

Data: The New Corporate Resource

The development of database management systems, as well as the development of modern computers, came about as a result of society's recognition of the crucial importance of storing, managing, and retrieving its rapidly expanding volumes of business data. To understand how far we have come in this regard, it is important to know where we began and how the concept of managing data has developed. This chapter begins with the historical background of the storage and uses of data and then continues with a discussion of the importance of data to the modern corporation.

OBJECTIVES

Explain why humankind's interest in data dates back to ancient times.
Describe how data needs have historically driven many information technology developments.
Describe the evolution of data storage media during the last century.
Relate the idea of data as a corporate resource that can be used to gain a competitive advantage to the development of the database management systems environment.

CHAPTER OUTLINE

Introduction
The History of Data
 The Origins of Data
 Data Through the Ages
 Early Data Problems Spawn
 Calculating Devices
 Swamped with Data
 Modern Data Storage Media
Data in Today's Information Systems Environment
Using Data for Competitive Advantage
Problems in Storing and Accessing Data
Data as a Corporate Resource
The Database Environment
Summary

Introduction

What a fascinating world we live in today! Technological advances are all around us in virtually every aspect of our daily lives. From cellular telephones to satellite television to advanced aircraft to modern medicine to computers—especially computers—high tech is with us wherever we look. Businesses of every description and size rely on computers and the information systems they support to a degree that would have been

unimaginable just a few short years ago. Businesses routinely use automated manufacturing and inventory-control techniques, automated financial transaction procedures, and high-tech marketing tools. As consumers, we take for granted being able to call our banks, insurance companies, and department stores to instantly get up-to-the-minute information on our accounts. And everyone, businesses and consumers alike, has come to rely on the Internet for instant worldwide communications. Beneath the surface, the foundation for all of this activity is data: the stored facts that we need to manage all of our human endeavors.

This book is about **data**. It's about how to think about data in a highly organized and deliberate way. It's about how to store data efficiently and how to retrieve it effectively. It's about ways of managing data so that the exact data that we need will be there when we need it. It's about the concept of assembling data into a highly organized collection called a "**database**" and about the sophisticated software known as a "**database management system**" that controls the database and oversees the **database environment**. It's about the various approaches people have taken to database management and about the roles people have assumed in the database environment. We will see many real-world examples of data usage throughout this book.

Computers came into existence because we needed help in processing and using the massive amounts of data we have been accumulating. Is the converse true? Could data exist without computers? The answer to this question is a resounding "yes." In fact, data has existed for thousands of years in some very interesting, if by today's standards crude, forms. Furthermore, some very key points in the history of the development of computing devices were driven, not by any inspiration about computing for computing's sake, but by a real need to efficiently handle a pesky data management problem. Let's begin by tracing some of these historical milestones in the evolution of data and data management.

The History of Data

The Origins of Data

What is data? To start, what is a single piece of data? A single piece of data is a single fact about something we are interested in. Think about the world around you, about your environment. In any environment, there are things that are important to you and there are facts about those things that are worth remembering. A "thing" can be an obvious object like an automobile or a piece of furniture. But the concept of an object is broad enough to include a person, an organization like a company, or an event that took place such as a particular meeting. A fact can be any characteristic of an object. In a university environment, it may be the fact that student Gloria Thomas has completed 96 credits; or it may be the fact that Professor Howard Gold graduated from Ohio State University; or it may be the fact that English 349 is being held in Room 830 of Alumni Hall. In a commercial environment, it may be the fact that employee John Baker's employee number is 137; or it may be the fact that one of a company's suppliers, the Superior Products Co., is located in Chicago; or it may be the fact that the refrigerator with serial number 958304 was manufactured on November 5, 2004.

Actually, people have been interested in data for at least the past 12,000 years. While today we often associate the concept of data with the computer, historically there have been many more primitive methods of data storage and handling.

In the ancient Middle East, shepherds kept track of their flocks with pebbles, Figure 1.1. As each sheep left its pen to graze, the shepherd placed one pebble in a

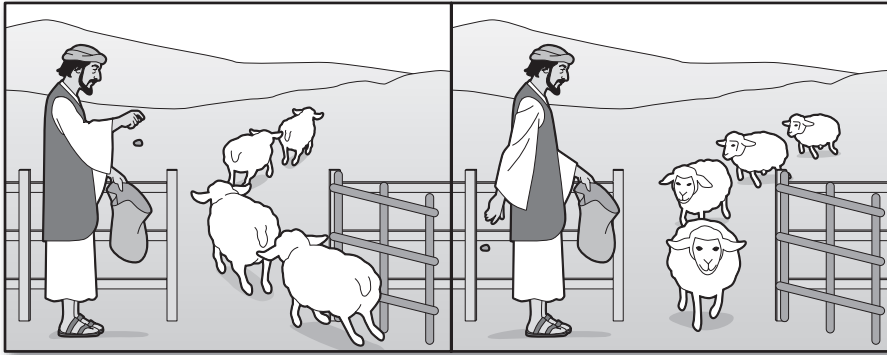


FIGURE 1.1 Shepherd using pebbles to keep track of sheep

small sack. When all of the sheep had left, the shepherd had a record of how many sheep were out grazing. When the sheep returned, the shepherd discarded one pebble for each animal, and if there were more pebbles than sheep, he knew that some of his sheep still hadn't returned or were missing. This is, indeed, a primitive but legitimate example of data storage and retrieval. What is important to realize about this example is that the count of the number of sheep going out and coming back in was all that the shepherd cared about in his "business environment" and that his primitive data storage and retrieval system satisfied his needs.

