαδικασία, καθώς, αφού ο χρήστης επιλέξει π.χ. το αεροπορικό εισιτήριο που επιθυμεί και προχωρήσει στην κράτησή του, το σύστημα πρέπει να διασφαλίσει ότι το συγκεκριμένο εισιτήριο θα δεσμευτεί μόνο από τον συγκεκριμένο χρήστη. Υποθέστε ότι υπάρχει διαθέσιμο μόνο ένα εισιτήριο και δύο ταξιδιωτικοί πράκτορες προσπαθούν να προχωρήσουν στην κράτησή του. Το σύστημα διαχείρισης της βάσης δεδομένων είναι αυτό που πρέπει να αναγνωρίσει τη συγκεκρι-

μένη κατάσταση και να επιτρέψει να ολοκληρωθεί μόνο η μία κράτηση, ενημερώνοντας τον άλλο πράκτορα ότι δεν υπάρχουν πλέον διαθέσιμα εισιτήρια.

Αξιολόγηση καταλύματος σε πλατφόρμα κοινωνικής δικτύωσης

Η προσθήκη σχολίων και αξιολογήσεων για καταλύματα, εστιατόρια κ.ά. σε διάφορες πλατφόρμες κοινωνικής δικτύωσης γίνεται μέσα από διάφορους τρόπους πρόσβασης και διαθέσιμες εφαρμογές των συγκεκριμένων πλατφορμών. Ο χρήστης, έχοντας συνδεθεί στην πλατφόρμα με τον λογαριασμό του, έχει τη δυνατότητα να προσθέσει σχόλια και να αξιολογήσει κάποιο ξενοδοχείο ή κατάλυμα, ένα εστιατόριο ή ακόμα και μια εμπειρία που είχε σε κάποιον προορισμό. Τα σχόλια και η αξιολόγηση που προσθέτει συσχετίζονται τόσο με τον ίδιο όσο και με το αντικείμενο της κριτικής του στη βάση δεδομένων.

4.1.4 Μοντελοποίηση των δεδομένων

Ένας βασικός σκοπός των συστημάτων βάσεων δεδομένων είναι να παρέχουν στους χρήστες ένα επίπεδο αφαίρεσης δεδομένων (data abstraction). Αυτό σημαίνει ότι, ενώ οι χρήστες έχουν πρόσβαση και διαχειρίζονται τα δεδομένα, το σύστημα αποκρύπτει τις λεπτομέρειες που αφορούν τον τρόπο αποθήκευσης, οργάνωσης και διαχείρισής τους. Για να επιτευχθεί αυτή η αφαίρεση, είναι απαραίτητη η χρήση ενός μοντέλου δεδομένων, μιας συλλογής εννοιών που μπορούν να χρησιμοποιηθούν για να περιγράψουν τη δομή μιας βάσης δεδομένων.

