CONCEPTS IN COMPUTING (PGCC)

TEC Team
Prince George's Community College

LibreTexts™

INT 1010: Concepts in Computing (PGCC)

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

This text was compiled on 01/15/2024

TABLE OF CONTENTS

About the Book

Introduction to Information Technology is a survey course in evolving information technology and its relevance to individuals and society. This book examines the categories of computing devices and different types of computer applications, software, and their uses. The emphasis is on enhancing skills in data analysis and programming. Additionally, the book evaluates ethical principles related to privacy, security, intellectual property and how these apply to their academic and professional life. The book also explores strategies to manage risks related to systems security threats. Lastly, the book includes basic principles of connectivity and data communications.

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing****.*

Acknowledgements

The development of this text was under the leadership of the Department of Technology, Engineering, and Construction at Prince George's Community College and Michael W. Smith, Department Chairperson. We would like to thank our colleague, Tammy Cameron-Allen, for her leadership in the development and editing of this work.

Finally, We would like to thank those professors who participated in creating, reviewing, and commenting on drafts of this text:

Joy S Choudhury

Lynn Gross-Butler

Manzoor I Hossain

Velma C Latson

Sean A Varnado

Sonia Washington

CHAPTER OVERVIEW

1: Introduction to INT 1010 - Introduction to Information Technology

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Describe the categories of computing devices;
- Compare and contrast types of computer applications and software and their uses;
- Demonstrate basic proficiency in programming;
- Demonstrate basic proficiency in data analysis using the software;
- Evaluate ethical principles in emerging technologies as they apply to academic and professional life;
- Describe risk management strategies that address information systems security threats; and
- Explore data communications concepts and basic principles of connectivity including wired and wireless technologies;

Introduction to Information Technology is a survey course in evolving information technology and its relevance to individuals and society. Students examine the categories of computing devices and different types of computer applications, software, and their uses. The emphasis in this this course is on enhancing students' skills in data analysis and programming. Additionally, students evaluate ethical principles related to privacy, security, intellectual property and how these apply to their academic and professional life. They also explore strategies to manage risks related to systems security threats. Lastly, students learn about the basic principles of connectivity and data communications.

1.1: Introduction

Introduction to Information Technology

INT 1010, is a survey course in evolving information technology and its relevance to individuals and society. The emphasis in this course is on enhancing students' skills in data analysis and programming and applying learned skills to their academic and professional life.

This course examines:

· Categories of computing devices and different types of computer applications.

· Software and their uses.

· Ethical principles related to privacy, security, intellectual property.

· Strategies to manage risks related to systems security threats.

· Students learn about the basic principles of connectivity and data communications.

Introduction to Technology course material utilizes Open Education Recourses (OER). Purchase of a book is not required. Content will be presented in the form of both text (online) and activities.

Each chapter contains:

- ***Objectives:***

Description of what your accomplishments should be at the end of the chapters.

- ***Content:***

This is content explained in each chapter about different topics.

- ***Learning Activities:***

These are activities assigned to assess your knowledge. They include:

· Individual and group activities

· Quizzes

· Exams

The key to success in INT 1010 is planning your study schedule, utilizing the instructor's office hours, using videos, and recognizing if you need to utilize the tutoring center.

1.3: Communication Etiquette

Netiquette 101

It’s never been easier to connect with professors and classmates online, but along with ease and convenience comes a certain level of responsibility when you are posting or mailing information. Keep in mind this basic rule: don’t do anything online that you wouldn’t like someone else to do to you (i.e. sending an aggressive email, posting an overly critical reply, forwarding messages to third parties without permission). Remember, your online activity leaves a permanent digital trail. Here are some “netiquette” tips to remember in an academic setting:

Act Within the Law:

- Never send a message that threatens, harasses, or blatantly offends any member of the KCC community.
- Always give credit for words or ideas that belong to someone else.
- Identify yourself by name. Don’t pretend to be someone else.

Be Courteous and Respectful:

- Avoid using profanity in an email or discussion forum.
- Keep personal matters between you and your classmates and/or professors private—both in online and face-to-face courses.
- Before you press “send” or post something in an online discussion forum, re-read your text to make sure it won’t be misunderstood. Your readers will not necessarily know your mood or be able to read your body language.
- If you decide to send a message via cellphone, take the time to punctuate your text properly.
- Don’t write in all lowercase letters.
- YOU’RE SHOUTING when you write in all capital letters!
- Abbreviations and emoticons can help convey tone or mood, but keep in mind that they wear thin on some readers. LOL

Emailing Your Professor - Ten Tips

1. Put your class/section in the subject line, followed by the nature of the communication. Example - Subject: (INT 1010 LD01) Assignments
2. Use an appropriate standard greeting, such as “Dear Professor Shaw” or “Good Morning.” Never “Hey.”
3. Keep the message on point. “My question about today’s assignment is this:”
4. Write in standard English—no “textspeak” or slang.
5. Always sign your name at the end (full name if you are unknown to the recipient or first name if you are known).
6. Never ask your professor if you missed anything important; of course you have.
7. Don’t share too much personal detail if you miss class. An absence is an absence.
8. If you will be missing class, always ask what you need to do to keep current.
9. Never forward jokes, memes, or chain letters to your professor.
10. Make readers want to respond. Your email is a reflection of you, your work habits, and your professionalism.

1.4: Summary

Summary

Introduction to Information Technology, INT 1010, is a survey course in evolving information technology and its relevance to individuals and society.

Each chapter contains learning activities and success tips.

Netiquette tips for online and in class.

CHAPTER OVERVIEW

2: Introduction to Information Systems

Introduction

If you are reading this, you are most likely taking a course in information systems, but do you even know what the course is going to cover? When you tell your friends or your family that you are taking a course in information systems, can you explain what it is about? For the past several years, I have taught an Introduction to Information Systems course. The first day of class I ask my students to tell me what they think an information system is. I generally get answers such as “computers,” “databases,” or “Excel.” These are good answers, but definitely incomplete ones. The study of information systems goes far beyond understanding some technologies. Let’s begin our study by defining information systems.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Define Information Systems
- Identify the five components of an Information system
- Describe the role of Information Systems
- Describe Information System users

2.1: Information Systems

Defining Information Systems

Almost all programs in business require students to take a course in something called *information systems*. But what exactly does that term mean? Let's take a look at some of the more popular definitions, first from Wikipedia and then from a couple of textbooks:

- "Information systems (IS) is the study of complementary networks of hardware and software that people and organizations use to collect, filter, process, create, and distribute data."[1]

  In some organizations, a matrix reporting structure was developed in which IT personnel were placed within a department and reported to both the department management and the functional management within IS. The advantages of dedicated IS personnel for each department must be weighed against the need for more control over the strategic information resources of the company.

- "Information systems are combinations of hardware, software, and telecommunications networks that people build and use to collect, create, and distribute useful data, typically in organizational settings."[2]
- "Information systems are interrelated components working together to collect, process, store, and disseminate information to support decision making, coordination, control, analysis, and visualization in an organization."[3]

As you can see, these definitions focus on two different ways of describing information systems: the *components* that make up an information system and the *role* that those components play in an organization. Let's take a look at each of these.

The Role of Information Systems

In this section, we will explore the role information systems play in an organization. Later in the text, you will learn how components collect, store, organize and distribute data throughout the organization. In fact, we might say that one of the roles of information systems is to take data and turn it into information, and then transform that into organizational knowledge. As technology has developed, this role has evolved into the backbone of the organization. To get a full appreciation of the role information systems to play, we will review how they have changed over the years.

Figure 2.1.1: IBM 704 Mainframe (CC-BY; Lawrence Livermore National Laboratory)

The Mainframe Era

From the late 1950s through the 1960s, computers were seen as a way to more efficiently do calculations. These first business computers were room-sized monsters, with several refrigerator-sized machines linked together. The primary work of these devices was to organize and store large volumes of information that were tedious to manage by hand. Only large businesses, universities, and government agencies could afford them, and they took a crew of specialized personnel and specialized facilities to maintain. These devices served dozens to hundreds of users at a time through a process called time-sharing. Typical functions included scientific calculations and accounting, under the broader umbrella of "data processing."

IBM

Registered trademark of International Business Machines

In the late 1960s, the Manufacturing Resources Planning (MRP) systems were introduced. This software, running on a mainframe computer, gave companies the ability to manage the manufacturing process, making it more efficient. From tracking inventory to creating bills of materials to scheduling production, the MRP systems (and later the MRP II systems) gave more businesses a reason to want to integrate computing into their processes. IBM became the dominant mainframe company. Nicknamed "Big Blue," the company became synonymous with business computing. Continued improvement in software and the availability of cheaper hardware eventually brought mainframe computers (and their little sibling, the minicomputer) into most large businesses.

The PC Revolution

In 1975, the first microcomputer was announced on the cover of *Popular Mechanics*: the Altair 8800. Microcomputers' immediate popularity sparked the imagination of entrepreneurs everywhere, and there were quickly dozens of companies making these "personal computers." Though at first just a niche product for computer hobbyists, improvements in usability and the availability of practical software led to growing sales. The most prominent of these early personal computer makers was a little company known as Apple Computer, headed by Steve Jobs and Steve Wozniak, with the hugely successful "Apple II." Not wanting to be left out of the revolution, in 1981 IBM (teaming with a little company called Microsoft for their operating-system software) hurriedly released their own version of the personal computer, simply called the "PC." Businesses, who had used IBM mainframes for years to run their businesses, finally had the permission they needed to bring personal computers into their companies, and the IBM PC took off. The IBM PC was named *Time* magazine's "Man of the Year" for 1982.

Because of the IBM PC's open architecture, it was easy for other companies to copy, or "clone" it. During the 1980s, many new computer companies sprung up, offering less expensive versions of the PC. This drove prices down and spurred innovation. Microsoft developed its Windows operating system and made the PC even easier to use. Common uses for the PC during this period included word processing, spreadsheets, and databases. These early PCs were not connected to any sort of network; for the most part, they stood alone as islands of innovation within the larger organization.

Client-Server

In the mid-1980s, businesses began to see the need to connect their computers together as a way to collaborate and share resources. This networking architecture was referred to as "client-server" because users would log in to the local area network (LAN) from their PC (the "client") by connecting to a powerful computer called a "server," which would then grant them rights to different resources on the network (such as shared file areas and a printer). Software companies began developing applications that allowed multiple users to access the same data at the same time. This evolved into software applications for communicating, with the first real popular use of electronic mail appearing at this time.

This networking and data sharing all stayed within the confines of each business, for the most part. While there was sharing of electronic data between companies, this was a very specialized function. Computers were now seen as tools to collaborate internally, within an organization. In fact, these networks of computers were becoming so powerful that they were replacing many of the functions previously performed by the larger mainframe computers at a fraction of the cost. It was during this era that the first Enterprise Resource Planning (ERP) systems were developed and run on the client-server architecture. An ERP system is a software application with a centralized database that can be used to run a company's entire business. With separate modules for accounting, finance, inventory, human resources, and many, many more, ERP systems, with Germany's SAP leading the way, represented the state of the art in information systems integration.

The World Wide Web and E-Commerce

First invented in 1969, the Internet was confined to use by universities, government agencies, and researchers for many years. Its rather arcane commands and user applications made it unsuitable for mainstream use in business. One exception to this was the ability to expand electronic mail outside the confines of a single organization. While the first e-mail messages on the Internet were sent in the early 1970s, companies who wanted to expand their LAN-based e-mail started hooking up to the Internet in the 1980s. Companies began connecting their internal networks to the Internet in order to allow communication between their employees and employees at other companies. It was with these early Internet connections that the computer truly began to evolve from a computational device to a communications device.

Figure 2.1.2: The historic **World Wide Web** logo, designed by Robert Cailliau. (1990; Robert Cailliau)

In 1989, Tim Berners-Lee developed a simpler way for researchers to share information over the network at CERN laboratories, a concept he called the World Wide Web.[4] This invention became the launching point of the growth of the Internet as a way for businesses to share information about themselves. As web browsers and Internet connections became the norm, companies rushed to grab domain names and create websites.

Figure 2.1.3: The CERN data centre in 2010 housing some WWW servers. The CERN datacenter with World Wide Web and Mail servers. The rear of the equipment racks are exposed to the room, indicating cold aisle containment is being practiced.

In 1991, the National Science Foundation, which governed how the Internet was used, lifted restrictions on its commercial use. The year 1994 saw the establishment of both eBay and Amazon.com, two true pioneers in the use of the new digital marketplace. A mad rush of investment in Internet-based businesses led to the dot-com boom through the late 1990s, and then the dot-com bust in 2000. While much can be learned from the speculation and crazy economic theories espoused during that bubble, one important outcome for businesses was that thousands of miles of Internet connections were laid around the world during that time. The world became truly “wired” heading into the new millennium, ushering in the era of globalization.

As it became more expected for companies to be connected to the Internet, the digital world also became a more dangerous place. Computer viruses and worms, once slowly propagated through the sharing of computer disks, could now grow with tremendous speed via the Internet. Software written for a disconnected world found it very difficult to defend against these sorts of threats. A whole new industry of computer and Internet security arose.

Web 2.0

As the world recovered from the dot-com bust, the use of technology in business continued to evolve at a frantic pace. Websites became interactive; instead of just visiting a site to find out about a business and purchase its products, customers wanted to be able to customize their experience and interact with the business. This new type of interactive website, where you did not have to know how to create a web page or do any programming in order to put information online, became known as web 2.0. Web 2.0 is exemplified by blogging, social networking, and interactive comments being available on many websites. This new web-2.0 world, in which online interaction became expected, had a big impact on many businesses and even whole industries. Some industries, such as bookstores, found themselves relegated to niche status. Others, such as video rental chains and travel agencies, simply began going out of business as they were replaced by online technologies. This process of technology replacing a middleman in a transaction is called disintermediation.

As the world became more connected, new questions arose. Should access to the Internet be considered a right? Can I copy a song that I downloaded from the Internet? How can I keep information that I have put on a website private? What information is acceptable to collect from children? Technology moved so fast that policymakers did not have enough time to enact appropriate laws, making for a Wild West–type atmosphere.

The Post-PC World

After thirty years as the primary computing device used in most businesses, sales of the PC are now beginning to decline as sales of tablets and smartphones are taking off. Just as the mainframe before it, the PC will continue to play a key role in business, but will no longer be the primary way that people interact and do business. The limited storage and processing power of these devices is being offset by a move to "cloud" computing, which allows for storage, sharing, and backup of the information on a massive scale. This will require new rounds of thinking and innovation on the part of businesses as technology continues to advance.

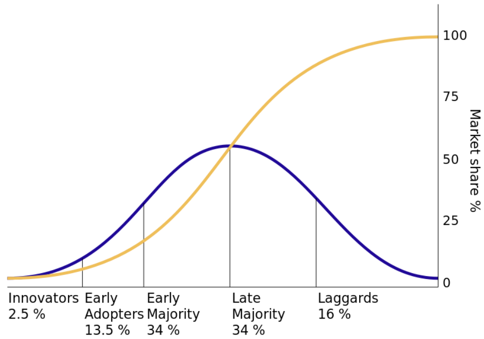
The Eras of Business Computing

Era	Hardware	Operating System	Applications
Mainframe (1960s)	Terminals connected to mainframe computer.	Time-sharing (TSO) on MVS	Custom-written MRP software
PC (mid-1980s)	IBM PC or compatible. Sometimes connected to mainframe computer via expansion card.	MS-DOS	WordPerfect, Lotus 1-2-3
Client-Server (late 80s to early 90s)	IBM PC "clone" on a Novell Network.	Windows for Workgroups	Microsoft Word, Microsoft Excel
World Wide Web (mid-90s to early 2000s)	IBM PC "clone" connected to company intranet.	Windows XP	Microsoft Office, Internet Explorer
Web 2.0 (mid-2000s to present)	Laptop connected to company Wi-Fi.	Windows 7	Microsoft Office, Firefox
Post-PC (today and beyond)	Apple iPad	iOS	Mobile-friendly websites, mobile apps

Information Systems Users – Types of Users

Besides the people who work to create, administer, and manage information systems, there is one more extremely important group of people, namely, the users of information systems. This group represents a very large percentage of an organization's employees. If the user is not able to successfully learn and use an information system, the system is doomed to failure.

Technology adoption user types



Diffusion of Innovation

One tool that can be used to understand how users will adopt a new technology comes from a 1962 study by Everett Rogers. In his book, *Diffusion of Innovation*,[1]Rogers studied how farmers adopted new technologies and noticed that the adoption rate started slowly and then dramatically increased once adoption hit a certain point. He identified five specific types of technology adopters:

- **Innovators**. Innovators are the first individuals to adopt new technology. Innovators are willing to take risks, are the youngest in age, have the highest social class, have great financial liquidity, are very social, and have the closest contact with scientific sources and interaction with other innovators. Risk tolerance is high so there is a willingness to adopt technologies that may ultimately fail. Financial resources help absorb these failures (Rogers, 1962, p. 282).
- **Early adopters**. The early adopters are those who adopt innovation soon after a technology has been introduced and proven. These individuals have the highest degree of opinion leadership among the other adopter categories, which means that these adopters can influence the opinions of the largest majority. Characteristics include being younger in age, having a higher social status, possessing more financial liquidity, having advanced education, and being more socially aware than later adopters. These adopters are more discrete in adoption choices than innovators, and realize the judicious choice of adoption will help them maintain a central communication position (Rogers, 1962, p. 283).
- **Early majority**. Individuals in this category adopt an innovation after a varying degree of time. This time of adoption is significantly longer than the innovators and early adopters. This group tends to be slower in the adoption process, has above average social status, has contact with early adopters, and seldom holds positions of opinion leadership in a system (Rogers, 1962, p. 283).
- **Late majority**. The late majority will adopt an innovation after the average member of the society. These individuals approach an innovation with a high degree of skepticism, have below average social status, very little financial liquidity, are in contact with others in the late majority and the early majority, and show very little opinion leadership.
- **Laggards**. Individuals in this category are the last to adopt an innovation. Unlike those in the previous categories, individuals in this category show no opinion leadership. These individuals typically have an aversion to change agents and tend to be advanced in age. Laggards typically tend to be focused on “traditions,” are likely to have the lowest social status and the lowest financial liquidity, be the oldest of all other adopters and be in contact with only family and close friends.[2]

These five types of users can be translated into information technology adopters as well, and provide additional insight into how to implement new information systems within the organization. For example, when rolling out a new system, IT may want to identify the innovators and early adopters within the organization and work with them first, then leverage their adoption to drive the rest of the implementation to the other users.

New Models of Organizations

The integration of information technology has influenced the structure of organizations. The increased ability to communicate and share information has led to a “flattening” of the organizational structure due to the removal of one or more layers of management.

The network-based organizational structure is another change enabled by information systems. In a network-based organizational structure, groups of employees can work somewhat independently to accomplish a project. People with the right skills are brought together for a project and then released to work on other projects when that project is over. These groups are somewhat informal and allow for all members of the group to maximize their effectiveness.

- **9: The People in Information Systems** by David T. Bourgeois, James L. Smith, Shouhong Wang & Joseph Mortati is licensed CC BY-NC 4.0. Original source: https://digitalcommons.biola.edu/open-textbooks/1.
- **2.3: Components of an Information System** by David T. Bourgeois, James L. Smith, Shouhong Wang & Joseph Mortati is licensed CC BY-NC 4.0.