Excavations in the Zagros region of Iran, dated to 8500 B.C., have unearthed clay **tokens** or counters that we think were used for record keeping in primitive forms of accounting. Such tokens have been found at sites from present-day Turkey to Pakistan and as far afield as the present-day Khartoum in Sudan, dating as long ago as 7000 B.C. By 3000 B.C., in the present-day city of Susa in Iran, the use of such tokens had reached a greater level of sophistication. Tokens with special markings on them, Figure 1.2, were sealed in hollow clay vessels that accompanied commercial goods in transit. These primitive bills of lading certified the contents of the shipments. The tokens represented the quantity of goods being shipped and, obviously, could not be

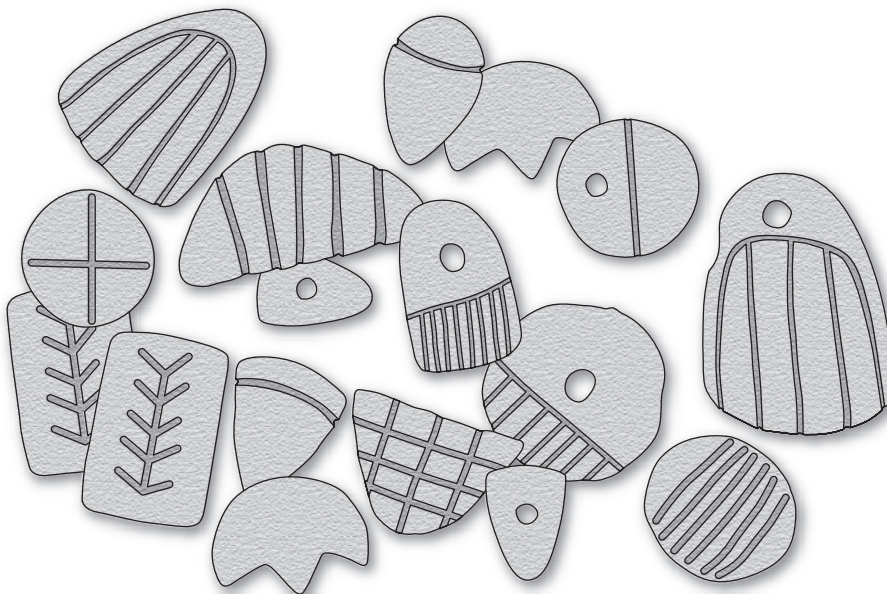


FIGURE 1.2 Ancient clay tokens used to record goods in transit

tampered with without the clay vessel being broken open. Inscriptions on the outside of the vessels and the seals of the parties involved provided a further record. The external inscriptions included such words or concepts as “deposited,” “transferred,” and “removed.”

At about the same time that the Susa culture existed, people in the city-state of Uruk in Sumeria kept records in clay texts. With pictographs, numerals, and ideographs, they described land sales and business transactions involving bread, beer, sheep, cattle, and clothing. Other Neolithic means of record keeping included storing tallies as cuts and notches in wooden sticks and as knots in rope. The former continued in use in England as late as the medieval period; South American Indians used the latter.

Data Through the Ages

As in Susa and Uruk, much of the very early interest in data can be traced to the rise of cities. Simple subsistence hunting, gathering, and, later, farming had only limited use for the concept of data. But when people live in cities they tend to specialize in the goods and services they produce. They become dependent on one another, **bartering** and using money to trade these goods and services for mutual survival. This trade encouraged **record keeping**—the recording of data—to track how much someone has produced and what it can be bartered or sold for.

As time went on, more and different kinds of data and records were kept. These included calendars, census data, surveys, land ownership records, marriage records, records of church contributions, and family trees, Figure 1.3. Increasingly sophisticated merchants had to keep track of inventories, shipments, and wage payments in addition to production data. Also, as farming went beyond the subsistence level and progressed to the feudal manor stage, there was a need to keep data on the amount of produce to consume, to barter with, and to keep as seed for the following year.

The Crusades took place from the late eleventh to the late thirteenth centuries. One side effect of the Crusades was a broader view of the world on the part of the Europeans, with an accompanying increase in interest in trade. A common method of trade in that era was the establishment of temporary partnerships among merchants, ship



FIGURE 1.3 New types of data with the advance of civilization

captains, and owners to facilitate commercial voyages. This increased level of commercial sophistication brought with it another round of increasingly complex record keeping, specifically, *double-entry bookkeeping*.

