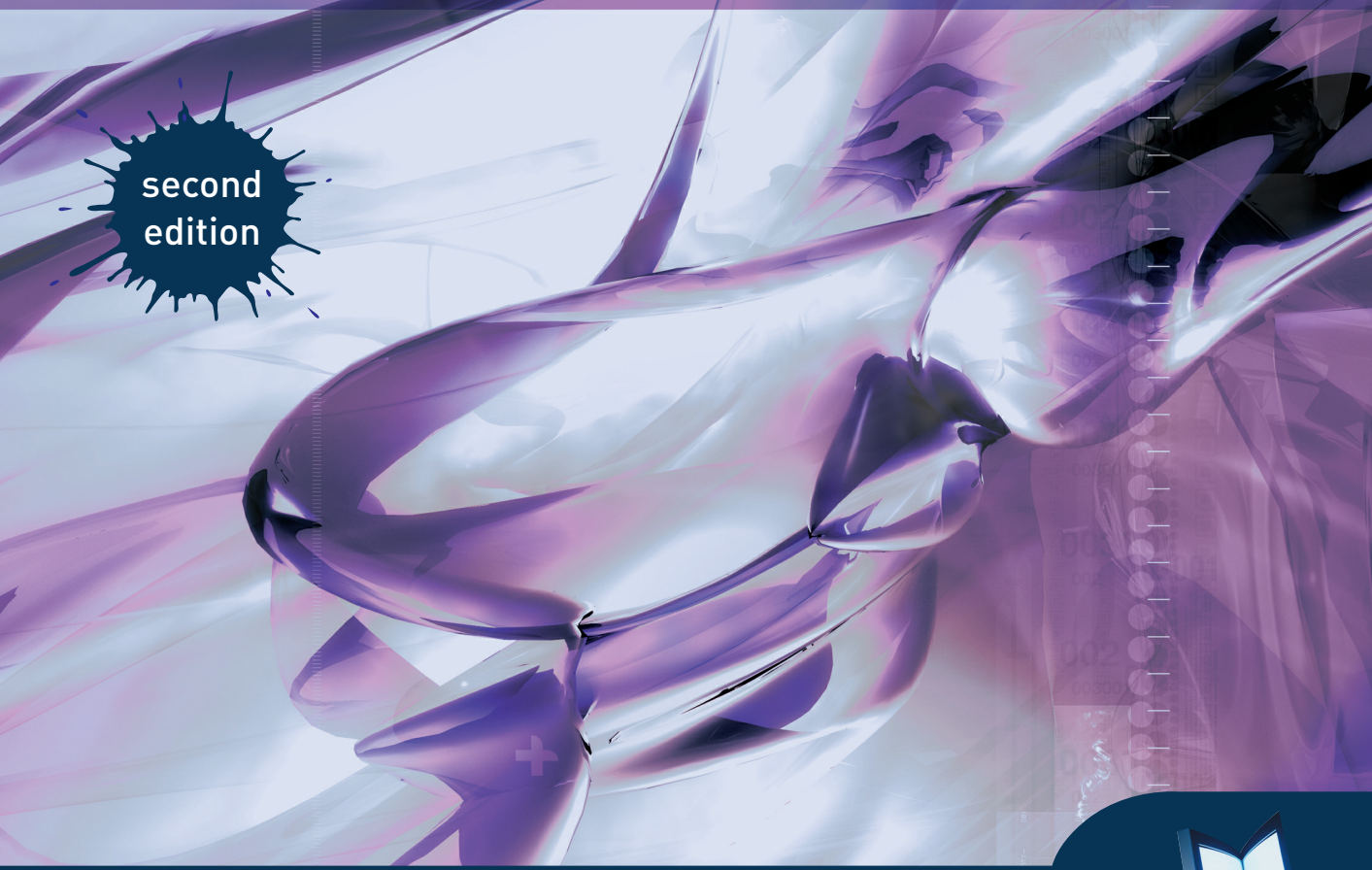
Thomas Connolly & Carolyn Begg

Database Solutions

A step-by-step guide to building databases

second edition



PEARSON
Addison Wesley

Use the online resources
for this book at
www.booksites.net

Database Solutions

A step-by-step guide to building databases

(u) List the titles of all videos in a specified category, ordered by title.