2.2: What is Competitive Advantage

Competitive Advantage

What does it mean when a company has a competitive advantage? What are the factors that play into it? While there are entire courses and many different opinions on this topic, let's go with one of the most accepted definitions, developed by Michael Porter in his book *Competitive Advantage: Creating and Sustaining Superior Performance*. A company is said to have a competitive advantage over its rivals when it is able to sustain profits that exceed average for the industry. According to Porter, there are two primary methods for obtaining competitive advantage: cost advantage and differentiation advantage.

One of the ways that information systems participated in competitive advantage is through integrating the supply chain electronically. This is primarily done through a process called electronic data interchange, or EDI. EDI can be thought of as the *computer-to-computer exchange of business documents in a standard electronic format between business partners*. By integrating suppliers and distributors via EDI, a company can vastly reduce the resources required to manage the relevant information. Instead of manually ordering supplies, the company can simply place an order via the computer and the next time the order process runs, it is ordered.

Using Information Systems for Competitive Advantage

Now that we have an understanding of competitive advantage and some of the ways that IT may be used to help organizations gain it, we will turn our attention to some specific examples. A strategic information system is an information system that is designed specifically to implement an organizational strategy meant to provide a competitive advantage. These sorts of systems began popping up in the 1980s, as noted in a paper by Charles Wiseman entitled "Creating Competitive Weapons From Information Systems."[4]

Specifically, a strategic information system is one that attempts to do one or more of the following:

- deliver a product or a service at a lower cost;
- deliver a product or service that is differentiated;
- help an organization focus on a specific market segment;
- enable innovation.

Following are some examples of information systems that fall into this category.

Collaborative Systems

As organizations began to implement networking technologies, information systems emerged that allowed employees to begin collaborating in different ways. These systems allowed users to brainstorm ideas together without the necessity of physical, face-to-face meetings. Utilizing tools such as discussion boards, document sharing, and video, these systems made it possible for ideas to be shared in new ways and the thought processes behind these ideas to be documented.

Broadly speaking, any software that allows multiple users to interact on a document or topic could be considered collaborative. Electronic mail, a shared Word document, social networks, and discussion boards would fall into this broad definition. However, many software tools have been created that are designed specifically for collaborative purposes. These tools offer a broad spectrum of collaborative functions. Here is just a short list of some collaborative tools available for businesses today:

- Google Drive. Google Drive offers a suite of office applications (such as a word processor, spreadsheet, drawing, presentation) that can be shared between individuals. Multiple users can edit the documents at the same time and threaded comments are available.
- Microsoft SharePoint. SharePoint integrates with Microsoft Office and allows for collaboration using tools most office workers are familiar with. SharePoint was covered in more detail in chapter 5.
- Cisco WebEx. WebEx is a business communications platform that combines video and audio communications and allows participants to interact with each other's computer desktops. WebEx also provides a shared whiteboard and the capability for text-based chat to be going on during the sessions, along with many other features. Mobile editions of WebEx allow for full participation using smartphones and tablets.
- Atlassian Confluence. Confluence provides an all-in-one project-management application that allows users to collaborate on documents and communicate progress. The mobile edition of Confluence allows the project members to stay connected throughout the project.

- IBM Lotus Notes/Domino. One of the first true “groupware” collaboration tools, Lotus Notes (and its web-based cousin, Domino) provides a full suite of collaboration software, including integrated e-mail.

2.3: Components of an Information System

The Components of Information Systems

Many students understand that an information system has something to do with databases or spreadsheets. Others mention computers and e-commerce. And they are all right, at least in part: information systems are made up of different components that work together to provide value to an organization.

The first way I describe information systems to students is to tell them that they are made up of five components: hardware, software, data, people, and process. The first three, hardware, software, data, fitting under the category *technology*, are generally what most students think of when asked to define information systems. But the last two, people and process, are really what separate the idea of information systems from more technical fields, such as computer science. In order to fully understand information systems, students must understand how all of these components work together to bring value to an organization.

Technical Components of Information Systems

Technology can be thought of as the application of scientific knowledge for practical purposes. From the invention of the wheel to the harnessing of electricity for artificial lighting, technology is a part of our lives in so many ways that we tend to take it for granted. As discussed before, the first three components of information systems – hardware, software, and data – all fall under the category of technology. Each of these will get its own chapter and a much lengthier discussion, but we will take a moment here to introduce them so we can get a full understanding of an information system.

Hardware

The first component is hardware of an information system. Hardware is the part of an information system you can touch – the physical components of the technology. Computers, keyboards, disk drives, iPads, and flash drives are all examples of information systems hardware. We will spend some time going over these components in future chapters. Figure 3.1 displays internal parts on a computer.

Software

Software, the second component of an information system, is a set of instructions that tells the hardware what to do. Software is not tangible – it cannot be touched. When programmers create software programs, what they are really doing is simply typing out lists of instructions telling the hardware what to do. There are several categories of software, with the two main categories being operating-system software, which makes the hardware usable, and application software, which does something useful. Examples of operating systems include Microsoft Windows on a personal computer and Google's Android on a mobile phone. Examples of application software are Microsoft Excel and Angry Birds.

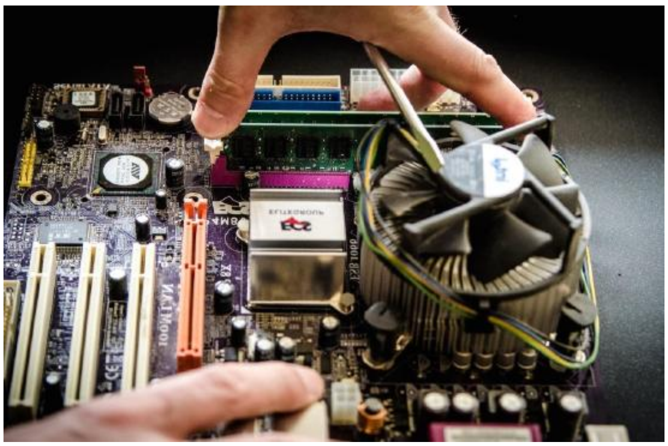


Figure 2.3.1: Computer Hardware. (Michal Jarmoluk, CCO, via pixnio.com)

Below is a diagram showing how the user interacts with application software on a typical desktop computer. The application software layer interfaces with the operating system, which in turn communicates with the hardware. The arrows indicate information flow.

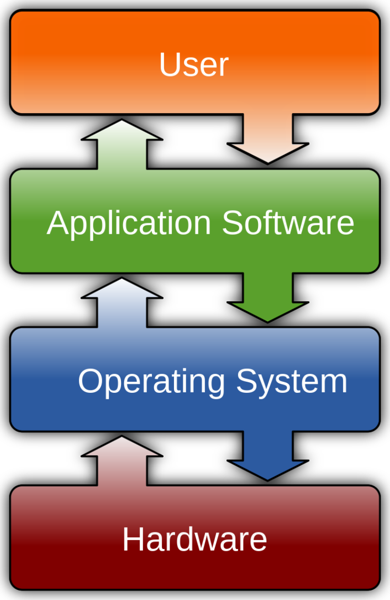


Figure 2.3.2: Computer Software : (CC BY-SA 3.0, via Wikimedia Commons)

Data

The third component is data. You can think of data as a collection of facts. For example, your street address, the city you live in, and your phone number are all pieces of data. Like software, data is also intangible. By themselves, pieces of data are not really very useful. But aggregated, indexed, and organized together into a database, data can become a powerful tool for businesses. In fact, all of the definitions presented at the beginning of this chapter focused on how information systems manage data. Organizations collect all kinds of data and use it to make decisions. These decisions can then be analyzed as to their effectiveness and the organization can be improved.

Networking Communication: A Fourth Technology Piece?

Besides the components of hardware, software, and data, which have long been considered the core technology of information systems, it has been suggested that one other component should be added: communication. An information system can exist without the ability to communicate – the first personal computers were stand-alone machines that did not access the Internet. However, in today's hyper-connected world, it is an extremely rare computer that does not connect to another device or to a network. Technically, the networking communication component is made up of hardware and software, but it is such a core feature of today's information systems that it has become its own category.

Figure 2.3.3: Data: Different Types of Data. (CC BY 3.0, via Wikimedia Commons)

Non-Technical Components of Information Systems

People

When thinking about information systems, it is easy to get focused on the technology components and forget that we must look beyond these tools to fully understand how they integrate into an organization. A focus on the people involved in information systems is the fourth component. From the front-line help-desk workers to systems analysts, to programmers, all the way up to the chief information officer (CIO), the people involved with information systems are an essential element that must not be overlooked.

Process

Figure 2.3.4: People. (geralt, CC, via pixabay.com)

The last component of information systems is process. We have all heard the term *process* before, but what exactly does it mean? A process is a series of steps undertaken to achieve a desired outcome or goal. Processes are something that businesses go through every day in order to accomplish their mission. The better their processes, the more effective the business. Some businesses see their processes as a strategy for achieving competitive advantage. A process that achieves its goal in a unique way can set a company apart. A process that eliminates costs can allow a company to lower its prices (or retain more profit).

The simplest way to document a process is to simply create a list. The list shows each step in the process; each step can be checked off upon completion. For example, a simple process, such as how to create an account on eBay, might look like this:

1. Go to ebay.com.
2. Click on "register."
3. Enter your contact information in the "Tell us about you" box.
4. Choose your user ID and password.
5. Agree to User Agreement and Privacy Policy by clicking on "Submit."

For processes that are not so straightforward, documenting the process as a checklist may not be sufficient. For example, here is the process for determining if an article for a term needs to be added to Wikipedia:

1. Search Wikipedia to determine if the term already exists.
2. If the term is found, then an article is already written, so you must think of another term. Go to 1.
3. If the term is not found, then look to see if there is a related term.
4. If there is a related term, then create a redirect.
5. If there is not a related term, then create a new article.

This procedure is relatively simple – in fact, it has the same number of steps as the previous example – but because it has some decision points, it is more difficult to track with as a simple list.

Information systems are becoming more and more integrated with organizational processes, bringing more productivity and better control to those processes. But simply automating activities using technology is not enough – businesses looking to effectively utilize information systems do more. Using technology to manage and improve processes, both within a company and externally with suppliers and customers is the ultimate goal. Technology buzzwords such as "business process reengineering," "business process management," and "enterprise resource planning" all have to do with the continued improvement of these business procedures and the integration of technology with them. Businesses hoping to gain an advantage over their competitors are highly focused on this component of information systems.

1. Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press↵
2. Rogers, E. M. (1962). *Diffusion of innovations*. New York: Free Press↵

2.4: Summary

Summary

Defining information systems is interpreted indifferently.

Five components of an Information System are hardware, data, people, and process.

Information Systems had roles during the Mainframe, PC revolution, Client Server, World Wide Web, E-Commerce, and Web 2.0 eras.

Information Systems are thought to create competitive advantages.

CHAPTER OVERVIEW

3: Hardware Components of an Information System

Introduction

An information system is made up of five components: hardware, software, data, people, and process. The physical parts of computing devices – those that you can actually touch – are referred to as hardware. In this chapter, we will take a look at this component of information systems, learn a little bit about how it works, and discuss some of the current trends surrounding it.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Describe different computer hardware that makes up the computer system;
- Identify peripheral devices that connect to the computer via various connection ports;
- Identify output devices that transform computer-readable data back into an information format;

3.1: Hardware

Computer Hardware

Different Computer Systems

There are a number of different types of computer systems in use today. The personal computer, also known as the PC or desktop, is one of the most common types of computer due to its versatility and relatively low price. Laptops (or Notebooks) are generally very similar, although may use lower-power or reduced size components. Others include tablets, mobile phones, and e-readers for work, home, and leisure activities.

Laptops

Portable computing improved tremendously over the years moving much of computing to the Internet. Laptop and Notebook computer users are making use of “the cloud” for data and application storage. The “luggable” computer has given way to a much lighter clamshell computer that weighs from 4 to 6 pounds and runs on batteries. In fact, the most recent advances in technology give us a new class of laptops that is quickly becoming the standard: these laptops are extremely light and portable and use less power than their larger counterparts. The MacBook Air is a good example of this: it weighs less than three pounds and is only 0.68 inches thick!

Figure 4.1.1: The original 2008 MacBook Air. (Tim Malabuyo via CC BY-2.0, via Wikimedia Commons)

Mobile Phones

The first modern-day mobile phone was invented in 1973. Resembling a brick and weighing in at two pounds, it was priced out of reach for most consumers at nearly four thousand dollars. Since then, mobile phones have become smaller and less expensive; today mobile phones are a modern convenience available to all levels of society. As mobile phones evolved, they became more like small computers. These smartphones have many of the same characteristics as a personal computer, such as an operating system and memory. The first smartphone was the IBM Simon, introduced in 1994.

In January of 2007, Apple introduced the iPhone. Its ease of use and intuitive interface made it an immediate success and solidified the future of smartphones. Running on an operating system called iOS, the iPhone was really a small computer with a touch-screen interface. In 2008, the first Android phone was released, with similar functionality.

Figure 4.1.2: A decade of evolution of mobile phones, from a 1994 Motorola 8900X-2 to the 2004 HTC Typhoon, an early smartphone. (Tim Malabuyo , CC BY-2.0, via Wikimedia Commons)

Figure 4.1.3: Personal Handy-phone System mobiles and modems, 1997–2003. (Marus, Public Domain, via Wikimedia Commons)

Figure 4.1.4: SAMSUNG Galaxy Note10, SAMSUNG Galaxy Note10+ , SAMSUNG Galaxy Note10 Lite. (Google, CC BY-3.0, via Wikimedia Commons)

Tablet Computers

A tablet computer is one that uses a touch screen as its primary input and is small enough and light enough to be carried around easily. They generally have no keyboard and are self-contained inside a rectangular case. The first tablet computers appeared in the early 2000s and used an attached pen as a writing device for input. These tablets ranged in size from small personal digital assistants (PDAs), which were handheld, to full-sized, 14-inch devices. Most early tablets used a version of an existing computer operating system, such as Windows or Linux.

These early tablet devices were, for the most part, commercial failures. In January 2010, Apple introduced the iPad, which ushered in a new era of tablet computing. Instead of a pen, the iPad used the finger as the primary input device. Instead of using the operating system of their desktop and laptop computers, Apple chose to use iOS, the operating system of the iPhone. Because the iPad had a user interface that was the same as the iPhone, consumers felt comfortable and sales took off. The iPad has set the standard for tablet computing. After the success of the iPad, computer manufacturers began to develop new tablets that utilized operating systems that were designed for mobile devices, such as Android.

Figure 4.1.5: Asus Transformer Pad, a 2-in-1 detachable tablet, powered by the Android operating system. (Dinominant at English Wikipedia - Own work, CCO, via Wikimedia Commons)

The Rise of Mobile Computing

Mobile computing is having a huge impact on the business world today. The use of smartphones and tablet computers is rising at double-digit rates each year. The Gartner Group, in a report issued in April 2013, estimates that over 1.7 million mobile phones will ship in the US in 2013 as compared to just over 340,000 personal computers. Over half of these mobile phones are smartphones.[2] Almost 200,000 tablet computers are predicted to ship in 2013. According to the report, PC shipments will continue to decline as phone and tablet shipments continue to increase. [3]

Integrated Computing

Along with advances in computers themselves, computing technology is being integrated into many everyday products. From automobiles to refrigerators to airplanes, computing technology is enhancing what these devices can do and is adding capabilities that would have been considered science fiction just a few years ago. Here are two of the latest ways that computing technologies are being integrated into everyday products:

- The Smart House
- The Self-Driving Car

Next, we will look at the components of these devices.

Computer System Components

Computer hardware (usually simply called hardware when a computing context is concerned) is the collection of physical elements that constitutes a computer system. Computer hardware is the physical parts or components of a computer, such as a monitor, mouse, keyboard, computer data storage, hard disk drive (HDD), graphic cards, sound cards, memory, motherboard, and so on, all of which are physical objects that are tangible. In contrast, the software is instructions that can be stored and run by hardware.

Chassis

The computer chassis or case is a plastic or metal enclosure that houses most of the components. Those found on desktop computers are usually small enough to fit under a desk, however in recent years more compact designs have become more commonplace, such as the all-in-one style designs. Though a case can basically be big or small, what matters more is which form factor of motherboard it’s designed for. Laptops are computers that usually come in a clamshell form factor, again however in more recent years deviations from this form factor have started to emerge such as laptops that have a detachable screen that becomes tablet computers in their own right. In the following sections, we will examine the internal and external components of a computer.

Motherboard History

Prior to the invention of the microprocessor, the digital computer consisted of multiple printed circuit boards in a card-cage case with components connected by a backplane, a set of interconnected sockets. In very old designs the wires were discrete connections between card connector pins, but printed circuit boards soon became the standard practice. The Central Processing Unit (CPU), memory, and peripherals were housed on individual printed circuit boards, which were plugged into the backplate.

During the late 1980s and 1990s, it became economical to move an increasing number of peripheral functions onto the motherboard. In the late 1980s, personal computer motherboards began to include single ICs (also called Super I/O chips) capable of supporting a set of low-speed peripherals: keyboard, mouse, floppy disk drive, serial ports, and parallel ports. By the late-1990s, many personal computer motherboards supported a full range of audio, video, storage, and networking functions without the need for any expansion cards at all; higher-end systems for 3D gaming and computer graphics typically retained only the graphics card as a separate component.

What is a Motherboard

A motherboard (sometimes alternatively known as the mainboard, system board, planar board, or logic board, colloquially, a mobo) is the main component of a computer. It holds and allows communication between many of the crucial electronic components of a system that connects the other parts of the computer including the CPU, the RAM, the disk drives (CD, DVD, hard disk, or any others) as well as any peripherals connected via the ports or the expansion slots. Components directly attached to or part of the motherboard include a central processing unit, chipset, random-access memory, read-only memory buses, and CMOS battery.

Central Processing Unit

The **CPU** (Central Processing Unit) also called processor or microprocessor, performs most of the calculations which enable a computer to function, and is sometimes referred to as the “brain” of the computer. It is usually cooled by a heat sink and fan. Most microprocessors include an on-die Graphics Processing Unit (GPU), memory, peripheral interfaces, and other components such as integrated devices also known as microcontrollers or systems on a chip (SoC). Some computers employ a multi-core processor, which is a single chip containing two or more CPUs called "cores"; in that context, single chips are sometimes referred to as "sockets". Array processors or vector processors have multiple processors that operate in parallel, with no unit considered central.

Chipset

The chipset is a set of electronic components in an integrated circuit that manages the data flow between the processor, memory, and peripherals. It is designed to work with a specific family of microprocessors. Because it controls communications between the processor and external devices, the chipset plays a crucial role in determining system performance.

Ransom-Access Memory

The Random-Access Memory (RAM), a volatile memory chip (where stored information is lost if power is removed) stores the code and data that are being actively accessed by the CPU. Today, random-access memory takes the form of integrated circuits. Other types of non-volatile memory exist that allow random access for reading operations, but either do not allow write operations or have limitations on them. These include most types of ROM and a type of flash memory called *NOR-Flash*.

Read-Only Memory

The Read-Only Memory (ROM), a non-volatile memory chip, stores data by permanently enabling or disabling selected transistors, such that the memory cannot be altered. It stores the BIOS that runs when the computer is powered on or otherwise begins execution, a process known as Bootstrapping, or "booting" or "booting up". The **BIOS** (Basic Input Output System) includes boot firmware and power management firmware. Newer motherboards use Unified Extensible Firmware Interface (UEFI) instead of BIOS. There are other types of non-volatile memory that are not based on solid-state IC technology, including Optical storage media, such CD-ROM which is read-only (analogous to masked ROM). CD-R is Write Once Read Many (analogous to PROM), while CD-RW supports erase-rewrite cycles (analogous to EEPROM); both are designed for backward compatibility with CD-ROM.