Double-entry bookkeeping originated in the trading centers of fourteenth-century Italy. The earliest known example, from a merchant in Genoa, dates to the year 1340. Its use gradually spread, but it was not until 1494, in Venice (about 25 years after Venice's first movable type printing press came into use), that a Franciscan monk named Luca Pacioli published his “Summa de Arithmetica, Geometrica, Proportioni et Proportionalita,” a work important in spreading the use of double-entry bookkeeping. Of course, as a separate issue, the increasing use of paper and the printing press furthered the advance of record keeping as well.

As the dominance of the Italian merchants declined, other countries became more active in trade and thus in data and record keeping. Furthermore, as the use of temporary trading partnerships declined and more stable long-term mercantile organizations were established, other types of data became necessary. For example, annual as opposed to venture-by-venture statements of profit and loss were needed. In 1673, the “Code of Commerce” in France required every businessman to draw up a **balance sheet** every two years. Thus the data had to be periodically accumulated for reporting purposes.

Early Data Problems Spawn Calculating Devices

It was also in the seventeenth century that data began to prompt people to take an interest in devices that could “automatically” *process their data*, if only in a rudimentary way. Blaise Pascal produced one of the earliest and best known such devices in France in the 1640s, reputedly to help his father track the data associated with his job as a tax collector, Figure 1.4. This was a small box containing interlocking gears that was capable of doing addition and subtraction. In fact, it was the forerunner of today's mechanical automobile odometers.