Αφαίρεση δεδομένων

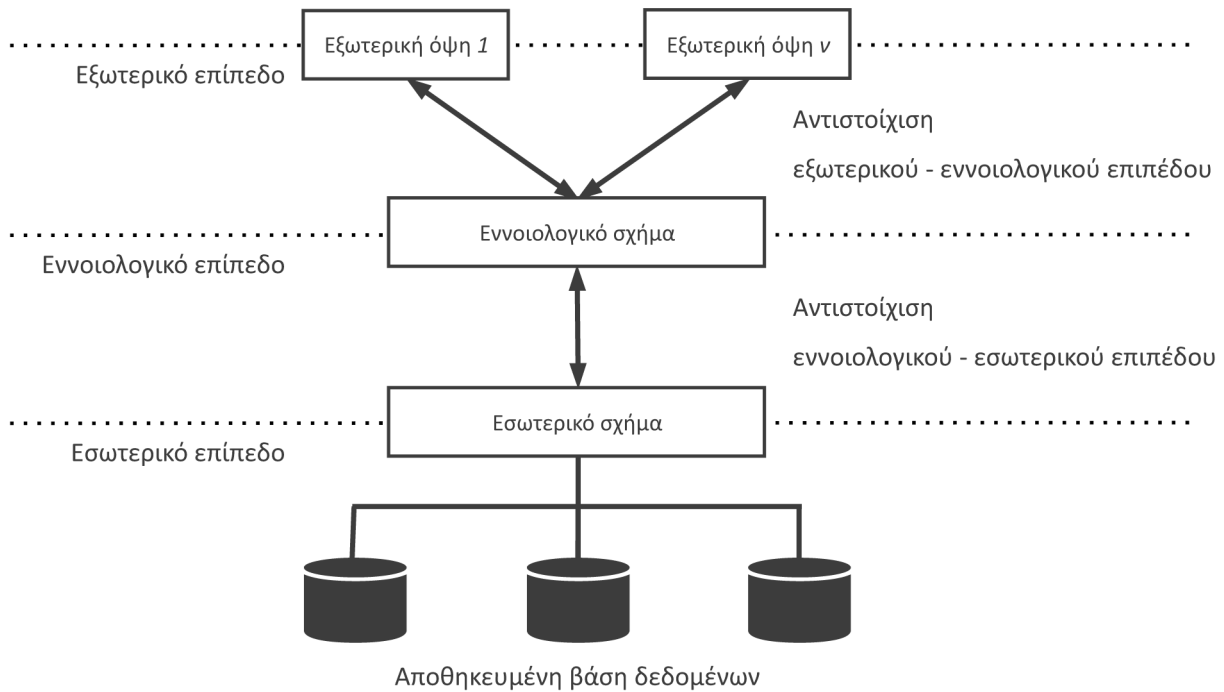
Για την αναπαράσταση των δεδομένων σε μια βάση δεδομένων χρησιμοποιούνται πολύπλοκες δομές δεδομένων. Εφόσον μεγάλο τμήμα των χρηστών των συστημάτων βάσεων δεδομένων δεν σχετίζονται με την επιστήμη της πληροφορικής, οι προγραμματιστές προσπαθούν να αποκρύψουν αυτή την πολυπλοκότητα μέσω τριών επιπέδων αφαίρεσης απλοποιώντας την αλληλεπίδραση των χρηστών με το σύστημα. Τα τρία επίπεδα αφαίρεσης επιτρέπουν την περιγραφή των δεδομένων σε τρία διακριτά επίπεδα και οδηγούν σε μια αρχιτεκτονική τριών επιπέδων (Tsichritzis & Klug, 1978· Tsichritzis & Lochovsky, 1982· ANSI, 1975), η οποία αποτελείται από το εξωτερικό, το εννοιολογικό και το εσωτερικό επίπεδο (βλ. Σχήμα 4.2).

- **Εξωτερικό επίπεδο ή επίπεδο όψης:** Αφορά τον τρόπο με τον οποίο αντιλαμβάνονται τα δεδομένα οι χρήστες. Οι χρήστες ενός συστήματος βάσης δεδομένων δεν χρειάζονται όλες τις πληροφορίες που είναι αποθηκευμένες και πρέπει να έχουν πρόσβαση μόνο σε ένα μέρος της βάσης δεδομένων. Το σύστημα μπορεί να παρέχει πολλές όψεις χρηστών για την ίδια βάση δεδομένων και κάθε όψη περιγράφει το μέρος της βάσης δεδομένων που ενδιαφέρει μια συγκεκριμένη ομάδα χρηστών, ενώ αποκρύπτει τα υπόλοιπα δεδομένα.
- **Εννοιολογικό επίπεδο ή λογικό επίπεδο:** Περιγράφει ποια δεδομένα αποθηκεύονται στη βάση δεδομένων και τις σχέσεις που υπάρχουν μεταξύ αυτών, αποτυπώνοντας μια καθολική εικόνα της βάσης δεδομένων μέσω ενός εννοιολογικού σχήματος.
- **Εσωτερικό επίπεδο ή φυσικό επίπεδο:** Περιγράφει τη φυσική δομή της βάσης δεδομένων και τον τρόπο με τον οποίο αποθηκεύονται πραγματικά τα δεδομένα.

Ανεξαρτησία δεδομένων

Η αρχιτεκτονική των τριών επιπέδων επιτρέπει την ανεξαρτησία δεδομένων, δηλαδή τη δυνατότητα να αλλάζουμε το σχήμα ενός επιπέδου σε ένα σύστημα βάσης δεδομένων χωρίς να χρειάζεται να αλλάξουμε το σχήμα του αμέσως ανώτερου επιπέδου. Υπάρχουν δύο τύποι ανεξαρτησίας δεδομένων:

- **Λογική ανεξαρτησία δεδομένων:** Είναι η δυνατότητα να αλλάζει το εννοιολογικό σχήμα, π.χ. με την επέκταση μιας βάσης δεδομένων, χωρίς να επηρεάζονται τα εξωτερικά σχήματα ή τα προγράμματα εφαρμογών.
- **Φυσική ανεξαρτησία δεδομένων:** Είναι η δυνατότητα να αλλάζουν το εσωτερικό σχήμα και ο τρόπος αποθή-



ΣΧΗΜΑ 4.2

Η αρχιτεκτονική τριών επιπέδων (γνωστή και ως αρχιτεκτονική ANSI/SPARC από την American National Standards Institute/ Standards Planning And Requirements Committee, που την πρότεινε) διαχωρίζει τις εφαρμογές χρήστη από τη φυσική βάση δεδομένων.

κευσης των δεδομένων, με στόχο, μεταξύ άλλων να βελτιωθεί η απόδοση χωρίς να επηρεαστεί το εννοιολογικό σχήμα.

Μοντέλα δεδομένων

Μοντέλο δεδομένων (data model) είναι μια ολοκληρωμένη συλλογή εννοιών για την περιγραφή και τη διαχείριση των δεδομένων, των σχέσεων μεταξύ των δεδομένων και των περιορισμών αυτών των δεδομένων. Ένα μοντέλο δεδομένων περιγράφει τον σχεδιασμό μιας βάσης δεδομένων σε φυσικό, λογικό επίπεδο και σε επίπεδο όψης. Τα μοντέλα δεδομένων μπορούν να ταξινομηθούν σε διάφορες κατηγορίες:

- **Σχεσιακό μοντέλο:** Το σχεσιακό μοντέλο (relational model) χρησιμοποιεί μια συλλογή πινάκων για να αναπαραστήσει τόσο τα δεδομένα όσο και τις σχέσεις μεταξύ τους. Η βασική έννοια σε αυτό το μοντέλο είναι η σχέση, η οποία μπορεί να αποτυπωθεί ως ένας δισδιάστατος πίνακας με γραμμές και στήλες. Η γραμμή κάθε πίνακα αναφέρεται ως εγγραφή ή πλειάδα και περιγράφει μια συλλογή σχετιζόμενων δεδομένων. Η στήλη κάθε πίνακα έχει ένα μοναδικό όνομα και περιγράφει ένα γνώρισμα της σχέσης. Το σχεσιακό μοντέλο είναι το πιο ευρέως διαδεδομένο μοντέλο δεδομένων και αυτό στο οποίο βασίζεται η πλειονότητα των σύγχρονων συστημάτων βάσεων δεδομένων.
- **Μοντέλο οντοτήτων-συσχετίσεων:** Το μοντέλο οντοτήτων-συσχετίσεων (entity-relationship model) χρησιμοποιεί ως βασική έννοια την οντότητα, η οποία περιγράφει ένα αντικείμενο ή μια έννοια του πραγματικού κόσμου και τις συσχετίσεις μεταξύ των οντοτήτων. Το μοντέλο οντοτήτων-συσχετίσεων χρησιμοποιείται στον σχεδιασμό των βάσεων δεδομένων.
- **Αντικειμενοστραφές μοντέλο:** Το αντικειμενοστραφές μοντέλο (object-oriented data model) περιγράφει τόσο τα δεδομένα όσο και τις μεταξύ τους σχέσεις με μια ενιαία δομή, το αντικείμενο, και έχει επηρεαστεί από τον αντικειμενοστραφή προγραμματισμό.
- **Μοντέλο ημιδομημένων δεδομένων:** Το μοντέλο ημιδομημένων δεδομένων επιτρέπει την περιγραφή δεδομένων όπου δεδομένα ίδιου τύπου μπορεί να έχουν διαφορετικά σύνολα χαρακτηριστικών, καθώς αυτό δεν επιτρέπεται από τα προηγούμενα μοντέλα δεδομένων. Τα μοντέλα ημιδομημένων δεδομένων περιλαμβάνουν

νουν την XML (Extensible Markup Language), τα μοντέλα που βασίζονται στην αποθήκευση ζεύγους κλειδιού-τιμής, καθώς και τα NoSQL μοντέλα, που χρησιμοποιούνται για τη διαχείριση μεγάλων δεδομένων.

Στις επόμενες ενότητες θα μελετήσουμε αναλυτικά το μοντέλο οντοτήτων-συσχετίσεων και το σχεσιακό μοντέλο, ενώ θα παρουσιάσουμε και τις NoSQL βάσεις δεδομένων.