Buses

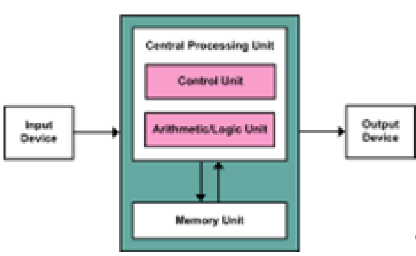
Buses are communication systems that transfer data between components inside a computer, or between computers. This expression covers all related hardware components (wire, optical fiber, etc.) and software, including communication protocols. An internal bus (also known as an internal data bus, memory bus, system bus, or Front-Side-Bus) expand cards for graphics and sound. Internal data buses are also referred to as local bus because they are intended to connect to local devices. This bus is typically rather quick and independent of the rest of the computer operations. The external bus, or expansion bus, is made up of the electronic pathways that connect the different external devices, such as printers, etc., to the computer.

CMOS Battery

The CMOS battery is also attached to the motherboard. This battery is the same as a watch battery or a battery for a remote to a car's central locking system. Most batteries are CR2032, which powers the memory for date and time in the BIOS chip.

Von Neumann Architecture

The template for all modern computers is the Von Neumann architecture, detailed in a 1945 paper by Hungarian mathematician John von Neumann.



This describes a design architecture for an electronic digital computer with subdivisions of a processing unit consisting of an arithmetic logic unit and processor registers, a control unit containing an instruction register and program counter, a memory to store both data and instructions, external mass storage, and input and output mechanisms.[3] The meaning of the term has evolved to mean a stored-program computer in which an instruction fetch and a data operation cannot occur at the same time because they share a common bus. This is referred to as the Von Neumann bottleneck and often limits the performance of the system.

Modern Motherboards

Since the 1990's, modern motherboards have had more internal components added for more efficiency. In addition to the CPU, chipset, RAM, ROM, buses, and CMOS battery, today you will find in some devices sockets, slots, clock generators, power connectors, and connectors for hard drives. They are discussed in detail in the following section.

Sockets

Sockets (or slots) in which one or more microprocessors may be installed. In the case of CPUs in ball grid array packages, such as the VIA C3, the CPU is directly soldered to the motherboard. A CPU socket (central processing unit) attaches to a Printed Circuit Board (PCB) and is designed to house a CPU (also called a microprocessor). It is a special type of integrated circuit socket designed for very high pin counts. A CPU socket provides many functions, including a physical structure to support the CPU, support for a heat sink, facilitating replacement (as well as reducing cost), and most importantly, forming an electrical interface both with the CPU and the PCB. CPU sockets on the motherboard can most often be found in most desktop and server computers (laptops typically use surface mount CPUs), particularly those based on the Intel x86 architecture. A CPU socket type and motherboard chipset must support the CPU series and speed.

Slots

Slots into which the system's main memory is to be installed (typically in the form of DIMM modules containing DRAM chips). **Slots for expansion cards** (the interface to the system via the buses supported by the chipset)

Clock Generator

A clock generator produces the system clock signal to synchronize the various components.

Power Connectors

Power connectors receive electrical power from the computer power supply and distribute it to the CPU, chipset, main memory, and expansion cards. As of 2007, some graphics cards (e.g. GeForce 8 and Radeon R600) require more power than the motherboard can provide, and thus dedicated connectors have been introduced to attach them directly to the power supply.

Connectors for Hard Drives

Connectors for hard drives are typically SATA only. Disk drives also connect to the power supply.

Additionally, nearly all motherboards include logic and connectors to support commonly used input devices, such as PS/2 connectors for a mouse and keyboard. Occasionally video interface hardware was also integrated into the motherboard.

Given the high thermal design power of high-speed computer CPUs and components, modern motherboards nearly always include heat sinks and mounting points for fans to dissipate excess heat.

Peripheral Devices

A peripheral device by definition is what a computer can read to receive or output information. The mode of receiving this information is via removable media, storage devices, or Bluetooth. In order for a personal computer to be useful, it must have channels for receiving input from the user and channels for delivering output to the user. These input and output devices connect to the computer via various connection ports, which generally are part of the motherboard and are accessible outside the computer case.

Storage Devices

Besides fixed storage components, removable storage media are also used in most personal computers. Removable media allows you to take your data with you. And just as with all other digital technologies, these media have gotten smaller and more powerful as the years have gone by. Early computers used floppy disks, which could be inserted into a disk drive in the computer. Data was stored on a magnetic disk inside an enclosure. These disks ranged from 8″ in the earliest days down to 3 1/2″.

Floppy-disk evolution (8″ to 5 1/4″ to 3 1/2″) (Public Domain)

Around the turn of the century, a new portable storage technology was being developed: the USB flash drive. This device attaches to the universal serial bus (USB) connector, which became standard on all personal computers beginning in the late 1990s. As with all other storage media, flash drive storage capacity has skyrocketed over the years, from initial capacities of eight megabytes to current capacities of 64 gigabytes, and still growing.

In early personal computers, specific ports were designed for each type of output device. The configuration of these ports has evolved over the years, becoming more and more standardized over time. Today, almost all devices plug into a computer through the use of a USB port. This port type, first introduced in 1996, has increased in its capabilities, both in its data transfer rate and power supplied.

Bluetooth

How do these devices connect? Besides USB, some input and output devices connect to the computer via a wireless-technology standard called Bluetooth. Bluetooth was first invented in the 1990s and exchanges data over short distances using radio waves. Bluetooth generally has a range of 100 to 150 feet. For devices to communicate via Bluetooth, both the personal computer and the connecting device must have a Bluetooth communication chip installed. The latest version is 5.2.

Input devices

Data may enter an information system in a variety of different ways, and the input device that is most appropriate will usually depend on the type of data being entered into the system, how frequently this is done, and who is responsible for the activity. For example, it would be more efficient to scan a page of typed text into an information system rather than retyping it, but if this happens very seldom, and if typing staff are readily available, then the cost of the scanner might not be justified. However, all of the input devices described in this chapter have at least one thing in common: the ability to translate non-digital data types such as text, sound or graphics into digital (i.e. binary) format for processing by a computer.

The keyboard

A lot of input still happens by means of a keyboard. Usually, the information that is entered by means of a keyboard is displayed on the monitor. The layout of most keyboards is similar to that of the original typewriter on which it was modeled. Ironically, this "QWERTY" keyboard layout was originally designed to slow the operator down, so that the keys of the typewriter would not get stuck against each other. This layout now works counter-productively since a computer can process keyboard input many times faster than even the fastest typist can manage. A number of attempts have been made to design alternative layouts by rearranging the keys (the Dvorak keyboard) or by reducing the number of keys. None of these alternative designs has really caught on. Special keyboards have also been designed for countries that use a non-Roman alphabet, and also for disabled people.

Pointing devices

The now-ubiquitous electronic *mouse* is an essential input device for use with any graphical user interface. It consists of a plastic molded housing, designed to fit snugly in the palm of the hand, with a small ball at its bottom. Moving the mouse across a flat surface will translate the movements into a rolling action of the ball. This is translated into electronic signals that direct the corresponding movement of a cursor on the computer monitor. Buttons on the mouse can then be used to select icons or menu items, or the cursor can be used to trace drawings on the screen.

The less popular *trackball* operates exactly like an “upside-down” mouse except that the ball is much larger and, instead of the mouse being moved over a surface, the user manipulates the ball directly. Since the trackball can be built into the side of the keyboard, it obviates the need for a free surface area and is therefore handy in situations where desktop surface area is at a premium or not available. Originally popular in educational laboratory settings and for laptop computers, trackballs are now mainly confined to exhibition displays and other public terminals.

Touch-screens are computer monitors that incorporate sensors on the screen panel itself or its sides. The user can indicate or select an area or location on the screen by pressing a finger onto the monitor. *Light and touch pens* work on a similar principle, except that a stylus is used, allowing for much finer control. Touch pens are more commonly used with handheld computers such as personal organizers or digital assistants. They have a pen-based interface whereby a stylus (a pen without ink) is used on the small touch-sensitive screen of the handheld computer, mainly by means of ticking off pre-defined options, although the fancier models support data entry either by means of a stylized alphabet, which resembles a type of shorthand, or some other more sophisticated handwriting recognition interface.

Digitizer tablets also use a pressure-sensitive area with a stylus. This can be used to trace drawings. A similar conceptual approach is used for the touchpad that can be found on the majority of new notebook computers, replacing the more awkward joystick or trackball. The user controls the cursor by moving a finger across a fairly small rectangular touch-sensitive area below the keyboard, usually about 5 cm by 7 cm.

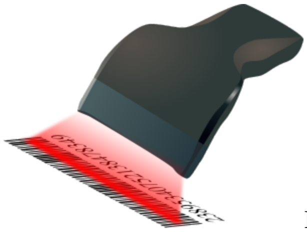
A large number of game interfaces have been developed to provide a more realistic and natural interface in various gaming situations and simulations: the joystick, steering wheel, foot pedal, and other gaming devices. They all perform functions similar to the mouse in that they allow the user to control a cursor or simulate generally real-time motion control. Contact your nearest game arcade for details.

Although the data glove also fits under the previous category, it is technically a lot more complex. It looks like a hand glove but contains a large number of sensors and has a data cable attached; though the latter is being replaced by means of infrared cordless data transmission. Not only does the data glove allow for full three-dimensional movement but it also senses the position of individual fingers, translating this into a grip. The glove is currently used in virtual reality simulators where the user moves around in an artificially rendered environment projected onto tiny LCD screens fitted into vision goggles. The computer generates various imaginary objects, which the user can “pick up” and manipulate by means of the glove. Advanced models even allow for tactile feedback by means of small pressure pockets built into the glove.

Optical scanners and readers

There are a number of different optical scanner technologies on the market.

- Optical Scanners use light-emitting devices to illuminate the printing on paper. Depending on how much light is reflected, a light-sensor determines the position and darkness (or color) of the markings on the paper. Special-purpose optical scanners are in use by postal services to read and / interpret hand-written postal codes. General-purpose scanners are used with personal computers to scan in images or text. These vary from handheld devices (see picture) to flatbed scanners which feed input documents one sheet at a time. A common use of optical scanners is the scanning of black-and-white or color images and pictures. When scanning text, it is necessary to load additional optical character recognition (OCR) software that converts the scanned raster-image of the text into the equivalent character symbols, so that they can be edited using word processing software.

-  Barcode scanners detect sequences of vertical lines of different widths, the ubiquitous barcode is found also on the back of this book. These scanners have become very popular with retailers due to the fact that all pre-packaged products are now required to have a product bar code on their packaging, following the standard laid down by the South African Article Numbering Association (SAANA). Libraries and video shops now also commonly use bar code scanners. They are more generally used for tracking and routing large numbers of physical items such as for asset inventory purposes in many larger organizations, postal items by the postal services and courier services, or for luggage handling by airlines.
- Optical mark readers are capable of reading dark marks on specially designed forms. The red multiple-choice answer sheets in use at many educational and testing institutions are a good example.

Other input devices

A magnetic card reader reads the magnetized stripe on the back of plastic credit-card size cards. These cards need to be pre-recorded following certain standards. Although the cards can hold only a tiny amount of information, they are very popular for access (door) control and financial transactions (ATMs and point-of-sale terminals).

Magnetic ink character recognition (MICR) uses a special ink (containing magnetizable elements) and a distinct font type. It is used mainly in the banking sector for the processing of cheques.

Touch-tone devices can use a voice telephone to contact computer-based switchboards or enter information directly into remote computers. Many corporate telephone help-lines rely on the customer pressing the touch-tone telephone buttons to route his/her call to the correct operator by selecting through a menu of possible options. South African banks also enable their clients to perform a number of banking transactions via telephone.

Digital cameras allow you to take pictures of physical objects directly in a digital, i.e. computer-readable, format. Relatively low-cost digital still-picture cameras are now available that capture images directly on electronic disk or RAM media instead of the traditional film. Apart from being very compact, most of these digital cameras can also interface directly with personal computers and are thus becoming a popular tool to capture pictures for e-mailing or loading on the worldwide Web.

Biometric devices are used to verify personal identity based on fingerprints, iris or retinal scanning, hand geometry, facial characteristics, etc. A scanning device is used to capture key measurements and compare them against a database of previously stored information. This type of authentication is becoming increasingly important in the control of physical access.

Finally, voice input devices are coming of age. Voice recognition has recently made a strong entry into the market with the availability of low-cost systems that work surprisingly well with today's personal computers. These systems allow for voice control of most standard applications (including the operating system). With voice control, the computer recognizes a very limited number (50 or less) of frequently used, programmable system commands ("save", "exit", "print"…) from a variety of users. In fact, these systems are not only used for the interface of computer programs; they are also slowly making an appearance in consumer appliances, novelty items, and even motor cars!

Much more difficult to achieve than voice control is true voice dictation used to dictate e.g. a letter to your word processor. The difficulty is that the computer must not only distinguish between many tens of thousands of possible words but it must also recognize the almost unnoticeable breaks in between words, different accents, and intonations. Therefore, voice dictation typically requires a user to train the voice recognition software by reading standard texts aloud. Nevertheless, for personal purposes and slow typists, voice recognition is rapidly becoming a viable alternative to the keyboard.

Output Devices

The final stage of information processing involves the use of output devices to transform computer-readable data back into an information format that can be processed by humans. As with input devices, when deciding on an output device you need to consider what sort of information is to be displayed, and who is intended to receive it.

One distinction that can be drawn between output devices is that of hardcopy versus softcopy devices. Hardcopy devices (printers) produce a tangible and permanent output whereas softcopy devices (display screens) present a temporary, fleeting image.

Display screens

The desk-based computer screen is the most popular output device. The standard monitor works on the same principle as the normal TV tube: a “ray” gun fires electrically charged particles onto a specially coated tube (hence the name Cathode-Ray Tube or CRT). Where the particles hit the coating, the “coating” is being “excited” and emits light. A strong magnetic field guides the particle stream to form the text or graphics on your familiar monitor.

CRTs vary substantially in size and resolution. Screen size is usually measured in inches diagonally across from corner to corner and varies from as little as 12 or 14 inches for standard PCs, to as much as 40+ inches for large demonstration and video-conferencing screens. The screen resolution depends on a number of technical factors.

A technology that has received much impetus from the fast-growing laptop and notebook market is the *liquid crystal display* (LCD). LCDs have matured quickly, increasing in resolution, contrast, and color quality. Their main advantages are lower energy requirements and their thin, flat size. Although alternative technologies are already being explored in research laboratories, they currently dominate the “flat display” market.

Organic light-emitting diodes (OLED) can generate brighter and faster images than LED technology and require thinner screens, but they have less stable color characteristics, making them more suitable for cellular telephone displays than for computers.

Another screen-related technology is the *video projection unit*. Originally developed for the projection of video films, the current trend towards more portable LCD-based lightweight projectors is fuelled by the need for computer-driven public presentations. Today’s units fit easily into a small suitcase and project a computer presentation in very much the same way a slide projector shows a slide presentation. They are rapidly replacing the flat transparent LCD panels that needed to be placed on top of an overhead projection unit. Though the LCD panels are more compact, weigh less, and are much cheaper, their image is generally of much poorer quality and less bright.

Printers and plotters

Printers are the most popular output device for producing permanent, paper-based computer output. Although they are all hardcopy devices, a distinction can be made between impact and non-impact printers. With impact printers, a hammer or needle physically hits an inked ribbon to leave an ink impression of the desired shape on the paper. The advantage of the impact printer is that it can produce more than one simultaneous copy by using carbon or chemically-coated paper. Non-impact printers, on the other hand, have far fewer mechanically moving parts and are therefore much quieter and tend to be more reliable.

The following are the main types of printers currently in use.

- *Dot-matrix printers* used to be the familiar low-cost printers connected to many personal computers. The print head consists of a vertical row of needles each of which is individually controlled by a magnet. As the print head moves horizontally across the paper, the individual needles strike the paper (and ribbon in between) as directed by the control mechanism to produce text characters or graphics. A close inspection of a dot-matrix printout will reveal the constituent dots that make up the text. Although it is one of the cheapest printer options, its print quality is generally much lower than that of laser and ink-jet printers. However, today’s models are quick and give a much better quality by increasing the number of needles.
- *Laser printers* are quickly growing in market share. They work on the same principle as the photocopier. A laser beam, toggled on and off very quickly, illuminates selected areas on a photosensitive drum, where the light is converted into electrical charge. As the drum rotates into a “bed” of carbon particles (“toner”) with the opposite charge, these particles will adhere to the drum. The blank paper is then pressed against the drum so that the particles “rub off onto the paper sheet. The sheet then passes through a high-temperature area so that the carbon particles are permanently fused onto the paper. Current high-end laser printers can cope with extremely large printing volumes, as is required e.g. by banks to print their millions of monthly account statements. Laser technology continues to develop in tandem with photocopier technology. Laser printers can now handle color printing, double-sided printing, or combine with mail equipment to perforate, fold, address and seal automatically into envelopes. At the lower end of the scale are the low-cost “personal” laser printers, which give a very good printing quality at a relatively modest cost.
- *Thermal printers* use heat to print. The older thermal printers used heat-sensitive paper, similar to the special fax paper. A slight heat or pressure will leave a darker area. This produced very cheap but low-quality output. Currently, thermal-printing technology is used mainly for high-quality color printing. These new thermal printers use colored wax sticks and melt the wax onto the paper. Although they are slower than competing color laser and inkjet technologies, they give a much more vibrant, color-saturated image.

-  *Inkjet printers* are probably the most popular low-cost printing technology. Liquid ink is squirted onto the paper in the form of tiny droplets. These printers are about the same price as dot-matrix printers, albeit more expensive in terms of consumables. Their quality is close to that of the laser printers. Their great advantage is that the printers can easily be adapted to use colored ink, thus making popular color printers.
- *Plotters* are mainly used for engineering and architectural drawings. A plotter consists of one (or several—in the case of color plotters) pen(s) affixed to an arm. As the arm moves across the sheet of paper, the pen draws lines onto the paper. It is ideal for line drawings such as plans, especially in cases where the paper size exceeds that which can be accommodated by the other types of printers.
- *Chain and line printers* are still popular in mainframe environments for the quick production of large volumes of internal printing. The line printer consists of a horizontal, rotating "drum" with 132 cylinders, each containing a full character set. As the 132-column wide paper moves up past the drum, a line at a time, each one of the 132 hammers on the other side of the paper strikes at the exact moment that the corresponding cylinder "shows" the correct character. The hammer hits the drum (and ink ribbon) and leaves an imprint of the character on the paper. The chain printer works on the same principle, but uses a horizontally rotating chain with engraved characters, instead of a drum. As anyone with some working experience in a large organization knows, the print quality of these "computer printouts" is not very high.

Figure 4 compares the various output devices in terms of a number of characteristics.