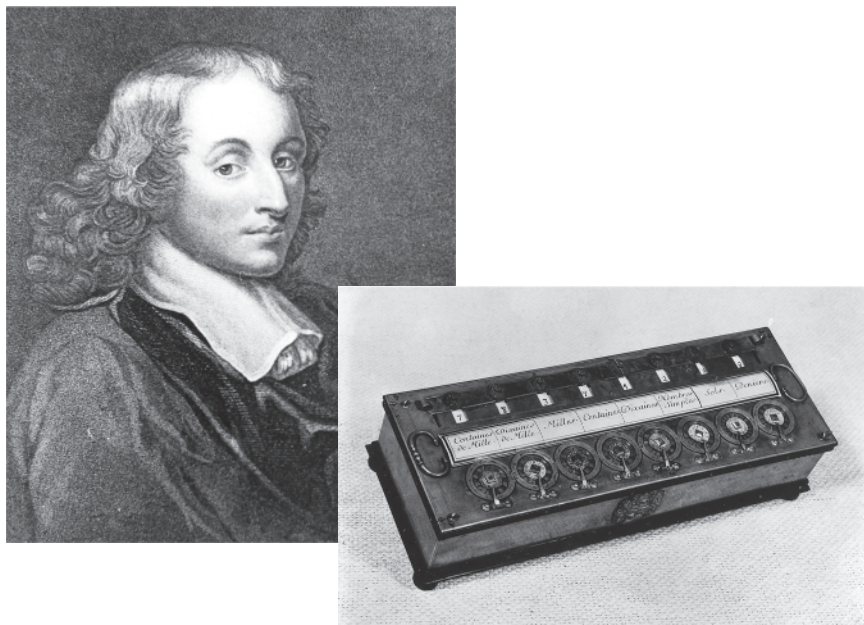


FIGURE 1.4 Blaise Pascal and his adding machine

Courtesy of IBM Archives

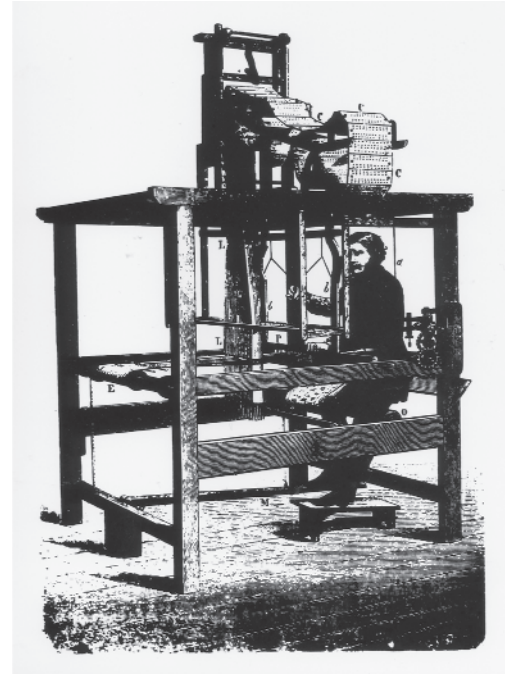


FIGURE 1.5 The Jacquard loom recorded patterns in punched cards

Courtesy of IBM Archives

In 1805, Joseph Marie Jacquard of France invented a device that automatically reproduced patterns used in textile weaving. The heart of the device was a series of cards with holes punched in them; the holes allowed strands of material to be interwoven in a sequence that produced the desired pattern, Figure 1.5. While Jacquard's loom wasn't a calculating device as such, his method of storing fabric patterns, a form of graphic data, as holes in **punched cards** was a very clever means of **data storage** that would have great importance for computing devices to follow. Charles Babbage, a nineteenth-century English mathematician and inventor, picked up Jacquard's concept of storing data in punched cards. Beginning in 1833, Babbage began to think about an invention that he called the "Analytical Engine." Although he never completed it (the state of the art of machinery was not developed enough), included in its design were many of the principles of modern computers. The Analytical Engine was to consist of a "store" for holding data items and a "mill" for operating upon them. Babbage was very impressed by Jacquard's work with punched cards. In fact, the Analytical Engine was to be able to store calculation instructions in punched cards. These would be fed into the machine together with punched cards containing data, would operate on that data, and would produce the desired result.

Swamped with Data

In the late 1800s, an enormous (for that time) data storage and retrieval problem and greatly improved machining technology ushered in the era of modern **information processing**. The 1880 U.S. **Census** took about seven years to compile by hand. With a rapidly expanding population fueled by massive immigration, it was estimated that with the same manual techniques, the compilation of the 1890 census would not be completed until after the 1900 census data had begun to be collected. The solution to processing census data was provided by a government engineer named Herman Hollerith. Basing his work on Jacquard's punched-card concept, he arranged to have

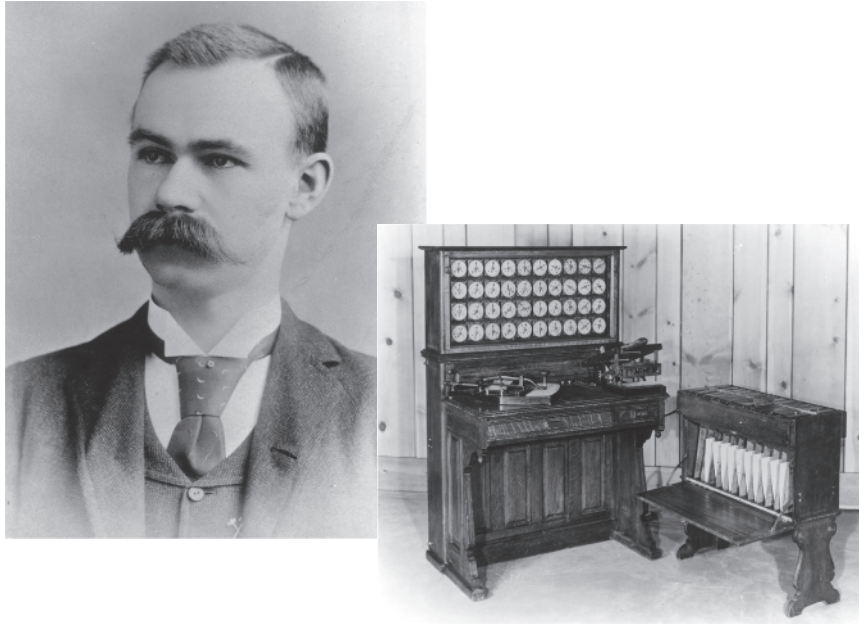


FIGURE 1.6 Herman Hollerith and his tabulator/sorter, circa 1890.

Bell, C. M. / Wikimedia Commons / Public domain; Courtesy of IBM Archives.

the census data stored in punched cards. He built devices to punch the holes into cards and devices to sort the cards, Figure 1.6. Wire brushes touching the cards completed circuits when they came across the holes and advanced counters. The equipment came to be classified as “electromechanical,” “electro” because it was powered by electricity and “mechanical” because the electricity powered mechanical counters that tabulated the data. By using Hollerith’s **electromechanical equipment**, the total population count of the 1890 census was completed a month after all the data was in. The complete set of tabulations, including data on questions that had never before even been practical to ask, took two years to complete. In 1896, Hollerith formed the Tabulating Machine Company to produce and commercially market his devices. That company, combined with several others, eventually formed what is today the International Business Machines Corporation (IBM).

Toward the turn of the century, immigrants kept coming and the U.S. population kept expanding. The Census Bureau, while using Hollerith’s equipment, continued experimenting on its own to produce even more advanced data-tabulating machinery. One of its engineers, James Powers, developed devices to automatically feed cards into the equipment and automatically print results. In 1911, he formed the Powers Tabulating Machine Company, which eventually formed the basis for the UNIVAC division of the Sperry Corporation, which eventually became the Unisys Corporation.

From the days of Hollerith and Powers through the 1940s, commercial data processing was performed on a variety of electromechanical punched-card-based devices. They included calculators, punches, sorters, collators, and printers. The data was stored in punched cards, while the processing instructions were implemented as collections of wires plugged into specially designed boards that in turn were inserted into slots in the electromechanical devices. Indeed, electromechanical equipment overlapped with **electronic computers**, which were introduced commercially in the mid-1950s.