4.2 Σχεσιακές βάσεις δεδομένων

Η πλειονότητα των σύγχρονων συστημάτων βάσεων δεδομένων βασίζονται στο σχεσιακό μοντέλο, το οποίο παρουσιάστηκε από τον E. F. Codd (Codd, 1970) και παρέχει μια μοντελοποίηση των δεδομένων στο λογικό επίπεδο.

4.2.1 Σχεσιακό μοντέλο

Το σχεσιακό μοντέλο βασίζεται στη μαθηματική έννοια της σχέσης, η οποία αναπαρίσταται ως ένας πίνακας, και μια βάση δεδομένων αποτελείται από μια συλλογή πινάκων. Για την ανάλυση του σχεσιακού μοντέλου θα βασιστούμε σε ένα απλό και οικείο παράδειγμα βάσης δεδομένων, μια βάση δεδομένων για ένα ξενοδοχείο στην οποία αποθηκεύονται δεδομένα που αφορούν τους πελάτες, τα δωμάτια, τις κρατήσεις και τις αξιολογήσεις των πελατών.

Σχέση

Μια σχέση είναι ένας δισδιάστατος πίνακας από γραμμές και στήλες. Για παράδειγμα, οι πληροφορίες για τους πελάτες του ξενοδοχείου αποθηκεύονται στη σχέση ΠΕΛΑΤΗΣ. Η σχέση ΠΕΛΑΤΗΣ αποτυπώνεται ως ένας δισδιάστατος πίνακας, ο οποίος αποτελείται από 10 γραμμές (πλειάδες/εγγραφές) και 9 στήλες (γνωρίσματα). Στον πίνακα αυτό κάθε γραμμή περιγράφει έναν συγκεκριμένο πελάτη. Το σχήμα μιας σχέσης είναι το όνομα της σχέσης σε συνδυασμό με τα ονόματα των γνωρισμάτων της και τα πεδία ορισμού τους. Το στιγμιότυπο μιας σχέσης είναι το σύνολο των πλειάδων με τις συγκεκριμένες τιμές των γνωρισμάτων τους μία δεδομένη στιγμή, όπως π.χ. αποτυπώνεται στον Πίνακα 4.1.

Γνώρισμα

Ένα γνώρισμα (*attribute*) μιας σχέσης είναι κάθε στήλη του πίνακα· κάθε γνώρισμα έχει ένα ορισμένο, και μοναδικό για τη σχέση, όνομα. Στη σχέση ΠΕΛΑΤΗΣ υπάρχουν 9 γνωρίσματα, και ενδεικτικά γνωρίσματα είναι ο ΚωδΠε-

ΠΙΝΑΚΑΣ 4.1

Ένα στιγμιότυπο της σχέσης ΠΕΛΑΤΗΣ μιας σχεσιακής βάσης δεδομένων.

ΠΕΛΑΤΗΣ								
ΚωδΠελάτη	Επώνυμο	Όνομα	Email	Τηλέφωνο	Διεύθυνση	Πόλη	Χώρα	ΕτοςΓεν
1	Walker	Tate	hermanab@outlook.com	00285085161482	315 Bridle St.	Adrian	Liechtenstein	1986
2	Izaiah	Randall	csilvers@live.com	00236468169687	8407 South Lane	Allentown	Portugal	1995
3	Rodney	Branch	oevens@att.net	00368927081481	740 Green Hill Road	Anoka	Luxembourg	1951
4	Natalia	Stevenson	hermanab@live.com	00565736102826	675 53rd Drive	Bemidji	Monaco	1975
5	Yael	Suarez	pjacklam@yahoo.com	<NULL>	121 Hilltop St.	Bergenfield	Montenegro	1967
6	Kaitlin	Koch	jmcnamara@yahoo.ca	00109455029343	114 Birchpond Street	Bountiful	Italy	1954
7	Marques	Cox	panolex@live.com	00263485620222	130 Cedar St.	Bowie	Spain	1937
8	Taniyah	Petty	redingtn@verizon.net	00213362644108	91 S. Swanson Street	Braintree	Serbia	1938
9	Curtis	Sloan	draper@sbcglobal.net	00303495087449	8262 East Temple St.	Branford	Gibraltar	1950
10	Barbara	Butler	larry@comcast.net	00322129854777	19 Hartford Dr.	Brookline	Croatia	1994