Figure 4: Comparison of output devices

Device	Technology	Quality	Speed	Duplicates?	Graphics?	Fonts?	Colour?
CRT	softcopy	high	very fast	n/a	yes	yes	yes
LCD	softcopy	fair	very fast	n/a	yes	yes	yes
Plotter	hardcopy	fair	slow	no	yes	yes	yes
Chain/line printer	hardcopy	low	very fast	yes	no	no	no
Laser printer	hardcopy	high	fast/fair	no	yes	yes	yes
Dot-Matrix printer	hardcopy	fair	fast/fair	yes	yes	yes	some
Inkjet printer	hardcopy	good	fair	no	yes	yes	yes

Audio-output devices

A type of output that is becoming increasingly popular is different types of audio output.

- *Sound output* is required by most multimedia applications and sophisticated games. The sound card in many of today's personal computers synthesizes sound by drawing from a library of stored sounds, essentially using the same process as found in music keyboards. More advanced multimedia workstations are equipped for full stereo multi-channel surround sound and easily surpass many a modern hi-fi system in cabling and speaker complexity.
- *MIDI in/output*. Modern-day music production would be impossible without a vast array of electronic instruments and keyboards. These are typically controlled by a personal computer by means of Musical Instrument Digital Interface (MIDI), a common standard for linking, controlling, and processing electronic music.
- *Speech synthesis* is the production of speech-like output using an artificial voice. Although the lack of intonation still makes the voice sound artificial, the technology is reasonably mature and can be found anywhere from talking clocks and luxury cars to automated responses for telephonic directory inquiries.

Other Output Devices

Many other, extremely specialized input and output devices have been developed. Process control, for example, is a very specialized field but extremely important for automated factories (car manufacturing, canneries), continuous process environments (nuclear plants, refineries), or hazardous places (microbiological research laboratories, space exploration). For these applications, the computer relies on a multitude of sensors for its inputs: temperatures, speed, pressure, flow rates, weight, position, … These sensor inputs are then processed by the computers, which in turn control directly robot arms and other mechanical devices such as cutters, welding equipment, valves, switches, mixers, etc.

3.2: Summary

Summary

The main ***internal*** components of the system or device are Motherboard (system board), Memory module, Processor slot(s), processor chip,(central processing unit, or CPU, Adapter card slot(s) or expansion card slot(s), Adapter card (expansion card), Internal drive bay, Hard drive, and Power supply.

The main ***external*** components are ***Bays*** are used to house CD and DVD drives, ***Slots*** are used to insert modem cards, wireless-access cards, and other PC cards that may not be internal to the system unit, and ***Ports*** are used to connect additional input, output, storage, and communication devices.

CHAPTER OVERVIEW

4: Software Component of an Information System

Software and hardware cannot function without each other. Without software, hardware is useless. Without hardware, the software has no hardware to run on. This chapter discusses the types of software, their purpose, and how they support different hardware devices, individuals, groups, and organizations.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Define the term software;
- Describe the two primary categories of software;
- Describe the role ERP software plays in an organization;
- Describe the process to write a computer program;
- Describe cloud computing and its advantages and disadvantages for use in an organization;
- Define the term open-source and identify its primary characteristics; and
- Determine what operating system you have;
- Learn to manage files.

4.1: Computer Software

Introduction to Software

The second component of an information system is software. Software is the means to take a user's data and process it to perform its intended action. Software translates what users want to do into a set of instructions that tell the hardware what to do. A set of instructions is also called a computer program. For example, when a user presses the letter 'A" key on the keyboard when using a word processing app, it is the word processing software that tells the hardware that the user pressed the key 'A' on the keyboard and fetches the image of the letter A to display on the screen as feedback to the user that the user's data is received correctly.

Software is created through the process of programming. We will cover the creation of software in this chapter. In essence, hardware is the machine, and software is the intelligence that tells the hardware what to do. Without software, the hardware would not be functional.

Types of Software

The software component can be broadly divided into two categories: system software and application software.

The system software is a collection of computer programs that provide a software platform for other software programs. It also insulates the hardware's specifics from the applications and users as much as possible by managing the hardware and the networks. It consists of

1. Operating System
2. Utilities

Application software is a computer program that delivers a specific activity for the users (i.e., create a document, draw a picture). It can be for either

1. a general-purpose (i.e., Microsoft Word, Google doc) or
2. for a particular purpose (i.e., weather forecast, CAD engineering)

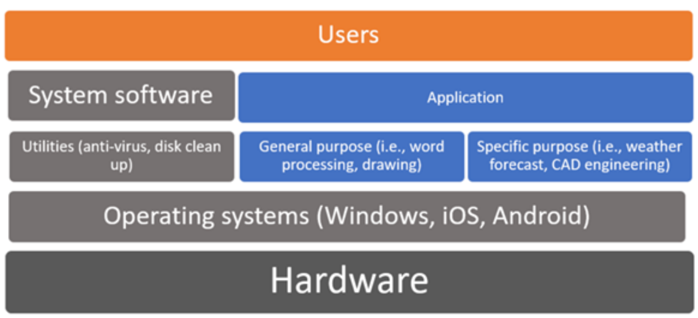


Figure 4.1.1 : Overview of software types. Image by Ly-Huong Pham is licensed under CC BY-NC

System Software

Operating Systems

The operating system provides several essential functions, including:

1. Managing the hardware resources of the computer
2. Providing the user-interface components
3. Providing a platform for software developers to write applications.

All computing devices run an operating system (OS), a key component of the system software. An OS is a set of programs that coordinate hardware components and other programs and acts as an interface with application software and networks.

Early personal-computer operating systems were simple by today's standards; they did not provide multitasking and required the user to type commands to initiate an action. The amount of memory that early operating systems could handle was limited as well, making large programs impractical to run. The most popular of the early operating systems was IBM's Disk Operating System, or DOS, which was actually developed for them by Microsoft.

For personal computers, some of the most popular operating systems today are Microsoft's Windows, Apple's OS X, Chrome, and different versions of Linux. Figure 5.1.2 displays how the operating system accepts input from various input devices such as a mouse, a keyboard, a digital pen, or a speech recognition, outputs to various output devices such as screen monitor or a printer; acts an intermediary between applications and apps, and access the internet via network devices such as a router or a web server.

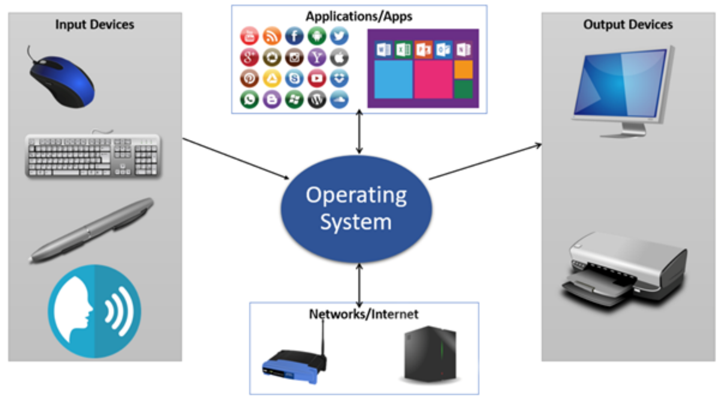


Figure 4.1.2 : Operating System Role. Image by Ly-Huong T. Pham is licensed by CC BY NC

In 1984, Apple introduced the Macintosh computer, featuring an operating system with a graphical user interface, now known as macOS. Apple has different names for its OS running on different devices such as iOS, iPadOS, watchOS, and tvOS.

In 1986, as a response to Apple, Microsoft introduced the Microsoft Windows Operating Systems, commonly known as Windows, as a new graphical user interface for their then command-based operating system, known as MS-DOS, which was developed for IBM's Disk Operating System or IBM-DOS. By the 1990s, Windows dominated the desktop personal computers market as the top OS and overtaken Apple's OS.

Since 1990, both Apple and Microsoft have released many new versions of their operating systems, with each release adding the ability to process more data at once and access more memory. Features such as multitasking, virtual memory, and voice input have become standard features of both operating systems.

A third personal-computer operating system family that is gaining in popularity is Linux. Linux is a version of the Unix operating system that runs on a personal computer. Unix is an operating system used primarily by scientists and engineers on larger minicomputers. These computers, however, are costly, and software developer Linus Torvalds wanted to find a way to make Unix run on less expensive personal computers: Linux was the result. Linux has many variations and now powers a large percentage of web servers in the world. It is also an example of open-source software, a topic we will cover later in this chapter.

Figure 4.1.3 : Tux, Linux's Mascot. Image by lewing@isc.tamu.edu Larry Ewing and The GIMP is licensed under Creative Commons CC0 1.0 Universal Public Domain Dedication

Smartphones and tablets run operating systems as well, such as Apple's iOS^{TM}, Google's Android (introduced in 2007) Microsoft's Windows Mobile, and Blackberry. It is based on the Linux kernel, and a consortium of developers developed other open-source software. Android quickly became the top OS for mobile devices and overtook Microsoft.

Operating systems have continuously improved with more and more features to increase speed and performance to process more data at once and access more memory. Features such as multitasking, virtual memory, and voice input have become standard features of both operating systems.

All computing devices run an operating system, as shown in the below table. The most popular operating systems are Microsoft's Windows, Apple's operating system, and different Linux versions for personal computers. Smartphones and tablets run operating systems as well, such as Apple's iOS and Google's Android and Chrome.

Computing devices and Operating system

Operating Systems	**Desktop**	**Mobile**
Microsoft Windows	Windows 10	Windows 10
Apple OS	Mac OS	iOS
Various versions of Linux	Ubuntu	Android (Google)

According to netmarketshare.com (2020), from August 2019 to August 2020, Windows still retains the desktop's dominant position with over 87% market share. Still, it is losing in the mobile market share, to Android with over 70% market share, followed by Apple's iOS with over 28% market share.

Sidebar: Why Is Microsoft Software So Dominant in the Business World?

Almost all businesses used IBM mainframe computers back in the 1960s and 1970s. These same businesses shied away from personal computers until IBM released the PC in 1981. Initially, business decisions were low-risk decisions since IBM was dominant, a safe choice. Another reason might be that once a business selects an operating system as the standard solution, it will invest in additional software, hardware, and services built for this OS. The switching cost to another OS becomes a hurdle both financially and for the workforce to be retrained.

Utility

Utility software includes software that is specific-purposed and focused on keeping the infrastructure healthy. Examples include antivirus software to scan and stop computer viruses and disk desegmentation software to optimize files' storage. Over time, some of the popular utilities were absorbed as features of an operating system.

Application or App Software

The second major category of software is application software. While system software focuses on running the computers, application software allows the end-user to accomplish some goals or purposes. Examples include word processing, photo editor, spreadsheet, or a browser. Applications software are grouped in many categories, including:

- Killer app
- Productivity
- Enterprise
- Mobile

The "Killer" App

When a new type of digital device is invented, there are generally a small group of technology enthusiasts who will purchase it just for the joy of figuring out how it works. A "killer" application runs only on one OS platform and becomes so essential that many people will buy a device on that OS platform just to run that application. For the personal computer, the killer application was the spreadsheet. In 1979, VisiCalc, the first personal-computer spreadsheet package, was introduced. It was an immediate hit and drove sales of the Apple II. It also solidified the value of the personal computer beyond the relatively small circle of technology geeks. When the IBM PC was released, another spreadsheet program, Lotus 1-2-3, was the killer app for business users. Today, Microsoft Excel dominates as the spreadsheet program, running on all the popular operating systems.

Figure 4.1.4:*VisiCalc. Image by Gortu is licensed under Public Domain*

Productivity Software

Along with the spreadsheet, several other software applications have become standard tools for the workplace. These applications, called productivity software, allow office employees to complete their daily work. Many times, these applications come packaged together, such as in Microsoft's Office suite. Here is a list of these applications and their basic functions:

- Word processing: This class of software provides for the creation of written documents. Functions include the ability to type and edit text, format fonts and paragraphs, and add, move, and delete text throughout the document. Most modern word-processing programs also have the ability to add tables, images, voice, videos, and various layout and formatting features to the document. Word processors save their documents as electronic files in a variety of formats. The most popular word-processing package is Microsoft Word, which saves its files in the Docx format. This format can be read/written by many other word-processor packages or converted to other formats such as Adobe's PDF.
- Spreadsheet: This class of software provides a way to do numeric calculations and analysis. The working area is divided into rows and columns, where users can enter numbers, text, or formulas. The formulas make a spreadsheet powerful, allowing the user to develop complex calculations that can change based on the numbers entered. Most spreadsheets also include the ability to create charts based on the data entered. The most popular spreadsheet package is Microsoft Excel, which saves its files in the XLSX format. Just as with word processors, many other spreadsheet packages can read and write to this file format.
- Presentation: This software class provides for the creation of slideshow presentations that can be shared, printed, or projected on a screen. Users can add text, images, audio, video, and other media elements to the slides. Microsoft's PowerPoint remains the most popular software, saving its files in PPTX format.
- Office Suite: Microsoft popularized the idea of the office-software productivity bundle with their release of Microsoft Office. Some office suites include other types of software. For example, Microsoft Office includes Outlook, its e-mail package, and OneNote, an information-gathering collaboration tool. The professional version of Office also includes Microsoft Access, a database package. (Databases are covered more in chapter 4.) This package continues to dominate the market, and most businesses expect employees to know how to use this software. However, many competitors to Microsoft Office exist and are compatible with Microsoft's file formats (see table below). Microsoft now has a cloud-based version called Microsoft Office 365. Similar to Google Drive, this suite allows users to edit and share documents online utilizing cloud-computing technology.

Category Suite:	**Word Processing**	**Spreadsheet**	**Presentation**	**Other**
Microsoft Office	Word	Excel	PowerPoint	Outlook (email), Access (database), OneNote (information gathering) and many other
Apple iWork	pages	numbers	keynote	Integrates with iTunes, iCloud, and other Apple software

Category Suite:	Word Processing	Spreadsheet	Presentation	Other
OpenOffice	Writer	Calc	Impress	Base (database), Draw (drawing, Math (equations)
Google Drive	Document	Spreadsheet	Presentation	Gmail,(email), Forms (online form data collection, Draw (drawing) and many others.

Sidebar: "PowerPointed" to Death

As presentation software, specifically Microsoft PowerPoint, has gained acceptance as the primary method to formally present information in a business setting, the art of giving an engaging presentation is becoming rare. Many presenters now just read the bullet points in the presentation and immediately bore those in attendance who can already read it for themselves.

The real problem is not with PowerPoint as much as it is with the person creating and presenting. The book Presentation Zen by Garr Reynolds is highly recommended to anyone who wants to improve their presentation skills.

New opportunities have been presented to make presentation software more effective. One such example is Prezi. Prezi is a presentation tool that uses a single canvas for the presentation, allowing presenters to place text, images, and other media on the canvas and then navigate between these objects as they present.

Enterprise Software

As the personal computer proliferated inside organizations, control over the information generated by the organization began splintering. For example, the customer service department creates a customer database to track calls and problem reports. The sales department also creates a database to keep track of customer information. Which one should be used as the master list of customers? As another example, someone in sales might create a spreadsheet to calculate sales revenue, while someone in finance creates a different one that meets their department's needs. However, the two spreadsheets will likely come up with different totals for revenue. Which one is correct? And who is managing all this information? This type of example presents challenges to management to make effective decisions.

Enterprise Resource Planning

In the 1990s, the need to bring the organization's information back under centralized control became more apparent. The enterprise resource planning (ERP) system (sometimes just called enterprise software) was developed to bring together an entire organization in one software application. Key characteristics of an ERP include:

- An integrated set of modules: Each module serves different functions in an organization, such as Marketing, Sales, Manufacturing.
- A consistent user interface: An ERP is a software application that provides a common interface across all modules of the ERP and is used by an organization's employees to access information
- A common database: All users of the ERP edit and save their information from the data source. This means that there is only one customer database, there is only one calculation for revenue, etc.
- Integrated business processes: All users must follow the same business rules and process throughout the entire organization": ERP systems include functionality that covers all of the essential components of a business, such as how organizations track cash, invoices, purchases, payroll, product development, supply chain.

Figure 4.1.5: ERP Modules. Image by Shing Hin Yeung is licensed under CC-BY-SA

ERP systems were originally marketed to large corporations, given that they are costly. However, as more and more large companies began installing them, ERP vendors began targeting mid-sized and even smaller businesses. Some of the more well-known ERP systems include those from SAP, Oracle, and Microsoft.

To effectively implement an ERP system in an organization, the organization must be ready to make a full commitment, including the cost to train employees as part of the implementation.

All aspects of the organization are affected as old systems are replaced by the ERP system. In general, implementing an ERP system can take two to three years and several million dollars.

So why implement an ERP system? If done properly, an ERP system can bring an organization a good return on its investment. By consolidating information systems across the enterprise and using the software to enforce best practices, most organizations see an overall improvement after implementing an ERP. Business processes as a form of competitive advantage will be covered in chapter 9.

Customer Relationship Management

A customer relationship management (CRM) system is a software application designed to manage customer interactions, including customer service, marketing, and sales. It collects all data about the customers. The objectives of a CRM are:

- Personalize customer relationship to increase customer loyalty
- Improve communication
- Anticipate needs to retain existing or acquire new customers

Some ERP software systems include CRM modules. An example of a well-known CRM package in Salesforce

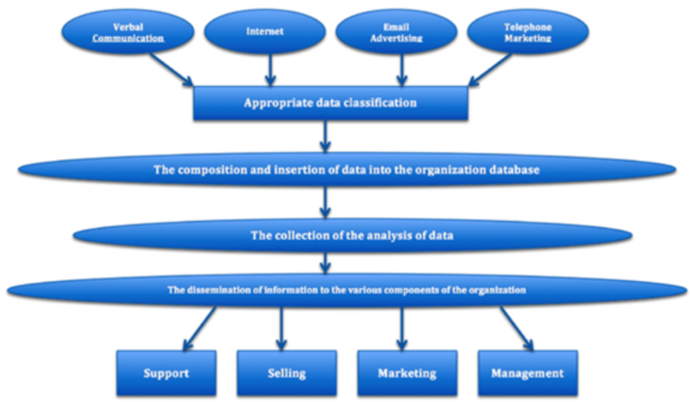


Figure 4.1.6: Components in the different types of CRM. Image by Bgrigorov is licensed under CC-BY-SA

Supply Chain Management

Many organizations must deal with the complex task of managing their supply chains. At its simplest, a supply chain is a linkage between an organization’s suppliers, its manufacturing facilities, and its products' distributors. Each link in the chain has a multiplying effect on the complexity of the process. For example, if there are two suppliers, one manufacturing facility, and two distributors, then there are 2 x 1 x 2 = 4 links to handle. However, if you add two more suppliers, another manufacturing facility, and two more distributors, then you have 4 x 2 x 4 = 32 links to manage.

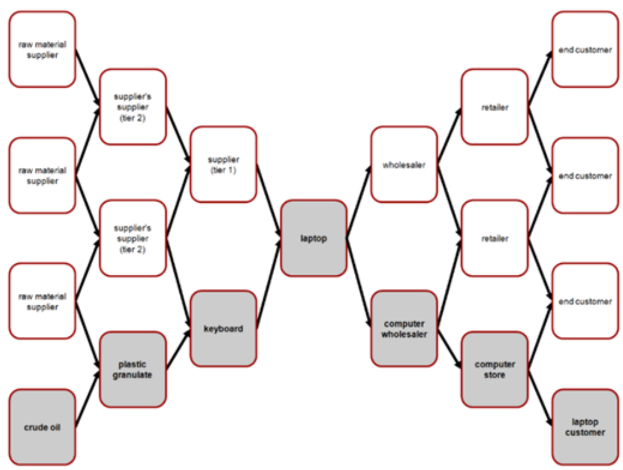


Figure 4.1.7: A supply and demand network. Image by Andreas Wieland is licensed under CC-BY-SA 3.0

A supply chain management (SCM) system manages the interconnection between these links and the products' inventory in their various development stages. The Association provides a full definition of a supply chain management system for Operations Management: “The design, planning, execution, control, and monitoring of supply chain activities to create net value, building a competitive infrastructure, leveraging worldwide logistics, synchronizing supply with demand, and measuring performance globally.” 2 Most ERP systems include a supply chain management module.