In fact, the introduction of electronic computers in the mid-1950s coincided with a tremendous boom in economic development that raised the level of data storage and retrieval requirements another notch. This was a time of rapid commercial growth in

the post–World War II United States as well as the rebuilding of Europe and the Far East. From this time onward, the furious pace of new data storage and retrieval requirements with more and more commercial functions and procedures were automated and the technological advances in computing devices has been one big blur. From this point on, it would be virtually impossible to tie advances in computing devices to specific, landmark data storage and retrieval needs. And there is no need to try to do so.

Modern Data Storage Media

Paralleling the growth of equipment to process data was the development of new media on which to store the data. The earliest form of modern data storage was **punched paper tape**, which was introduced in the 1870s and 1880s in conjunction with early teletype equipment. Of course we’ve already seen that Hollerith in the 1890s and Powers in the early 1900s used punched cards as a storage medium. In fact, punched cards were the only data storage medium used in the increasingly sophisticated electromechanical accounting machines of the 1920s, 1930s, and 1940s. They were still used extensively in the early computers of the 1950s and 1960s and could even be found well into the 1970s in smaller information systems installations, to a progressively reduced degree.

Your Turn

1.1 The Development of Data

The need to organize and store data has arisen many times and in many ways throughout history. In addition to the data-focused events presented in this chapter, what other historical events can you think of that have made people think about organizing and storing data? As a hint, you might think about the exploration and conquest of new lands, wars, changes in type of governments such

as the introduction of democracy, and the implications of new inventions such as trains, printing presses, and electricity.

Question:

Develop a timeline showing several historical events that influenced the need to organize and store data. Include a few noted in this chapter as well as a few that you can think of independently.

The middle to late 1930s saw the beginning of the era of erasable magnetic storage media, with Bell Laboratories experimenting with magnetic tape for sound storage. By the late 1940s, there was early work on the use of magnetic tape for recording data. By 1950, several companies, including RCA and Raytheon, were developing the **magnetic tape** concept for commercial use. Both UNIVAC and Raytheon offered commercially available magnetic tape units in 1952, followed by IBM in 1953, Figure 1.7. During the mid-1950s and into the mid-1960s, magnetic tape gradually became the dominant data-storage medium in computers. Magnetic tape technology has been continually improved since then and is still in limited use today, particularly for archived data.

The original concept that eventually grew into the **magnetic disk** actually began to be developed at MIT in the late 1930s and early 1940s. By the early 1950s, several companies including UNIVAC, IBM, and Control Data had developed prototypes of **magnetic “drums”** that were the forerunners of magnetic disk technology. In 1953, IBM began work on its 305 RAMAC (Random Access Memory Accounting Machine)

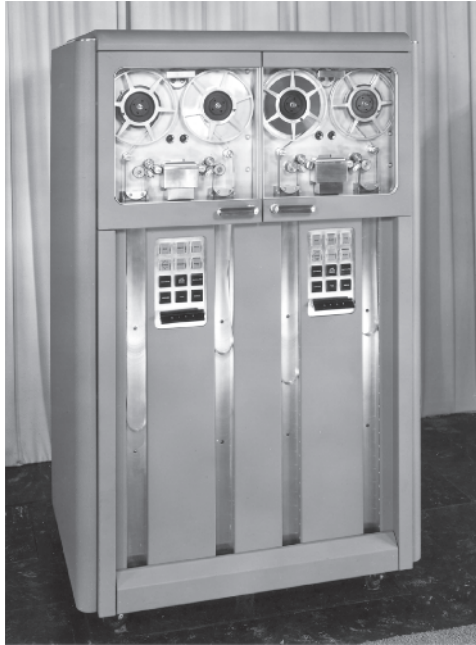


FIGURE 1.7 Early magnetic tape drive, circa 1953

Courtesy of IBM Archives

fixed disk storage device. By 1954, there was a multi-platter version, which became commercially available in 1956, Figure 1.8.

During the mid-1960s, a massive conversion from tape to magnetic disk as the pre-eminent data storage medium began and disk storage is still the data storage medium of choice today. After the early fixed disks, the disk storage environment became geared toward the removable disk-pack philosophy, with a dozen or more packs being juggled on and off a single drive as a common ratio. But, with the increasingly tighter environmental controls that fixed disks permitted, more data per square inch (or square



FIGURE 1.8 IBM RAMAC disk storage device, circa 1956

Courtesy of IBM Archives

centimeter) could be stored on fixed disk devices. Eventually, the **disk drives** on mainframes and servers, as well as the fixed disks or “hard drives” of PCs, all became non-removable, sealed units. But the removable disk concept stayed with us a while in the form of PC diskettes, the Iomega Corp.’s Zip Disks, and external hard drives that can be easily moved from one computer to another simply by plugging them into a USB port. These have been joined by the laser-based, optical technology **compact disk** (CD), introduced as a data storage medium in 1985. Originally, data could be recorded on these CDs only at the factory and once created, they were non-erasable. Later, data could be recorded on them, erased, and re-recorded in a standard PC. Eventually, solid-state technology became so miniaturized and inexpensive that a popular option for removable media developed as the **flash drive**. Finally, as solid-state technology continues on the path of being more compact and cheaper, rotating disk technology is gradually being replaced by “solid-state disks” (SSDs).