(v) List the total number of videos in each video category at each branch, ordered by branch number.

(w) List the total cost of the videos at all branches.

(x) List the total number of videos featuring each actor, ordered by actor name.

(y) List the total number of members at each branch who joined in 1999, ordered by branch number.

(z) List the total possible daily rental for videos at each branch, ordered by branch number.

Creating the systems specification for the StayHome database system

The systems specification should list all the important features for the *StayHome* database system. Examples of the types of features that should be described in the systems specification include:

- Initial database size
- Database rate of growth
- The types and average number of record searches
- Networking and shared access requirements
- Performance
- Security
- Backup and recovery
- User interface
- Legal issues.

Initial database size

(a) There are approximately 20 000 video titles and 400 000 videos for rent distributed over 100 branches. There are an average of 4000 and a maximum of 10 000 videos for rent at each branch.

(b) There are approximately 2000 staff working across all branches. There are an average of 15 and a maximum of 25 members of staff working at each branch.

(c) There are approximately 100 000 members registered across all branches. There are an average of 1000 and a maximum of 1500 members registered at each branch.

(d) There are approximately 400 000 video rentals across all branches. There are an average of 4000 and a maximum of 10 000 video rentals at each branch.

(e) There are approximately 1000 directors and 30 000 main actors in 60 000 starring roles.

(f) There are approximately 50 video suppliers and 1000 video orders.

Database rate of growth

(a) Approximately 100 new video titles and 20 copies of each video are added to the database each month.

(b) Once a copy of a video is no longer suitable for renting out (this includes those of poor visual quality, lost, or stolen), the corresponding record is deleted from the database. Approximately 100 records of videos for rent are deleted each month.

(c) Approximately 20 members of staff join and leave the company each month. The records of staff who have left the company are deleted after one year. Approximately 20 staff records are deleted each month.

(d) Approximately 1000 new members register at branches each month. If a member does not rent out a video at any time within a period of two years, his or her record is deleted. Approximately 100 member records are deleted each month.

(e) Approximately 5000 new video rentals are recorded across 100 branches each day. The details of video rentals are deleted two years after the creation of the record.

(f) Approximately 50 new video orders are placed each week. The details of video orders are destroyed two years after the creation of the record.

The types and average number of record searches

(a) Searching for the details of a branch – approximately 10 per day.

(b) Searching for the details of a member of staff at a branch – approximately 20 per day.

(c) Searching for the details of a given video – approximately 5000 per day (Sunday to Thursday), approximately 10 000 per day (Friday and Saturday). Peak workload 6–9pm daily.

(d) Searching for the details of a copy of a video – approximately 10 000 per day (Sunday to Thursday), approximately 20 000 per day (Friday and Saturday). Peak workload 6–9pm daily.

(e) Searching for the details of a specified member – approximately 100 per day.

(f) Searching for the details of a rental agreement for a member renting a video – approximately 10 000 per day (Sunday to Thursday), approximately 20 000 per day (Friday and Saturday). Peak workload 6–9pm daily.

Networking and shared access requirements

(a) All branches should be securely networked to a centralized database located at the company's HQ in Seattle.

(b) The system should allow for at least three people concurrently accessing the system from each branch. Consideration needs to be given to the licensing requirements for this number of concurrent accesses.

Performance

(a) During opening hours but not during peak periods expect less than 1 second response for all single record searches. During peak periods (6–9pm daily) expect less than 5 second response for all single record searches.

(b) During opening hours but not during peak periods expect less than 5 second response for all multiple record searches. During peak periods (6–9pm daily) expect less than 10 second response for all multiple record searches.

(c) During opening hours but not during peak periods expect less than 1 second response for all updates/saves. During peak periods (6–9pm daily) expect less than 5 second response for all updates/saves.

Security

(a) The database should be password protected.