Mobile Software

A mobile application, commonly called a mobile app, is a software application programmed to run specifically on a mobile device such as smartphones and tablets.

Smartphones and tablets are becoming a dominant form of computing, with many more smartphones being sold than personal computers. This means that organizations will have to get smart about developing software on mobile devices to stay relevant. With the rise of mobile devices' adoption, the number of apps is exploding in the millions (Forbes.com, 2020), and there is an app for just about anything a user is looking to do. Examples include apps as a flashlight, a step counter, a plant identifier, and games.

Software Creation

We just discussed different types of software and now can ask: How is software created? If the software is the set of instructions that tells the hardware what to do, how are these instructions written? If a computer reads everything as one and zero, do we have to learn how to write software that way? Thankfully, another software type is written, especially for software developers to write system software and applications - called programming languages. The people who can program are called computer programmers or software developers.

Analogous to a human language, a programming language consists of keywords, comments, symbols, and grammatical rules to construct statements as valid instructions understandable by the computer to perform certain tasks. Using this language, a programmer writes a program (called the source code). Another software then processes the source code to convert the programming statements to a machine-readable form, the ones, and zeroes necessary to execute the CPU. This conversion process is often known as compiling, and the software is called the compiler. Most of the time, programming is done inside a programming environment; when you purchase a copy of Visual Studio from Microsoft; It provides the developers with an editor to write the source code, a compiler, and help for many of Microsoft's programming languages. Examples of well-known programming languages today include Java, PHP, and C's various flavors (Visual C, C++, C#.)

Figure 4.1.8: Convert a computer program to an executable. Image by Ly-Huong T. Pham is licensed under CC-BY-NC

Convert a computer program to an executable. Image by Ly-Huong T. Pham is licensed under CC-BY-NC

Thousands of programming languages have been created since the first programming language in 1883 by a woman named Ada Lovelace. One of the earlier English-like languages called COBOL has been in use since the 1950s to the present time in services that we still use today, such as payroll, reservation systems. The C programming language was introduced in the 1970s and remained a top popular choice. Some new languages such as C#, Swift are gaining momentum as well. Programmers select the best-matched language with the problem to be solved for a particular OS platform. For example, languages such as HTML and JavaScript are used to develop web pages.

It is hard to determine which language is the most popular since it varies. However, according to TIOBE Index, one of the companies that rank the popularity of the programming languages monthly, the top five in August 2020 are C, Java, Python, C++, and C# (2020). For more information on this methodology, please visit the TIOBE definition page. For those who wish to learn more about programming, Python is a good first language to learn because not only is it a modern language for web development, it is simple to learn and covers many fundamental concepts of programming that apply to other languages.

One person can write some programs. However, most software programs are written by many developers. For example, it takes hundreds of software engineers to write Microsoft Windows or Excel. To ensure teams can deliver timely and quality software with the least amount of errors, also known as bugs, formal project management methodologies are used.

Open-Source vs. Closed-Source Software

When the personal computer was first released, computer enthusiasts immediately banded together to build applications and solve problems. These computer enthusiasts were happy to share any programs they built and solutions to problems they found; this collaboration enabled them to innovate more quickly and fix problems.

As software began to become a business, however, this idea of sharing everything fell out of favor for some. When a software program takes hundreds of hours to develop, it is understandable that the programmers do not want to give it away. This led to a new business model of restrictive software licensing, which required payment for software to the owner, a model that is still dominant today. This model is sometimes referred to as closed source, as the source code remains private property and is not made available to others. Microsoft Windows, Excel, Apple iOS are examples of closed source software.

There are many, however, who feel that software should not be restricted. Like those early hobbyists in the 1970s, they feel that innovation and progress can be made much more rapidly if we share what we learn. In the 1990s, with Internet access connecting more and more people, the open-source movement gained steam.

Open-source software is software that has the source code available for anyone to copy and use. For non-programmers, it won't be of much use unless the compiled format is also made available for users to use. However, for programmers, the open-source movement has led to developing some of the world's most-used software, including the Firefox browser, the Linux operating system, and the Apache webserver.

Just about every type of commercial product has an open source equivalent. SourceForge.net lists over two hundred and thirty thousand such products[1]! Many of these products come with the installation tools, support utilities, and full documentation that make them difficult to distinguish from traditional commercial efforts (Woods, 2008). In addition to the LAMP products, some major examples include the following:

- Firefox—a Web browser that competes with Internet Explorer
- OpenOffice—a competitor to Microsoft Office
- Gimp—a graphic tool with features found in Photoshop
- Alfresco—collaboration software that competes with Microsoft Sharepoint and EMC's Documentum
- Marketcetera—an enterprise trading platform for hedge fund managers that competes with FlexTrade and Portware
- Zimbra—open source e-mail software that competes with Outlook server
- MySQL, Ingres, and EnterpriseDB—open source database software packages that each go head-to-head with commercial products from Oracle, Microsoft, Sybase, and IBM
- SugarCRM—customer relationship management software that competes with Salesforce.com and Siebel
- Asterix—an open source implementation for running a PBX corporate telephony system that competes with offerings from Nortel and Cisco, among others
- Free BSD and Sun's OpenSolaris—open source versions of the Unix operating sys

Some people are concerned that open-source software can be vulnerable to security risks since the source code is available. Others counter that because the source code is freely available, many programmers have contributed to open-source software projects, making the code less buggy and adding features, and fixing bugs much faster than closed-source software.

Many businesses are wary of open-source software precisely because the code is available for anyone to see. They feel that this increases the risk of an attack. Others counter that this openness decreases the risk because the code is exposed to thousands of programmers who can incorporate code changes to patch vulnerabilities quickly.

In summary, some benefits of the open-source model are:

- **The software is available for free:** Free alternatives to costly commercial code can be a tremendous motivator, particularly since conventional software often requires customers to pay for every copy used and to pay more for software that runs on increasingly powerful hardware. Big Lots stores lowered costs by as much as $10 million by finding viable OSS (Castelluccio, 2008) to serve their system needs. Online broker E*TRADE estimates that its switch to open source helped save over $13 million a year (King, 2008). And Amazon claimed in SEC filings that the switch to open source was a key contributor to nearly $20 million in tech savings (Shankland, et. al., 2001). Firms like TiVo, which use OSS in their own products, eliminate a cost spent either developing their own operating system or licensing similar software from a vendor like Microsoft.
- **Reliability:** There's a saying in the open source community, "Given enough eyeballs, all bugs are shallow" (Raymond, 1999). What this means is that the more people who look at a program's code, the greater the likelihood that an error will be caught and corrected. The open source community harnesses the power of legions of geeks who are constantly trawling OSS products, looking to squash bugs and improve product quality. And studies have shown that the quality of popular OSS products outperforms proprietary commercial competitors (Ljungberg, 2000). In one study, Carnegie Mellon University's Cylab estimated the quality of Linux code to be less buggy than commercial alternatives by a factor of two hundred (Castelluccio, 2008)!
- **Security:** OSS advocates also argue that by allowing "many eyes" to examine the code, the security vulnerabilities of open source products come to light more quickly and can be addressed with greater speed and reliability (Wheeler, 2003). High profile hacking contests have frequently demonstrated the strength of OSS products. In one well-publicized 2008 event, laptops

running Windows and Macintosh were both hacked (the latter in just two minutes), while a laptop running Linux remained uncompromised (McMillan, 2008). Government agencies and the military often appreciate the opportunity to scrutinize open source efforts to verify system integrity (a particularly sensitive issue among foreign governments leery of legislation like the USA PATRIOT Act of 2001) (Lohr, 2003). Many OSS vendors offer **security focused** (sometimes called *hardened*) versions of their products. These can include systems that monitor the integrity of an OSS distribution, checking file size and other indicators to be sure that code has not been modified and redistributed by bad guys who've added a back door, malicious routines, or other vulnerabilities.
- **Scalability:** Many major OSS efforts can run on everything from cheap commodity hardware to high-end supercomputing. **Scalability** allows a firm to scale from start-up to blue chip without having to significantly rewrite their code, potentially saving big on software development costs. Not only can many forms of OSS be migrated to more powerful hardware, packages like Linux have also been optimized to balance a server's workload among a large number of machines working in tandem. Brokerage firm E*TRADE claims that usage spikes following 2008 U.S. Federal Reserve moves flooded the firm's systems, creating the highest utilization levels in five years. But E*TRADE credits its scalable open source systems for maintaining performance while competitors' systems struggled (King, 2008).
- **Agility and Time to Market:** Vendors who use OSS as part of product offerings may be able to skip whole segments of the software development process, allowing new products to reach the market faster than if the entire software system had to be developed from scratch, in-house. Motorola has claimed that customizing products built on OSS has helped speed time-to-market for the firm's mobile phones, while the team behind the Zimbra e-mail and calendar effort built their first product in just a few months by using some forty blocks of free code (Guth, 2006).
- **The software source code is available:** It can be examined and reviewed before it is installed.
- **Quick Updates:** The large community of programmers who work on open-source projects leads to quick bug-fixing and feature additions.

Some benefits of the closed-source model are:

- Providing a financial incentive for software developers or companies
- Technical support from the company that developed the software.

Today there are thousands of open-source software applications available for download. An example of open-source productivity software is Open Office Suite. One good place to search for open-source software is sourceforge.net, where thousands of software applications are available for free download.

Why Give It Away? The Business of Open Source

Open source is a sixty-billion-dollar industry (Asay, 2008), but it has a disproportionate impact on the trillion-dollar IT market. By lowering the cost of computing, open source efforts make more computing options accessible to smaller firms. More reliable, secure computing also lowers costs for all users. OSS also diverts funds that firms would otherwise spend on fixed costs, like operating systems and databases, so that these funds can be spent on innovation or other more competitive initiatives. Think about Google, a firm that some estimate has over 1.4 million servers. Imagine the costs if it had to license software for each of those boxes!

Commercial interest in OSS has sparked an acquisition binge. Red Hat bought open source application server firm JBoss for $350 million. Novell snapped up SUSE Linux for $210 million. And Sun plunked down over $1 billion for open source database provider MySQL (Greenberg, 2008). And with Oracle's acquisition of Sun, one of the world's largest commercial software firms has zeroed in on one of the deepest portfolios of open source products.

But how do *vendors* make money on open source? One way is by selling support and consulting services. While not exactly Microsoft money, Red Hat, the largest purely OSS firm, reported half a billion dollars in revenue in 2008. The firm had two and a half million *paid* subscriptions offering access to software updates and support services (Greenberg, 2008). Oracle, a firm that sells commercial ERP and database products, provides Linux for free, selling high-margin Linux support contracts for as much as five hundred thousand dollars (Fortt, 2007). The added benefit for Oracle? Weaning customers away from Microsoft—a firm that sells many products that compete head-to-head with Oracle's offerings. Service also represents the most important part of IBM's business. The firm now makes more from services than from selling hardware and software (Robertson, 2009). And every dollar saved on buying someone else's software product means more money IBM customers can spend on IBM computers and services. Sun Microsystems was a leader in OSS, even before the Oracle acquisition bid. The firm has used OSS to drive advanced hardware sales, but the firm also sells proprietary products that augment its open source efforts. These products include special optimization, configuration management, and performance tools that can tweak OSS code to work its best (Preimesberger, 2008).

Here's where we also can relate the industry's evolution to what we've learned about standards competition in our earlier chapters. In the pre-Linux days, nearly every major hardware manufacturer made its own, incompatible version of the Unix operating system. These fractured, incompatible markets were each so small that they had difficulty attracting third-party vendors to write application software. Now, much to Microsoft's dismay, all major hardware firms run Linux. That means there's a large, unified market that attracts software developers who might otherwise write for Windows.

To keep standards unified, several Linux-supporting hardware and software firms also back the Linux Foundation, the nonprofit effort where Linus Torvalds serves as a fellow, helping to oversee Linux's evolution. Sharing development expenses in OSS has been likened to going in on a pizza together. Everyone wants a pizza with the same ingredients. The pizza doesn't make you smarter or better. So why not share the cost of a bigger pie instead of buying by the slice (Cohen, 2008)? With OSS, hardware firms spend less money than they would in the brutal, head-to-head competition where each once offered a "me too" operating system that was incompatible with rivals but offered little differentiation. Hardware firms now find their technical talent can be deployed in other value-added services mentioned above: developing commercial software add-ons, offering consulting services, and enhancing hardware offerings.

Linux on the Desktop?

While Linux is a major player in enterprise software, mobile phones, and consumer electronics, the Linux OS can only be found on a tiny fraction of desktop computers. There are several reasons for this. Some suggest Linux simply isn't as easy to install and use as Windows or the Mac OS. This complexity can raise the **total cost of ownership (TCO)** of Linux desktops, with additional end-user support offsetting any gains from free software. The small number of desktop users also dissuades third party firms from porting popular desktop applications over to Linux. For consumers in most industrialized nations, the added complexity and limited desktop application availability of desktop Linux just it isn't worth the one to two hundred dollars saved by giving up Windows.

But in developing nations where incomes are lower, the cost of Windows can be daunting. Consider the OLPC, Nicholas Negroponte's "one-hundred-dollar" laptop. An additional one hundred dollars for Windows would double the target cost for the nonprofit's machines. It is not surprising that the first OLPC laptops ran Linux. Microsoft recognizes that if a whole generation of first-time computer users grows up without Windows, they may favor open source alternatives years later when starting their own businesses. As a result, Microsoft has begun offering low-cost versions of Windows (in some cases for as little as seven dollars) in nations where populations have much lower incomes. Microsoft has even offered a version of Windows to the backers of the OLPC. While Microsoft won't make much money on these efforts, the low cost versions will serve to entrench Microsoft products as standards in emerging markets, staving off open source rivals and positioning the firm to raise prices years later when income levels rise.

MySQL: Turning a Ten-Billion-Dollars-a-Year Business into a One-Billion-Dollar One

Finland is not the only Scandinavian country to spawn an open source powerhouse. Uppsala Sweden's MySQL (pronounced "my sequel") is the "M" in the LAMP stack, and is used by organizations as diverse as FedEx, Lufthansa, NASA, Sony, UPS, and YouTube.

The "SQL" in name stands for the **structured query language**, a standard method for organizing and accessing data. SQL is also employed by commercial database products from Oracle, Microsoft, and Sybase. Even Linux-loving IBM uses SQL in its own lucrative DB2 commercial database product. Since all of these databases are based on the same standard, switching costs are lower, so migrating from a commercial product to MySQL's open source alternative is relatively easy. And that spells trouble for commercial firms. Granted, the commercial efforts offer some bells and whistles that MySQL doesn't yet have, but those extras aren't necessary in a lot of standard database use. Some organizations, impressed with MySQL's capabilities, are mandating its use on all new development efforts, attempting to cordon off proprietary products in legacy code that is maintained but not expanded.

Savings from using MySQL can be huge. The Web site PriceGrabber pays less than ten thousand dollars in support for MySQL compared to one hundred thousand to two hundred thousand dollars for a comparable Oracle effort. Lycos Europe switched from Oracle to MySQL and slashed costs from one hundred twenty thousand dollars a year to seven thousand dollars. And the travel reservation firm Sabre used open source products such as MySQL to slash ticket purchase processing costs by 80 percent (Lyons, 2004).

MySQL does make money, just not as much as its commercial rivals. While you can download a version of MySQL over the Net, the flagship product also sells for four hundred ninety-five dollars per server computer compared to a list price for Oracle that can climb as high as one hundred sixty thousand dollars. Of the roughly eleven million copies of MySQL in use, the company only gets paid for about one in a thousand (Ricadela, 2007). Firms pay for what's free for one of two reasons: (1) for MySQL service, and (2) for the right to incorporate MySQL's code into their own products (Kirkpatrick, 2004). Amazon, Facebook, Gap, NBC, and Sabre pay MySQL for support; Cisco, Ericsson, HP, and Symantec pay for the rights to the code (Ricadela, 2007). Top-level round-the-clock support for MySQL for up to fifty servers is fifty thousand dollars a year, still a fraction of the cost for commercial alternatives. Founder Marten Mickos has stated an explicit goal of the firm is "turning the $10-billion-a-year database business into a $1 billion one" (Kirkpatrick, 2004).

When Sun Microsystems spent over $1 billion to buy Mickos' MySQL in 2008, Sun CEO Jonathan Schwartz called the purchase the "most important acquisition in the company's history" (Shankland, 2008). Sun hoped the cheap database software could make the firm's hardware offerings seem more attractive. And it looked like Sun was good for MySQL, with the product's revenues growing 55 percent in the year after the acquisition (Asay, 2009).

But here's where it gets complicated. Sun also had a lucrative business selling hardware to support commercial ERP and database software from Oracle. That put Sun and partner Oracle in a relationship where they were both competitors and collaborators (the "coopetition" or "frenemies" phenomenon mentioned in Chapter 6 "Understanding Network Effects"). Then in spring 2009, Oracle announced it was buying Sun. Oracle CEO Larry Ellison mentioned acquiring the Java language was the crown jewel of the purchase, but industry watchers have raised several questions. Will the firm continue to nurture MySQL and other open source products, even as this software poses a threat to its bread-and-butter database products? Will the development community continue to back MySQL as the de facto standard for open source SQL databases, or will they migrate to an alternative? Or will Oracle find the right mix of free and fee-based products and services that allow MySQL to thrive while Oracle continues to grow? The implications are serious for investors, as well as firms that have made commitments to Sun, Oracle, and MySQL products. The complexity of this environment further demonstrates why technologists need business savvy and market monitoring skills and why business folks need to understand the implications of technology and tech-industry developments.

Legal Risks and Open Source Software: A Hidden and Complex Challenge

Open source software isn't without its risks. Competing reports cite certain open source products as being difficult to install and maintain (suggesting potentially higher total cost of ownership, or TCO). Adopters of OSS without support contracts may lament having to rely on an uncertain community of volunteers to support their problems and provide innovative upgrades. Another major concern is legal exposure. Firms adopting OSS may be at risk if they distribute code and aren't aware of the licensing implications. Some commercial software firms have pressed legal action against the users of open source products when there is a perceived violation of software patents or other unauthorized use of their proprietary code.

For example, in 2007 Microsoft suggested that Linux and other open source software efforts violated some two hundred thirty-five of its patents (Ricadela, 2007). The firm then began collecting payments and gaining access to the patent portfolios of companies that use the open source Linux operating system in their products, including Fuji, Samsung, and Xerox. Microsoft also cut a deal with Linux vendor Novell in which both firms pledged not to sue each other's customers for potential patent infringements.

Also complicating issues are the varying open source license agreements (these go by various names, such as GPL and the Apache License), each with slightly different legal provisions—many of which have evolved over time. Keeping legal with so many licensing standards can be a challenge, especially for firms that want to bundle open source code into their own products (Lacy, 2006). An entire industry has sprouted up to help firms navigate the minefield of open source legal licenses. Chief among these are products, such as those offered by the firm Black Duck, which analyze the composition of software source code and report on any areas of concern so that firms can honor any legal obligations associated with their offerings. Keeping legal requires effort and attention, even in an environment where products are allegedly "free." This also shows that even corporate lawyers had best geek-up if they want to prove they're capable of navigating a twenty-first-century legal environment.