Data in Today’s Information Systems Environment

Using Data for Competitive Advantage

Today’s computers are technological marvels. Their speeds, compactness, ease of use, price as related to capability, and, yes, their data storage capacities are truly amazing. And yet, our fundamental interest in computers is the same as that of the ancient Middle Eastern shepherds in their pebbles and sacks: they are the vehicles we need to store and utilize the data that is important to us in our environment.

Indeed, data has become indispensable in every kind of modern business and government organization. Data, the applications that process the data, and the computers on which the applications run are fundamental to every aspect of every kind of endeavor. When speaking of **corporate resources**, people used to list such items as capital, plant and equipment, inventory, personnel, and patents. Today, any such list of corporate resources must include the corporation’s data. It has even been suggested that data is the most important corporate resource because it describes all of the others.

Data can provide a crucial **competitive advantage** for a company. We routinely speak of data and the information derived from it as competitive weapons in hotly contested industries. For example, FedEx had a significant competitive advantage when it first provided access to its package tracking data on its Web site. Then, once one company in an industry develops a new application that takes advantage of its data, the other companies in the industry are forced to match it to remain competitive. This cycle continually moves the use of data to ever-higher levels, making it an ever more important corporate resource than before. Examples of this abound. Banks give their customers online access to their accounts. Package shipping companies provide up-to-the-minute information on the whereabouts of a package. Retailers send manufacturers product sales data that the manufacturers use to adjust inventories and production cycles. Manufacturers automatically send their parts suppliers inventory data and expect the suppliers to use the data to keep a steady stream of parts flowing.

Your Turn

1.2 Data as a Competitive Weapon

Think about a company with which you or your family regularly does business. This might be a supermarket, a department store, or a pharmacy, as examples. What kind of data do you think they collect about their suppliers, their inventory, their sales, and their customers? What kind of data do you think they should collect and how do you

think they might be able to use it to gain a competitive advantage?

Question:

Choose one of the companies that you or your family does business with and develop a plan for the kinds of data it might collect and the ways in which it might use the data to gain a business advantage over its competitors.

Problems in Storing and Accessing Data

But being able to store and provide efficient access to a company's data while also maintaining its accuracy so that it can be used to competitive advantage is anything but simple. In fact, several factors make it a major challenge. First and foremost, the volume or amount of data that companies have is massive and growing all the time. At the time of this writing, Walmart estimates that its data warehouse (a type of database we will explore later) alone contains 30 petabytes (30 quadrillion bytes) of data and is constantly growing. The number of people who want access to the data is also growing: at one time, only a select group of a company's own employees were concerned with retrieving its data, but this has changed. Now, not only do vastly more of a company's employees demand access to the company's data but also so do the company's customers and trading partners. All major banks today give their depositors Internet access to their accounts. Increasingly, tightly linked "supply chains" require that companies provide other companies, such as their suppliers and customers, with access to their data. The combination of huge volumes of data and large numbers of people demanding access to it has created a major performance challenge. How do you sift through so much data for so many people and give them the data that they want in an acceptably small amount of time? How much patience would you have with an insurance company that kept you on the phone for five or ten minutes while it retrieved claim data about which you had a question? Of course, the tremendous advances in computer hardware, including data storage hardware, have helped—indeed, it would have been impossible to have gone as far as we have in information systems without them. But as the hardware continues to improve, the volumes of data and the number of people who want access to it also increase, making it a continuing struggle to provide them with acceptable response times.

Other factors that enter into data storage and retrieval include data security, data privacy, and backup and recovery. Data security involves a company protecting its data from theft, malicious destruction, deliberate attempts to make phony changes to the data (e.g. someone trying to increase his own bank account balance), and even accidental damage by the company's own employees. Data privacy implies assuring that even employees who normally have access to the company's data (much less outsiders) are given access only to the specific data they need in their work. Put another

way, sensitive data such as employee salary data and personal customer data should be accessible only by employees whose job functions require it. Backup and recovery means the ability to reconstruct data if it is lost or corrupted, say in a hardware failure. The extreme case of backup and recovery is known as disaster recovery when an information system is destroyed by fire, a hurricane, or other calamity. Today, data in “the cloud” has changed the way that we think about some of these issues. We will explore this later in this book.

Another whole dimension involves maintaining the accuracy of a company’s data. Historically, and in many cases even today, the same data is stored several, sometimes many, times within a company’s information system. Why does this happen? For several reasons. Many companies are simply not organized to share data among multiple applications. Every time a new application is written, new data files are created to store its data. As recently as the early 1990s, I spoke to a database administration manager (more on this type of position later) in the securities industry who told me that one of the reasons he was hired was to reduce duplicate data appearing in as many as 60–70 files! Eventually, the advent of Enterprise Resource Planning (ERP) systems, which we will discuss later in this book, helped alleviate this problem.