(b) Each member of staff should be assigned database access privileges appropriate to a particular user view, namely Director, Manager, Supervisor, Assistant, or Buyer.

(c) Staff should see only the data necessary to do their job in a form that suits what they're doing.

Backup and recovery

The database should be backed up each day at 12 midnight.

User interface

The user interface should be menu-driven. Online help should be easy to locate and access.

Legal issues

Each country has laws that govern the way that the computerized storage of personal data is handled. As the *StayHome* database holds data on staff and members, any legal issues that must be complied with should be investigated and implemented.

6.4.5 The *StayHome* case study – database design

In this chapter, we demonstrated the creation of the users' requirements specification for the Branch user views and the systems specification for the *StayHome* database system. These documents are the source of information for the next stage of the lifecycle called database design. In Chapters 9, 10, and 12 to 16, we'll provide a step-by-step methodology for database design, and we'll use the documents created in this chapter to demonstrate the methodology in practice.

For those of you interested in developing more complex multi-user-view database systems, we'll demonstrate how the view integration approach works in practice in Appendix C using the branch and business user views of *StayHome*.

Chapter summary

- ✓ **Fact-finding** is the formal process of using techniques such as interviews and questionnaires to collect facts about systems, requirements, and preferences.
- ✓ Fact-finding is particularly crucial to the early stages of the database system development lifecycle, including the database planning, system definition, and requirements collection and analysis stages.
- ✓ The five most common fact-finding techniques are examining documentation, interviewing, observing the business in operation, research, and questionnaires.
- ✓ The first step in the **database planning stage** is to define clearly the mission statement and mission objectives for the database project. The **mission statement** defines the major aims of the database system. Each **mission objective** should identify a particular task that the database must support.
- ✓ The purpose of the **system definition stage** is to define the boundaries and user views of the database system.
- ✓ There are two main documents created during the **requirements collection and analysis stage**, namely the users' requirements specification and the systems specification.
- ✓ The **users' requirements specification** describes in detail the data to be held in the database and how the data is to be used.
- ✓ The **systems specification** describes any features to be included in the database system such as the required performance and the levels of security.

Review questions

6.1 Briefly describe what the process of fact-finding attempts to achieve for a database developer.

6.2 Describe how fact-finding is used throughout the stages of the database system development lifecycle.

6.3 For each stage of the database system development lifecycle identify examples of the facts captured and the documentation produced.

6.4 A database developer normally uses several fact-finding techniques during a single database project. The five most commonly used techniques are examining documentation, interviewing, observing the business in operation,

conducting research, and using questionnaires. Describe each fact-finding technique and identify the advantages and disadvantages of each.

6.5 Describe the purpose of defining a mission statement and mission objectives for a database system.

6.6 What is the purpose of the systems definition stage?

6.7 How do the contents of a users' requirements specification differ from a systems specification?

6.8 Describe one approach to deciding whether to use centralized, view integration, or a combination of both when developing a database system for multiple user views.

Chapter 7

Entity–Relationship modeling

In this chapter you will learn:

- How to use ER modeling in database design.
- The basic concepts of an ER model called entities, relationships, and attributes.
- A diagrammatic technique for displaying an ER model.
- How to identify and solve connection traps in an ER model.

Database system development lifecycle discussed in Chapter 4

In Chapter 6, you learned about techniques for gathering and capturing information about what the users require of the database system. Once the requirements collection and analysis stage of the database system development lifecycle is complete and you have documented the requirements for the database system, you are now ready to begin database design.

One of the most difficult aspects of database design is the fact that designers, programmers, and end-users tend to view data and its use in different ways. Unfortunately, unless we can gain a common understanding that reflects how the organization operates, the design we produce will fail to meet the users' requirements. To ensure that we get a precise understanding of the nature of the data and how the organization uses it, we need to have a model for communication that is non-technical and free of ambiguities. The Entity–Relationship (ER) model is one such example. Since the introduction of ER modeling in 1976, the model has been extended to include additional enhanced modeling concepts. We cover the basic ER concepts in this chapter and introduce some of the more popular enhanced ER concepts in Chapter 11.