Software Licenses

The companies or developers own the software they create. The software is protected by law either through patents, copyright, or licenses. It is up to the software owners to grant their users the right to use the software through the terms of the licenses.

For closed-source vendors, the terms vary depending on the price the users are willing to pay. Examples include single user, single installation, multi-users, multi-installations, per network, or machine.

They have specific permission levels for open-source vendors to grant using the source code and set the modified version conditions. Examples include free to distribute, remix, adapt for non-commercial use but with the condition that the newly revised source code must also be licensed under identical terms. While open-source vendors don't make money by charging for their software, they generate revenues through donations or selling technical support or related services. For example, Wikipedia is a widely popular and online free-content encyclopedia used by millions of users. Yet, it relies mainly on donations to sustain its staff and infrastructure.

There Are Now 8.9 Million Mobile Apps, And China Is 40% Of Mobile App Spending (2020, Feb 28). Retrieved September 4, 2020, from https://www.forbes.com/

Cloud Computing

Historically, for software to run on a computer, an individual copy of the software had to be installed on the computer, either from a disk or, more recently, after being downloaded from the Internet. The concept of "cloud" computing changes this model.

"The cloud" refers to applications, services, and data stored in data centers, server farms, and storage servers and accessed by users via the Internet. In most cases, the users don't know where their data is actually stored. Individuals and organizations use cloud computing.

You probably already use cloud computing in some forms. For example, if you access your email via your web browser, you are using a form of cloud computing. If you use Google Drive's applications, you are using cloud computing. Simultaneously, these are free versions of cloud computing, big business in providing applications and data storage over the web. Commercial and large applications can also exist on the cloud, such as the entire suite of CRM from Salesforce is offered via the cloud. Cloud computing is not limited to web applications: it can also be used for phone or video streaming services.

Advantages of Cloud Computing

- No software to install or upgrades to maintain.
- Available from any computer that has access to the Internet.
- Can scale to a large number of users easily.
- New applications can be up and running very quickly.
- Services can be leased for a limited time on an as-needed basis.
- Your information is not lost if your hard disk crashes or your laptop is stolen.
- You are not limited by the available memory or disk space on your computer.

Disadvantages of Cloud Computing

- You must have Internet access to use it. If you do not have access, you're out of luck.
- You are relying on a third party to provide these services.
- You don't know how your data is protected from theft or sold by your own cloud service provider.

Cloud computing can greatly impact how organizations manage technology. For example, why is an IT department needed to purchase, configure, and manage personal computers and software when all that is really needed is an Internet connection?

Using a Private Cloud

Many organizations are understandably nervous about giving up control of their data and applications using cloud computing. But they also see the value in reducing the need for installing software and adding disk storage to local computers. A solution to this problem lies in the concept of a private cloud. While there are various private cloud models, the basic idea is for the cloud service provider to rent a specific portion of their server space exclusive to a specific organization. The organization has full control over that server space while still gaining some of the benefits of cloud computing.

Cloud Emerging Technology/Current Trend

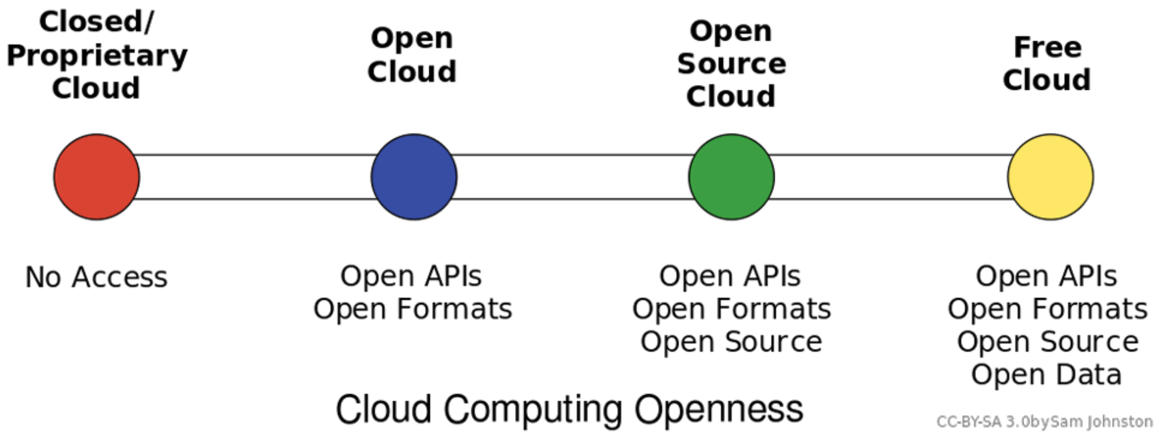


Figure 4.1.9: *Cloud Computing Styles*

Cloud computing can be loosely defined as the allocation of hardware and/or software under a service model (resources are assigned and consumed as needed). Typically, what we hear today referred to as cloud computing is the concept of business-to-business commerce revolving around "Company A" selling or renting their services to "Company B" over the Internet. A cloud can be public (hosted on a public internet, shared among consumers) or private (cloud concepts of provisioning and storage are applied to servers within a fire wall or internal network that is privately managed), and can also fall into some smaller subsets in between, as depicted in the graphic above.

Under Infrastructure as a Service (IaaS) computing model, which is what is most commonly associated with the term cloud computing, one or more servers with significant amounts of processing power, capacity, and memory, are configured through hardware and/or software methods to act as though they are multiple smaller systems that add up to their capacity. This is referred to as virtualizing, or virtual servers. These systems can be "right sized" where they only consume the resources they need on average, meaning many systems needing little resources can reside on one piece of hardware. When processing demands of one system expand or contract, resources from that server can be added or removed to account for the change. This is an alternative to multiple physical servers, where each would need the ability to serve not only the average but expected peak needs of system resources.

Software as a Service, Platform as a Service, and the ever-expanding list of "as-a-service" models follow the same basic pattern of balancing time and effort. Platforms as a service allow central control of end user profiles, and software as a service allows simplified (and/or automated) updating of programs and configurations. Storage as a service can replace the need to manually process backups and file server maintenance. Effectively, each "as-a-service" strives to provide the end user with an "as-good-if-not-better" alternative to managing a system themselves, all while trying to keep the cost of their services less than a self-managed solution.

Virtualization

The reduced costs and increased power of commodity hardware are not the only contributors to the explosion of cloud computing. The availability of increasingly sophisticated software tools has also had an impact. Perhaps the most important software tool in the cloud computing toolbox is **virtualization**. Virtualization is using software to create a virtual machine that simulates a computer with an operating system. For example, using virtualization, a single computer that runs Microsoft Windows can host a virtual machine that looks like a computer with a specific Linux-based OS. This ability maximizes the use of available resources on a single machine. Companies such as EMC provide virtualization software that allows cloud service providers to provision web servers to their clients quickly and efficiently. Organizations are also implementing virtualization to reduce the number of servers needed to provide the necessary services. For more detail on how virtualization works, see this informational page from VMWare.

Think of virtualization as being a kind of operating system for operating systems. A server running virtualization software can create smaller compartments in memory that each behave as a separate computer with its own operating system and resources. The most sophisticated of these tools also allow firms to combine servers into a huge pool of computing resources that can be allocated as needed (Lyons, 2008).

Virtualization can generate huge savings. Some studies have shown that on average, conventional data centers run at 15 percent or less of their maximum capacity. Data centers using virtualization software have increased utilization to 80 percent or more (Katz, 2009).This increased efficiency means cost savings in hardware, staff, and real estate. Plus it reduces a firm's IT-based energy consumption, cutting costs, lowering its carbon footprint, and boosting "green cred" (Castro, 2007). Using virtualization, firms can buy and maintain fewer servers, each running at a greater capacity. It can also power down servers until demand increases require them to come online.

While virtualization is a key software building block that makes public cloud computing happen, it can also be used in-house to reduce an organization's hardware needs, and even to create a firm's own private cloud of scalable assets. Bechtel, BT, Merrill Lynch, and Morgan Stanley are among the firms with large private clouds enabled by virtualization (Brodkin, 2008). Another kind of virtualization, **virtual desktops** allow a server to run what amounts to a copy of a PC—OS, applications, and all—and simply deliver an image of what's executing to a PC or other connected device. This allows firms to scale, back up, secure, and upgrade systems far more easily than if they had to maintain each individual PC. One game start-up hopes to remove the high-powered game console hardware attached to your television and instead put the console in the cloud, delivering games to your TV as they execute remotely on superfast server hardware. Virtualization can even live on your desktop. Anyone who's ever run Windows in a window on Mac OS X is using virtualization software; these tools inhabit a chunk of your Mac's memory for running Windows and actually fool this foreign OS into thinking that it's on a PC.

Interest in virtualization has exploded in recent years. VMware, the virtualization software division of storage firm EMC, was the biggest IPO of 2007. But its niche is getting crowded. Microsoft has entered the market, building virtualization into its server offerings. Dell bought a virtualization software firm for $1.54 billion. And there's even an open source virtualization product called Xen (Castro, 2007).

Virtualization Emerging Technology/Curent Trend

Server virtualization is the act of running multiple operating systems and other software on the same physical hardware at the same time, as we discussed in Cloud Computing. A hardware and/or software element is responsible for managing the physical system resources at a layer in between each of the operating systems and the hardware itself. Doing so allows the consolidation of physical equipment into fewer devices, and is most beneficial when the servers sharing the hardware are unlikely to demand resources at the same time, or when the hardware is powerful enough to serve all of the installations simultaneously.

The act of virtualizing is not just for use in cloud environments, but can be used to decrease the "server sprawl," or overabundance of physical servers, that can occur when physical hardware is installed on a one-to-one (or few-to-one) scale to applications and sites being served. Special hardware and/or software is used to create a new layer in between the physical resources of your computer and the operating system(s) running on it. This layer manages what each system sees as being the hardware available to it, and manages allocation of resources and the settings for all virtualized systems. Hardware virtualization, or the stand alone approach, sets limits for each operating system and allows them to operate independent of one another. Since hardware virtualization does not require a separate operating system to manage the virtualized system(s), it has the potential to operate faster and consume fewer resources than software virtualization. Software virtualization, or the host-guest approach, requires the virtualizing software to run on an operating system already in use, allowing simpler management to occur from the initial operating system and virtualizing program, but can be more demanding on system resources even when the primary operating system is not being used.

Ultimately, you can think of virtualization like juggling. In this analogy, your hands are the servers, and the balls you juggle are your operating systems. The traditional approach of hosting one application on one server is like holding one ball in each hand. If your hands are both "busy" holding a ball, you cannot interact with anything else without putting a ball down. If you juggle them, however, you can "hold" three or more balls at the same time. Each time your hand touches a ball is akin to a virtualized system needing resources, and having those resources allocated by the virtualization layer (the juggler) assigning resources (a hand), and then reallocating for the next system that needs them.

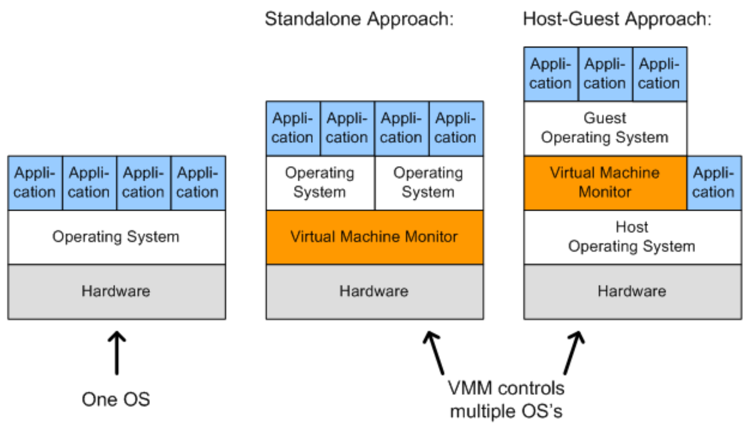


Figure 4.1.10: *Virtualization Styles*. (By Daniel Hirschbach [CC-BY-SA-2.0 Germany] via Wikimedia)

The addition of a virtual machine as shown above allows the hardware or software to see the virtual machine as part of the regular system. The monitor itself divides the resources allocated to it into subsets that act as their own computers.

Summary

The software gives the instructions that tell the hardware what to do. There are two basic categories of software: operating systems and applications. Operating systems provide access to the computer hardware and make system resources available. Application software is designed to meet a specific goal. Productivity software is a subset of application software that provides basic business functionality to a personal computer: word processing, spreadsheets, and presentations. An ERP system is a software application with a centralized database that is implemented across the entire organization. Cloud computing is a software delivery method that runs on any computer with a web browser and access to the Internet. Software is developed through a process called programming, in which a programmer uses a programming language to put together the logic needed to create the program. The software can be an open-source or a closed-source model, and users or developers are granted different licensing terms.

Learn more

Keywords, search terms: Cloud computing, virtualization,
virtual machines (VMs), software virtualization, hardware virtualization

Xen and the Art of Virtualization: http://li8-68.members.linode.com/~caker/xen/2003-xensosp.pdf

Virtualization News and Community: http://www.virtualization.net

Cloud Computing Risk Assessment: http://www.enisa.europa.eu/activities/risk-management/files/deliverables/cloud-computing-risk-assessment

Without formal legislation, judges and juries are placed in positions where they establish precedence by ruling on these issues, while having little guidance from existing law. As recently as March 2012 a file sharing case from 2007 reached the Supreme Court, where the defendant was challenging the constitutionality of a $222,000 USD fine for illegally sharing 24 songs on file sharing service Kazaa. This was the first case for such a lawsuit heard by a jury in the United States. Similar trials have varied in penalties up to $1.92 million US dollars, highlighting a lack of understanding of how to monetize damages. The Supreme Court denied hearing the Kazaa case, which means the existing verdict will stand for now. Many judges are now dismissing similar cases that are being brought by groups like the Recording Industry Association of America (RIAA10), as these actions are more often being seen as the prosecution using the courts as a means to generate revenue and not recover significant, demonstrable damages.

As these cases continue to move through courts and legislation continues to develop at the federal level, those decisions will have an impact on what actions are considered within the constructs of the law, and may have an effect on the contents or location of your site.

AutoAttribution

- **3.2: Types of Software** by Ly-Huong T. Pham, Tejal Desai-Naik, Laurie Hammond, & Wael Abdeljabbar, is licensed CC BY-NC-SA

4.2: Operating Systems

Operating System (OS)

An operating system (OS) is system software that manages computer hardware and software resources and provides common services for computer programs. Nearly every computer program requires an operating system to function. The two most common operating systems are Microsoft Windows and Apple's macOS. This course's main focus will be Windows 10 and 7.

Although this class will be focusing on Windows 10 and 7, the things you will learn in this module can be done by any version of Windows or macOS. If you are not running Windows 10 or 7, you can find directions online by searching for the task you are trying to do and the name of your operating system. (For example, you might search for "create folder windows vista.")

Windows vs. Mac

Apple logo

To identify your operating system, start by narrowing down which brand of OS you have. Often you can determine whether you are running Windows or macOS without even turning on your computer:

- Is there an Apple logo somewhere on your computer?
- Does your computer have a Command key (⌘) on the keyboard?
- Was your computer purchased at an Apple store?

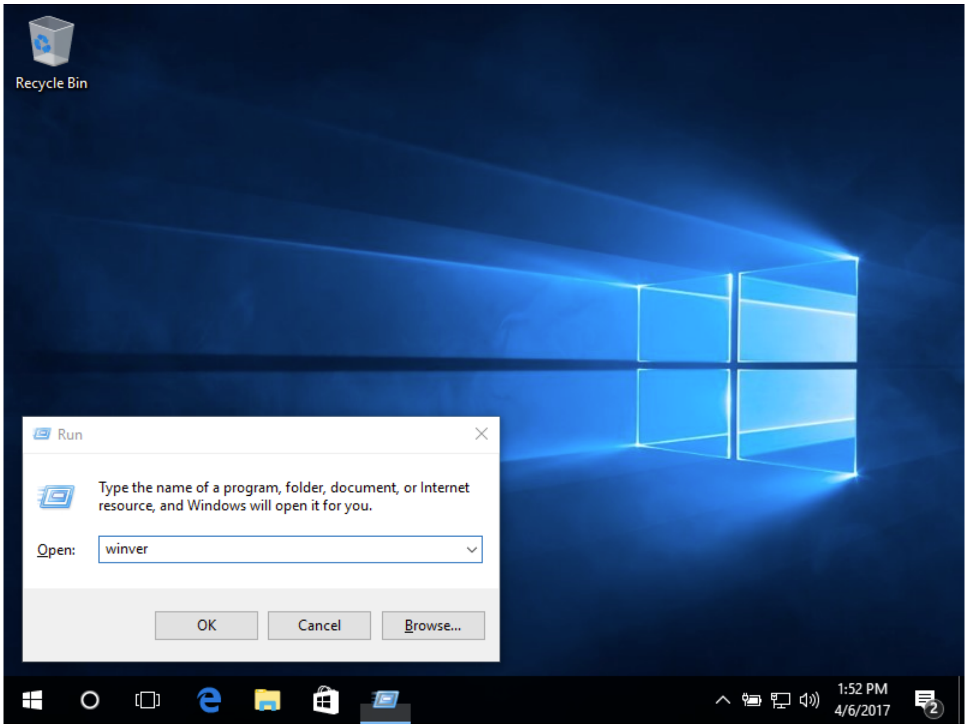
If the answer to any of those questions is *yes*, then you most likely have a computer that runs macOS.

Another good way to see whether you are running Windows or Mac is simply to turn the computer on. As the operating system boots up, it will display a logo. A Windows logo means you are running Windows, while an Apple logo means you are running macOS.

If you have determined that you are running Windows, there are many methods to determine your specific operating system, not all of which are listed here. **It doesn't matter which method you use**; what's important is figuring out which operating system you have. If one method doesn't work, try another.

Method 1

Press the Windows key (at the bottom of the keyboard, looks like four squares) and R key at the same time to open the Run dialog box. Type in "winver."