Furthermore, depending on how database files are designed, data can even be duplicated within a single file. We will explore this issue much more in this book, but for now, suffice it to say that duplicate data, either in multiple files or in a single file, can cause major data accuracy problems.

Data as a Corporate Resource

Every corporate resource must be carefully managed so that the company can keep track of it, protect it, and distribute it to those people and purposes in the company that need it. Furthermore, public companies have a responsibility to their shareholders to competently manage the company’s assets. Can you imagine a company’s money just sort of out there somewhere without being carefully managed? In fact, the chief financial officer with a staff of accountants and financial professionals is responsible for the money, with outside accounting firms providing independent audits of it. Typically vice presidents of personnel and their staffs are responsible for the administrative functions necessary to manage employee affairs. Production managers at various levels are responsible for parts inventories, and so on. Data is no exception.

But data may just be the most difficult corporate resource to manage. In data, we have a resource of tremendous volume, billions, trillions, and more individual pieces of data, each piece of which is different from the next. And it has the characteristic that much of it is in a state of change at any one time. It’s not as if we’re talking about managing a company’s employees. Even the largest companies have only a few hundred thousand of them, and they don’t change all that frequently. Or the money a company has: sure, there is a lot of it, but it’s all the same in the sense that a dollar that goes to payroll is the same kind of dollar that goes to paying a supplier for raw materials.

As far back as the early to mid-1960s, barely ten years after the introduction of commercially viable electronic computers, some forward-looking companies began to realize that storing each application’s data separately, in simple files, was becoming problematic and would not work in the long run, for just the reasons that we’ve talked about: the increasing volumes of data (even way back then), the increasing demand for data access, the need for data security, privacy, backup, and recovery, and the desire to share data and cut down on data redundancy. Several things were becoming clear. The task was going to require both a new kind of software to help manage the data

and progressively faster hardware to keep up with the increasing volumes of data and data access demands. And data-management specialists would have to be developed, educated, and made responsible for managing the data as a corporate resource.

Out of this need was born a new kind of software, the database management system (DBMS), and a new category of personnel, with titles like database administrator and data management specialist. And yes, hardware has progressively gotten faster and cheaper for the performance it provides. The integration of these advances adds up to much more than the simple sum of their parts. They add up to the database environment.

The Database Environment

Back in the early 1960s, the emphasis in what was then called data processing was on programming. Data was little more than a necessary afterthought in the application development process and in running the data-processing installation. There was a good reason for this. By today's standards, the rudimentary computers of the time had very small main memories and very simplistic operating systems. Even relatively basic application programs had to be shoehorned into main memory using low-level programming techniques and a lot of cleverness. But then, as we progressed further into the 1960s and beyond, two things happened simultaneously that made this picture change forever. One was that main memories became progressively larger and cheaper and operating systems became much more powerful. Plus, computers progressively became faster and cheaper on a price/performance basis. All these changes had the effect of permitting the use of higher-level programming languages that were easier for a larger number of personnel to use, allowing at least some of the emphasis to shift elsewhere. Well, nature hates a vacuum, and at the same time that all of this was happening, companies started becoming aware of the value of thinking of data as a corporate resource and using it as a competitive weapon.

The result was the development of database management systems (DBMS) software and the creation of the "database environment." Supported by ever-improved hardware and specialized database personnel, the database environment is designed largely to correct all the problems of the non-database environment. It encourages data sharing and the control of data redundancy with important improvements in data accuracy. It permits storage of vast volumes of data with acceptable access and response times for database queries. And it provides the tools to control data security, data privacy, and backup and recovery.

This book is a straightforward introduction to the fundamentals of database in the current information systems environment. It is designed to teach you the important concepts of the database approach and also to teach you specific skills, such as how to design relational databases, how to improve database performance, and how to retrieve data from relational databases using the SQL language. In addition, as you proceed through the book you will explore such topics as entity-relationship diagrams, object-oriented database, database administration, distributed database, data warehousing, NoSQL database management, blockchain, data in the cloud, and data-focused applications.

We start with the basics of database and take a step-by-step approach to exploring all the various components of the database environment. Each chapter progressively adds more to an understanding of both the technical and managerial aspects of the field. Database is a very powerful concept. Overall it provides ingenious solutions to a set of very difficult problems. As a result, it tends to be a multifaceted and complex subject that can appear difficult when one attempts to swallow it in one gulp. But

database is approachable and understandable if we proceed carefully, cautiously, and progressively step by step. And this is an understanding that no one involved in information systems can afford to be without.

Summary

Recognition of the commercial importance of data, of storing it, and of retrieving it can be traced back to ancient times. As trade routes lengthened and cities grew larger, data became increasingly important. Eventually, the importance of data led to the development of electromechanical calculating devices and then to modern electronic computers, complete with magnetic and optical disk-based data storage media.