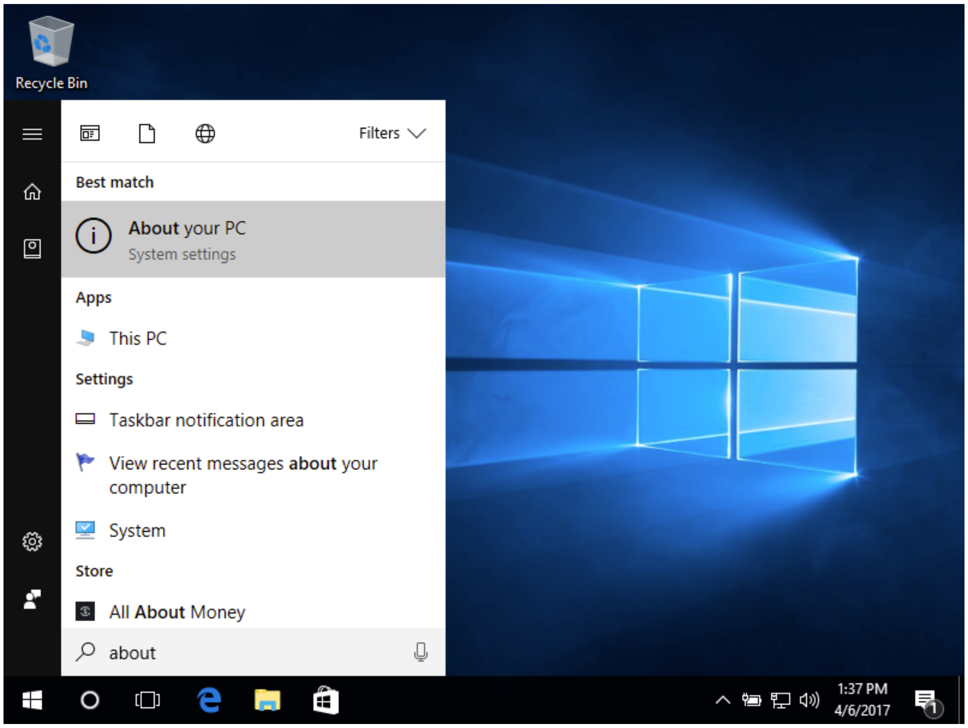


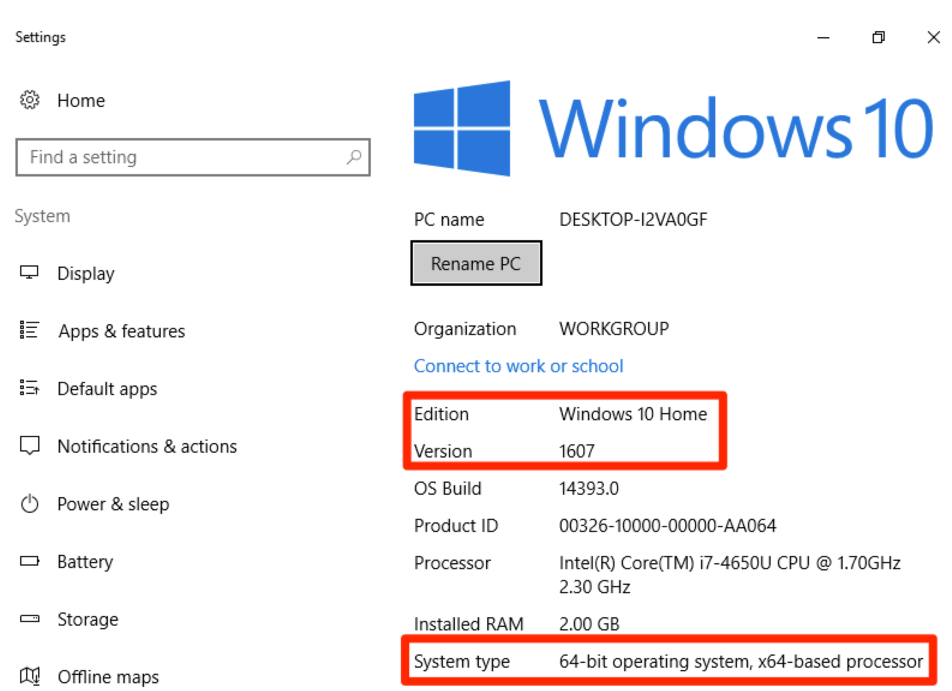
Click OK. A window should open called About Windows that tells you which operating system you are using.

This computer is using Windows 10.

Method 2

Open the Start menu by clicking the bottom left-hand corner of your screen. If you see a search text box, select it and then type "about" or "about your pc." If you do not see a search text box, then just start typing "about" or "about your pc." Select About your PC. If you do not see About your PC, try a different method.

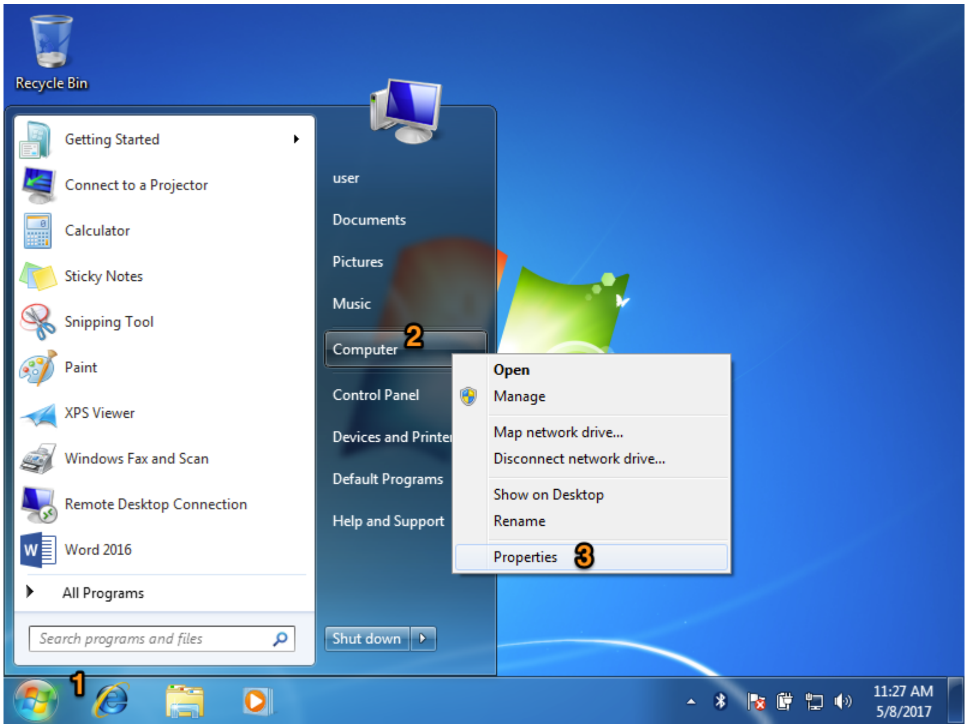


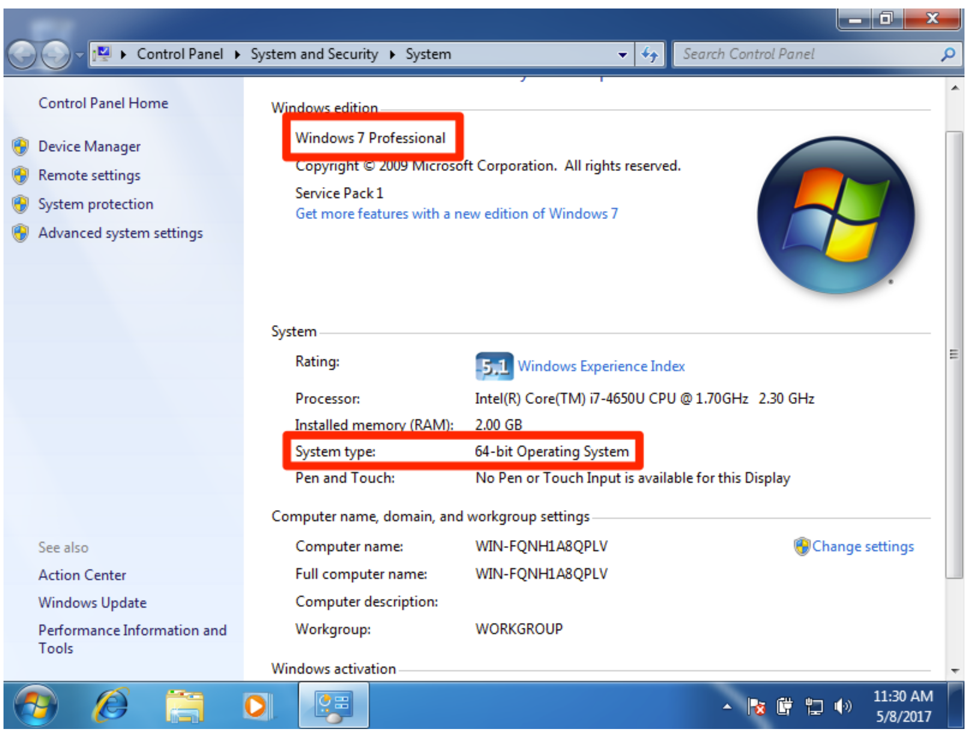


This computer is running Windows 10 Home Version 1607 with a 64-bit operating system.

Method 3

Open the Start menu by clicking the bottom left-hand corner of your screen. Right-click on Computer, then select Properties.

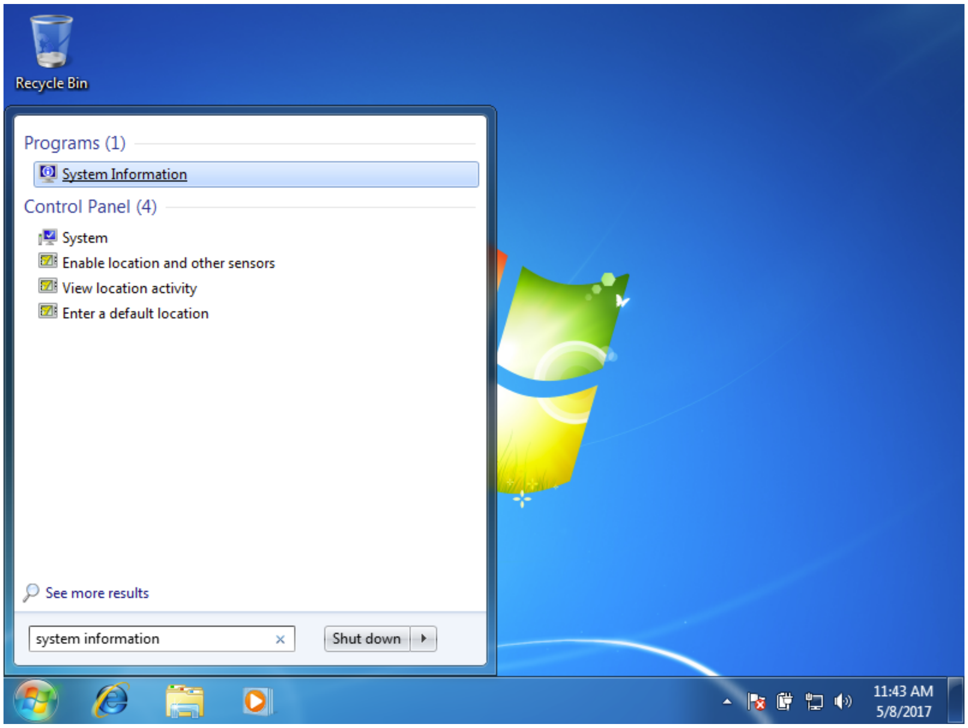


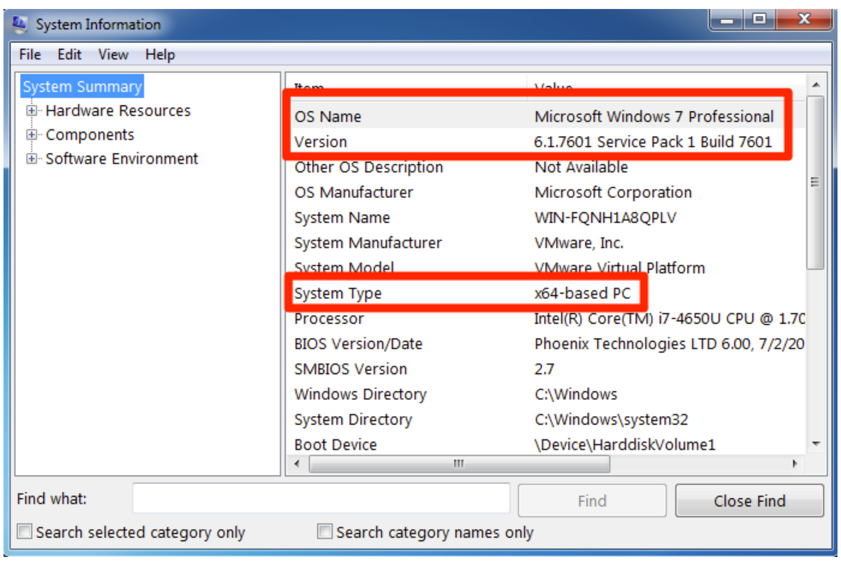


This computer is running Windows 7 Professional with a 64-bit operating system.

Method 4

Open the Start menu by clicking the bottom left-hand corner of your screen.If you see a search text box, select it and then type “System Information” in the search box. If you do not see a search text box, then just start typing “system” or “system information.” Select System Information under Programs. If you do not see System Information, try a different method.





This computer is running Windows 7 Professional with a 64-bit operating system.

Methods 2–4 also provide additional information on your computer that is useful for troubleshooting and determining what programs you will be able to run. For example, many software programs have a version for a x32-based processor and another for a x64-based processor. For the purposes of this course, you only need to know which edition of Windows you have.

Mac and Windows Operating Systems

Sidebar: Mac vs. Windows

Are you a Mac? Are you a PC? Ever since its introduction in 1984, users of the Apple Macintosh have been quite biased about their preference for the Macintosh operating system (now called OS X) over Microsoft's. When Microsoft introduced Windows, Apple sued Microsoft, claiming that they copied the "look and feel" of the Macintosh operating system. In the end, Microsoft successfully defended themselves.

Over the past few years, Microsoft and Apple have traded barbs with each other, each claiming to have a better operating system and software. While Microsoft has always had the larger market share (see sidebar), Apple has been the favorite of artists, musicians, and the technology elite. Apple also provides a lot of computers to elementary schools, thus gaining a following among the younger generation.

Sidebar: Why Is Microsoft Software So Dominant in the Business World?

If you've worked in the world of business, you may have noticed that almost all of the computers run a version of Microsoft's Windows operating system. Why is this? On almost all college campuses, you see a preponderance of Apple Macintosh laptops. In elementary schools, Apple reigns as well. Why has this not extended into the business world?

Almost all businesses used IBM mainframe computers back in the 1960s and 1970s. These same businesses shied away from personal computers until IBM released the PC in 1981. When executives had to make a decision about purchasing personal computers for their employees, they would choose the safe route and purchase IBM. The saying then was: "No one ever got fired for buying IBM." So over the next decade, companies bought IBM personal computers (or those compatible with them), which ran an operating system called DOS. DOS was created by Microsoft, so when Microsoft released Windows as the next iteration of DOS, companies took the safe route and started purchasing Windows.

Microsoft soon found itself with the dominant personal-computer operating system for businesses. As the networked personal computer began to replace the mainframe computer as the primary way of computing inside businesses, it became essential for Microsoft to give businesses the ability to administer and secure their networks. Microsoft developed business-level server products to go along with their personal computer products, thereby providing a complete business solution. And so now, the saying goes: "No one ever got fired for buying Microsoft."

Summary

The software gives the instructions that tell the hardware what to do. There are two basic categories of software: operating systems and applications. Operating systems provide access to the computer hardware and make system resources available. Application software is designed to meet a specific goal. Productivity software is a subset of application software that provides basic business functionality to a personal computer: word processing, spreadsheets, and presentations. An ERP system is a software application with a centralized database that is implemented across the entire organization. Cloud computing is a method of software delivery that runs on any computer that has a web browser and access to the Internet. Software is developed through a process called programming, in which a programmer uses a programming language to put together the logic needed to create the program. While most software is developed using a closed-source model, the open-source movement is gaining more support today.

4.3: File systems

Files and file systems

When a process completes (or crashes), any data stored in the main memory is lost. But data stored on a hard disk drive (HDD) or solid-state drive (SSD) is “persistent;” that is, it survives after the process completes, even if the computer shuts down.

Hard disk drives are complicated. Data is stored in blocks, which are laid out in sectors, which make up tracks, which are arranged in concentric circles on platters.

Solid-state drives are simpler in one sense because blocks are numbered sequentially, but they raise a different complication: each block can be written a limited number of times before it becomes unreliable.

As a programmer, you don’t want to deal with these complications. What you want is an appropriate abstraction of persistent storage hardware. The most common abstraction is called a “file system.”

Abstractly:

- A “file system” is a mapping from each file’s name to its contents. If you think of the names as keys, and the contents as values, a file system is a kind of key-value database (see https://en.Wikipedia.org/wiki/Key-value_database).
- A “file” is a sequence of bytes.

File names are usually strings, and they are usually “hierarchical”; that is, the string specifies a path from a top-level directory (or folder), through a series of subdirectories, to a specific file.

The primary difference between the abstraction and the underlying mechanism is that files are byte-based and persistent storage is block-based. The operating system translates byte-based file operations in the C library into block-based operations on storage devices. Typical block sizes are 1–8 KiB.

For example, the following code opens a file and reads the first byte:

```
FILE *fp = fopen("/home/downey/file.txt", "r");
char c = fgetc(fp);
fclose(fp);
```

When this code runs:

1. `fopen` uses the filename to find the top-level directory, called `/`, the subdirectory `home`, and the sub-subdirectory `downey`.
2. It finds the file named `file.txt` and “opens” it for reading, which means it creates a data structure that represents the file being read. Among other things, this data structure keeps track of how much of the file has been read, called the “file position”.

   In DOS, this data structure is called a File Control Block, but I want to avoid that term because in UNIX it means something else. In UNIX, there seems to be no good name for it. It is an entry in the open file table, so I will call it an OpenFileTableEntry.
3. When we call `fgetc`, the operating system checks whether the next character of the file is already in memory. If so, it reads the next character, advances the file position, and returns the result.
4. If the next character is not in memory, the operating system issues an I/O request to get the next block. Disk drives are slow, so a process waiting for a block from disk is usually interrupted so another process can run until the data arrives.
5. When the I/O operation is complete, the new block of data is stored in memory, and the process resumes. It reads the first character and stores it as a local variable.
6. When the process closes the file, the operating system completes or cancels any pending operations, removes data stored in memory, and frees the OpenFileTableEntry.

The process for writing a file is similar, but there are some additional steps. Here is an example that opens a file for writing and changes the first character.

```
FILE *fp = fopen("/home/downey/file.txt", "w");
fputc('b', fp);
```

```
    fclose(fp);
```

When this code runs:

1. Again, `fopen` uses the filename to find the file. If it does not already exist, it creates a new file and adds an entry in the parent directory, `/home/downey` .
2. The operating system creates an OpenFileTableEntry that indicates that the file is open for writing, and sets the file position to 0.
3. `fputc` attempts to write (or re-write) the first byte of the file. If the file already exists, the operating system has to load the first block into memory. Otherwise it allocates a new block in memory and requests a new block on disk.
4. After the block in memory is modified, it might not be copied back to the disk right away. In general, data written to a file is "buffered", which means it is stored in memory and only written to disk when there is at least one block to write.
5. When the file is closed, any buffered data is written to disk and the OpenFileTableEntry is freed.

To summarize, the C library provides the abstraction of a file system that maps from file names to streams of bytes. This abstraction is built on top of storage devices that are actually organized in blocks.

Everything is a file

The file abstraction is really a "stream of bytes" abstraction, which turns out to be useful for many things, not just file systems.

One example is the UNIX pipe, which is a simple form of inter-process communication. Processes can be set up so that output from one process is taken as input into another process. For the first process, the pipe behaves like a file open for writing, so it can use C library functions like `fputs` and `fprintf` . For the second process, the pipe behaves like a file open for reading, so it uses `fgets` and `fscanf` .

Network communication also uses the stream of bytes abstraction. A UNIX socket is a data structure that represents a communication channel between processes on different computers (usually). Again, processes can read data from and write data to a socket using "file" handling functions.