While the use of data has given many companies a competitive advantage in their industries, the storage and retrieval of today's vast amounts of data holds many challenges. These include speedy retrieval of data when many people try to access the data at the same time, maintaining the accuracy of the data, the issue of data security, and the ability to recover the data if it is lost.

The recognition that data is a critical corporate resource and that managing data is a complex task has led to the development and continuing refinement of specialized software known as database management systems, the subject of this book.

Key Terms

Balance sheet	Data	Double-entry	Magnetic disk
Barter	Database	bookkeeping	Magnetic drum
Calculating devices	Database environment	Electromechanical	Magnetic tape
Census	Database management	equipment	Punched cards
Compact disk	system	Electronic computer	Punched paper tape
Competitive advantage	Data storage	Flash drive	Record keeping
Corporate resource	Disk drive	Information processing	Token

Questions

1. What did the Middle Eastern shepherds' pebbles and sacks, Pascal's calculating device, and Hollerith's punched-card devices all have in common?
2. What did the growth of cities have to do with the need for data?
3. What did the growth of trade have to do with the need for data?
4. What did Jacquard's textile weaving device have to do with the development of data?
5. Choose what you believe to be the:
 - a. One most important
 - b. Two most important
 - c. Three most important landmark events in the history of data. Defend your choices.
6. Do you think that computing devices would have been developed even if specific data needs had not come along? Why or why not?
7. What did the need for data among ancient Middle Eastern shepherds have in common with the need for data of modern corporations?
8. List several problems in storing and accessing data in today's large corporations. Which do you think is the most important? Why?
9. How important an issue do you think data accuracy is? Explain.
10. How important a corporate resource is data compared to other corporate resources? Explain.
11. What factors led to the development of database management systems?

Exercises

1. Draw a timeline showing the landmark events in the history of data from ancient times to the present day. Do not include the development of computing devices in this timeline.
2. Draw a timeline for the last four hundred years comparing landmark events in the history of data to landmark events in the development of computing devices.
3. Draw a timeline for the last two hundred years comparing the development of computing devices to the development of data storage media.
4. Invent a fictitious company in *one* of the following industries and list several ways in which the company can use data to gain a competitive advantage.
 - a. Banking
 - b. Insurance
 - c. Manufacturing
 - d. Airlines
5. Invent a fictitious company in *one* of the following industries and describe the relationship between data as a corporate resource and the company's other corporate resources.
 - a. Banking
 - b. Insurance
 - c. Manufacturing
 - d. Airline

Minicases

1. Worldwide, vacation cruises on increasingly larger ships have been steadily growing in popularity. People like the all-inclusive price for food, room, and entertainment, the variety of shipboard activities, and the ability to unpack just once and still visit several different places. The first of the two minicases used throughout this book is the story of Happy Cruise Lines. Happy Cruise Lines has several ships and operates (begins its cruises) from a number of ports. It has a variety of vacation cruise itineraries, each involving several ports of call. The company wants to keep track of both its past and future cruises and of the passengers who sailed on the former and are booked on the latter. Actually, you can think of a cruise line as simply a somewhat specialized instance of any passenger transportation company, including airlines, trains, and buses. Beyond that, a cruise line is, after all, a business and like any other business of any kind it must be concerned about its finances, employees, equipment, and so forth.
 - a. Using this introductory description of (and hints about) Happy Cruise Lines, make a list of the things in Happy Cruise Lines' business environment about which you think the company would want to maintain data. Do some or all of these qualify as "corporate resources"? Explain.
 - b. Develop some ideas about how the data you identified in part a above can be used by Happy Cruise Lines to gain a competitive advantage over other cruise lines.
2. Sports are universally enjoyed around the globe. Whether the sport is a team or individual sport, whether a person is a participant or a spectator, and whether the sport is played at the amateur or professional level, one way or another this kind of activity can be enjoyed by people of all ages and interests. Furthermore, professional sports today are a big business involving very large sums of money. And so, the second of the two minicases to be used throughout this book is the story of the professional Super Baseball League. Like any sports league, the Super Baseball League wants to maintain information about its teams, coaches, players, and equipment, among other things. If you are not particularly familiar with baseball or simply prefer another sport, bear in mind that most of the issues that will come up in this minicase easily translate to any team sport at the amateur, college, or professional levels. After all, all team sports have teams, coaches, players, fans, equipment, and so forth. When specialized equipment or other baseball-specific items come up, we will explain them.
 - a. Using this introductory description of (and hints about) the Super Baseball League, list the things in the Super Baseball League's business environment about which you think the league would want to maintain data. Do some or all of these qualify as "corporate resources," where the term is broadened to include the resources of a sports league? Explain.
 - b. Develop some ideas about how the data that you identified in part a above can be used by the Super Baseball League to gain a competitive advantage over other sports leagues for the fans' interest and entertainment dollars (Euros, pesos, yen, etc.).