Reusing the file abstraction makes life easier for programmers, since they only have to learn one API (application program interface). It also makes programs more versatile, since a program intended to work with files can also work with data coming from pipes and other sources

4.4: Downloading Files

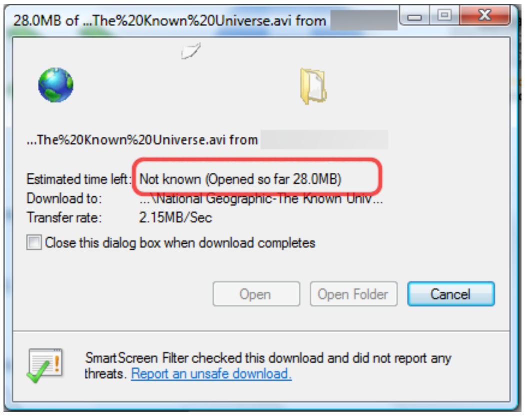


Figure 4.4.1: This Photo by Unknown Author is licensed under CC BY-SA

Links don't always go to another website. In some cases, they point to a file that can be downloaded, or saved, to your computer. If you click a link to a file, it may download automatically, but sometimes it just opens within your browser instead of downloading. To prevent it from opening in the browser, you can right-click the link and select Save link as (different browsers may use slightly different wording, like Save target as). Files downloaded from the internet are often downloaded to your download folder, which is a quick access folder accessible on the left navigation pane of windows explorer. Sometimes you may want to save an image from a website to your computer. To do this, right-click the image and select Save image as (or Save picture as). Make sure you trust the file that you downloaded to your computer, as downloading any file or program to your computer contains security risks to your computer (links).

4.5: File Management

Ribbon

In Windows 10, the File Explorer features a new ribbon toolbar, like the one featured in the recent versions of Microsoft Office. This ribbon contains buttons and commands for the most common tasks.

Home Tab of Windows 10 Ribbon

Figure 4.5.1: This Photo by Unknown Author is licensed under CC BY-SA

The Ribbon features four tabs, each with different commands. Among the tasks, you can perform from the Home tab is (Tutorialspoint.com, n.d.) −

- Copying and pasting files and folders from one place to another.
- Moving files and folders to another location.
- Copying files and folders to another location.
- Deleting a file or folder permanently or sending it to the Recycle Bin.
- Renaming a file or folder.
- Creating a new folder or other new items.
- Verifying or modifying the Properties of a document or folder.
- Opening a file or folder.
- Different options to select one or various files and folders.

Share Tab of the Windows 10 Ribbon

Figure 4.5.2: This Photo by Unknown Author is licensed under CC BY-SA

The Share ribbon gives you different options to share your files and folders. For example: (Tutorialspoint.com, n.d.)

- E-mailing or messaging a file.
- Compressing ("Zip") a folder to take less space.
- Printing or faxing documents.
- Sharing with other users or networks.

View Tab of the Windows 10 Ribbon

Figure 4.5.3: This Photo by Unknown Author is licensed under CC BY-SA

The View ribbon allows you to change the way Windows displays your files and folders. Some of the changes you can make here are (Tutorialspoint.com, n.d.) −

- Adding additional panes to show a preview or details of your files.
- Changing the layout of the files and folders from icons to list, and others.
- Sorting and arranging the contents of your folder.
- Hiding selected folders or files.

File Tab of the Windows 10 Ribbon

Figure 4.5.4: This Photo by Unknown Author is licensed under CC BY-SA

The File tab opens a menu with different options like

- Opening a new File Explorer window.
- Open command prompt (DOS) advanced users.
- Providing more options to change or configure file explorer.
- Frequent places that will link users to specific file folders that are often used.

Windows File Explorer

File Explorer is an application built into Windows 10 to help the user find, browse, and select folders and files. The main way to access the File Explorer is by clicking the folder icon in the Taskbar. After clicking the icon, the File Explorer window will open.

Figure 4.5.1: This Photo by Unknown Author is licensed under CC BY-SA

The initial File Explorer window is comprised of the File Explorer Ribbon, the Navigation Pane, Frequent Folders, and Recent Files.

Organize your files

Figure 4.5.1: This Photo by Unknown Author is licensed under CC BY-SA

A common method to organize your files is to place them in folders, which can contain files and other folders. Folders can be stored on multiple locations on your PC. The specific location or address allows your computer to keep track of where each file is located. You may also create a Windows File Explorer window for each folder you are working in. Files and Folders may be moved between folders by clicking on the file/folder and dragging the file/folder to the new location.

Cut, Copy, and Paste

One of the most common operations for users is cut, copy, or pasting files & folders. Pasting a file inserts the new file into the new location. Selecting "Cut" to a file will remove it from its current location and paste the file to its new location. The "Copy" command creates a replica of an original file and pastes the file in the selected location. To copy a file, select the file you want to copy and click Copy under the Home tab. In addition, you can copy or move a file without leaving your current location. Under the Home tab within the Organize group, click 'Copy to' or 'Move to' button, and choose the location if it's listed, or click the Choose Location button and select the location.

Check for Understanding

1. Demonstrate the creation of a new folder. Place a folder in that folder. Create another folder in that folder. Rename the Original Folder.
2. Perform the following: create a folder in your "Documents" folder. Drag the new folder into the "Downloads Folder"
3. Cut a newly created folder in the "Downloads folder" and paste it into the "Pictures" folder.

4.6: Summary

Summary

The software gives the instructions that tell the hardware what to do.

There are two basic categories of software: operating systems and applications.

Operating systems provide access to the computer hardware and make system resources available.

Application software is designed to meet a specific goal.

Productivity software is a subset of application software that provides basic business functionality to a personal computer: word processing, spreadsheets, and presentations.

An ERP system is a software application with a centralized database that is implemented across the entire organization.

Cloud computing is a software delivery method that runs on any computer with a web browser and access to the Internet.

Software is developed through a process called programming, in which a programmer uses a programming language to put together the logic needed to create the program.

The software can be an open-source or a closed-source model, and users or developers are granted different licensing terms.

Virtualization software allows one computing device to function as many.

Virtualization software can lower a firm's hardware needs, save energy, and boost scalability.

Data center virtualization software is at the heart of many so-called private clouds .

Virtualization also works on the desktop, allowing multiple operating systems (Mac OS X, Linux, Windows) to run simultaneously on the same platform.

Virtualization software can increase data center utilization to 80 percent or more.

A number of companies, including Microsoft and Dell, have entered the growing virtualization market.

CHAPTER OVERVIEW

5: Issues in Computing

As computing and networking resources have become more and more an integral part of business, they have also become a target of criminals. Organizations must be vigilant with the way they protect their resources. The same holds true for us personally: as digital devices become more and more intertwined with our lives, it becomes crucial for us to understand how to protect ourselves. Information systems have had an impact far beyond the world of business. New technologies create new situations that we have never dealt with before. How do we handle the new capabilities that these devices empower us with? What new laws are going to be needed to protect us from ourselves? This chapter will kick off with a discussion of the impact of information systems on how we behave (ethics). This will be followed by the new legal structures being put in place, with a focus on intellectual property and privacy.

Objectives

Upon successful completion of this chapter, you will be able to:

- Describe what the term information systems ethics means;
- Explain what a code of ethics is and describe the advantages and disadvantages;
- Define the term intellectual property and explain the protections provided by copyright, patent, and trademark; and
- Describe the challenges that information technology brings to individual privacy.

5.1: Information Systems Security

Authentication and Identification

What is Authentication?

Authentication is the process of verifying whether someone (or something) is, in fact, who (or what) it is declared to be. According to the National Institute of Standards and Technology authentication is defined as "Authentication: Verifying the identity of a user, process, or device, often as a prerequisite to allowing access to resources in an information system". Notice that this definition does not restrict authentication to human users. It includes processes or devices.

Authentication factors

The ways in which someone may be authenticated fall into four categories, based on what is known as the factors of authentication: something the user *knows*, something the user *has*, something the user *does*, and something the user *is*. Each **authentication factor** covers a range of elements used to authenticate or verify a person's identity prior to being granted access, approving a transaction request, signing a document or other work product, granting authority to others, and establishing a chain of authority.

Security research has determined that for a positive authentication, elements from at least two, and preferably all three, factors should be verified. The four factors (classes) and some of the elements of each factor are:

- the knowledge factors: Something the user knows (e.g., a password, partial password, passphrase, personal identification number (PIN), challenge-response (the user must answer a question or pattern), security question).
- the ownership factors: Something the user possesses (e.g., wrist band, ID card, security token, implanted device, cell phone with a built-in hardware token, software token, or cell phone holding a software token).
- the inherence factors: Something the user is or does (e.g., fingerprint, retinal pattern, DNA sequence (there are assorted definitions of what is sufficient), signature, face, voice, unique bio-electric signals, or other biometric identifiers).
- the location factors: Somewhere the user is (e.g. connection to a specific computing network or using a GPS signal to identify the location).

Multi-factor authentication

Multi-factor authentication is an electronic authentication method in which a computer user is granted access to a website or application only after successfully presenting two or more pieces of evidence (or factors) to an authentication mechanism: knowledge (something only the user knows), possession (something only the user has), and inherence (something only the user is). It protects the user from an unknown person trying to access their data such as personal ID details or financial assets.

Authentication takes place when someone tries to log into a computer resource (such as a network, device, or application). The resource requires the user to supply the identity by which the user is known to the resource, along with evidence of the authenticity of the user's claim to that identity. Simple authentication requires only one such piece of evidence (factor), typically a password. For additional security, the resource may require more than one factor—multi-factor authentication, or two-factor authentication in cases where exactly two pieces of evidence are to be supplied.

The use of multiple authentication factors to prove one's identity is based on the premise that an unauthorized actor is unlikely to be able to supply the factors required for access. If in an authentication attempt, at least one of the components is missing or supplied incorrectly, the user's identity is not established with sufficient certainty and access to the asset (e.g., a building, or data) being protected by multi-factor authentication then remains blocked.

Knowledge

Knowledge factors are the most commonly used form of authentication. In this form, the user is required to prove knowledge of a secret in order to authenticate.

A password is a secret word or string of characters that is used for user authentication. This is the most commonly used mechanism of authentication. Many multi-factor authentication techniques rely on passwords as one factor of authentication. Variations include both longer ones formed from multiple words (a passphrase) and the shorter, purely numeric, personal identification number (PIN) commonly used for ATM access. Traditionally, passwords are expected to be memorized.

Many secret questions such as "Where were you born?" are poor examples of a knowledge factor because they may be known to a wide group of people, or be able to be researched.

Possession

Possession factors ("something only the user has") have been used for authentication for centuries, in the form of a key to a lock. The basic principle is that the key embodies a secret that is shared between the lock and the key, and the same principle underlies possession factor authentication in computer systems. A security token is an example of a possession factor.

Disconnected tokens have no connections to the client's computer. They typically use a built-in screen to display the generated authentication data, which is manually typed in by the user. This type of token mostly uses a "one-time password" that can only be used for that specific session.

Connected tokens are devices that are physically connected to the computer to be used. Those devices transmit data automatically. There are a number of different types, including card readers, wireless tags, and USB tokens.

A software token (a.k.a. soft token) is a type of two-factor authentication security device that may be used to authorize the use of computer services. Software tokens are stored on a general-purpose electronic device such as a desktop computer, laptop, PDA, or mobile phone and can be duplicated. (Contrast hardware tokens, where the credentials are stored on a dedicated hardware device and therefore cannot be duplicated, absent physical invasion of the device.) A soft token may not be a device the user interacts with. Typically an X.509v3 certificate is loaded onto the device and stored securely to serve this purpose.

Inherent

These are factors associated with the user and are usually biometric methods, including fingerprint, face, voice, or iris recognition. Behavioral biometrics such as keystroke dynamics can also be used.

Location

Increasingly, a fourth factor is coming into play involving the physical location of the user. While hard wired to the corporate network, a user could be allowed to log in using only a pin code while off the network entering a code from a soft token as well could be required. This could be seen as an acceptable standard where access to the office is controlled.

Systems for network admission control work in similar ways where your level of network access can be contingent on the specific network your device is connected to, such as wifi vs wired connectivity. This also allows a user to move between offices and dynamically receive the same level of network access in each.

Mutual authentication

Mutual authentication or two-way authentication (not to be confused with two-factor authentication) refers to two parties authenticating each other at the same time in an authentication protocol. It was previously referred to as “mutual entity authentication,” as two or more entities verify the others' legality before any data or information is transmitted.

Mutual authentication is a desired characteristic in verification schemes that transmit sensitive data, in order to ensure data security. Mutual authentication is found in two types of schemes: username-password-based schemes and certificate-based schemes, and these schemes are often employed in the Internet of Things (IoT). Writing effective security schemes in IoT systems can become challenging, especially when needing schemes to be lightweight and have low computational costs. Mutual authentication is a crucial security step that can defend against many adversarial attacks, which otherwise can have large consequences if IoT systems (such as e-Healthcare servers) are hacked. In scheme analyses done of past works, a lack of mutual authentication had been considered a weakness in data transmission schemes.

Password

A password, sometimes called a passcode, is a memorized secret, typically a string of characters, usually used to confirm a user's identity. Using the terminology of the NIST Digital Identity Guidelines, "the secret is memorized by a party called the claimant while the party verifying the identity of the claimant is called the verifier. When the claimant successfully demonstrates knowledge of the password to the verifier through an established authentication protocol, the verifier is able to infer the claimant's identity".

In general, a password is an arbitrary string of characters including letters, digits, or other symbols. If the permissible characters are constrained to be numeric, the corresponding secret is sometimes called a personal identification number (PIN).

Despite its name, a password does not need to be an actual word; indeed, a non-word (in the dictionary sense) may be harder to guess, which is a desirable property of passwords. A memorized secret consisting of a sequence of words or other text separated by spaces is sometimes called a passphrase. A passphrase is similar to a password in usage, but the former is generally longer for added security.

Choosing a secure and memorable password

The easier a password is for the owner to remember generally means it will be easier for an attacker to guess. However, passwords that are difficult to remember may also reduce the security of a system because (a) users might need to write down or electronically store the password, (b) users will need frequent password resets and (c) users are more likely to re-use the same password across different accounts. Similarly, the more stringent the password requirements, such as "have a mix of uppercase and lowercase letters and digits" or "change it monthly", the greater the degree to which users will subvert the system. Others argue longer passwords provide more security (e.g., entropy) than shorter passwords with a wide variety of characters.

In The Memorability and Security of Passwords, Jeff Yan et al. examine the effect of advice given to users about a good choice of password. They found that passwords based on thinking of a phrase and taking the first letter of each word are just as memorable as naively selected passwords, and just as hard to crack as randomly generated passwords.

Combining two or more unrelated words and altering some of the letters to special characters or numbers is another good method, but a single dictionary word is not. Having a personally designed algorithm for generating obscure passwords is another good method.

However, asking users to remember a password consisting of a "mix of uppercase and lowercase characters" is similar to asking them to remember a sequence of bits: hard to remember, and only a little bit harder to crack (e.g. only 128 times harder to crack for 7-letter passwords, less if the user simply capitalizes one of the letters). Asking users to use "both letters and digits" will often lead to easy-to-guess substitutions such as 'E' → '3' and 'I' → '1', substitutions that are well known to attackers. Similarly typing the password one keyboard row higher is a common trick known to attackers.

In 2013, Google released a list of the most common password types, all of which are considered insecure because they are too easy to guess (especially after researching an individual on social media):

- The name of a pet, child, family member, or significant other
- Anniversary dates and birthdays
- Birthplace
- Name of a favorite holiday
- Something related to a favorite sports team
- The word "password"

Strong Password

A strong password is hard to detect both by humans and by the computer. Two things make a password stronger: (1) a larger number of characters (and the more characters, the stronger the password), and (2) mixing numeric digits, upper and lower case letters, and special characters ($, #, etc.). Passwords are typically case-sensitive, so a strong password contains **letters in both uppercase and lowercase.** A strong password is not a word that can be found in a dictionary or the name of a person, character, product, or organization, significantly different from your previous passwords, and easy for you to remember but difficult for others to guess. Consider using a memorable phrase like "6MonkeysRLooking^".

Factors in the security of a password system

The security of a password-protected system depends on several factors. The overall system must be designed for sound security, with protection against computer viruses, man-in-the-middle attacks, and the like. Physical security issues are also a concern, from deterring shoulder surfing to more sophisticated physical threats such as video cameras and keyboard sniffers. Passwords should be chosen so that they are hard for an attacker to guess and hard for an attacker to discover using any of the available automatic attack schemes. See password strength and computer security for more information.

Nowadays, it is a common practice for computer systems to hide passwords as they are typed. The purpose of this measure is to prevent bystanders from reading the password; however, some argue that this practice may lead to mistakes and stress, encouraging users to choose weak passwords. As an alternative, users should have the option to show or hide passwords as they type them.

Effective access control provisions may force extreme measures on criminals seeking to acquire a password or biometric token. Less extreme measures include extortion, rubber hose cryptanalysis, and side-channel attack.

The rate at which an attacker can try guessed passwords

The rate at which an attacker can submit guessed passwords to the system is a key factor in determining system security. Some systems impose a time-out of several seconds after a small number (e.g., three) of failed password entry attempts. In the absence of other vulnerabilities, such systems can be effectively secure with relatively simple passwords if they have been well chosen and are not easily guessed.

Many systems store a cryptographic hash of the password. If an attacker gets access to the file of hashed passwords guessing can be done offline, rapidly testing candidate passwords against the true password's hash value. In the example of a web-server, an online attacker can guess only at the rate at which the server will respond, while an off-line attacker (who gains access to the file) can guess at a rate limited only by the hardware on which the attack is running.

Passwords that are used to generate cryptographic keys (e.g., for disk encryption or Wi-Fi security) can also be subjected to high rate guessing. Lists of common passwords are widely available and can make password attacks very efficient. (See Password cracking.) Security in such situations depends on using passwords or passphrases of adequate complexity, making such an attack computationally infeasible for the attacker. Some systems, such as PGP and Wi-Fi WPA, apply a computation-intensive hash to the password to slow such attacks. See key stretching.

Limits on the number of password guesses

An alternative to limiting the rate at which an attacker can make guesses on a password is to limit the total number of guesses that can be made. The password can be disabled, requiring a reset, after a small number of consecutive bad guesses (say 5); and the user may be required to change the password after a larger cumulative number of bad guesses (say 30), to prevent an attacker from making an arbitrarily large number of bad guesses by interspersing them between good guesses made by the legitimate password owner. Attackers may conversely use knowledge of this mitigation to implement a denial of service attack against the user by intentionally locking the user out of their own device; this denial of service may open other avenues for the attacker to manipulate the situation to their advantage via social engineering.

Identification - what is it?

Identification is basically the process of someone claiming to be a specific person. They can identify themselves as “Pat”, show an id card of some type of card with a name on it, or have an email address showing their name.

In the current context of online transactions, users “identify” themselves by providing a name, an email address, or phone number to a web request. For example, using a process of identification alone, as long as a buyer has the card’s proper information that is associated with the card being used, the user is pretty much accepted as is.

A business that allows identification by itself is essentially saying, "We have no reason to doubt that you are indeed the person you claim to be", despite having not independently verified if the information is truthful. It’s like asking, “Who are you?” and simply accepting whatever answer is given. For transactions where there is not a lot at stake, like registering for a class or checking out a book, simply having someone declare their identity without providing any verification may be good enough.

It is becoming more and more frequent that identification alone is adequate. It’s like having a username without a password.

So how can we determine the person is who they say they are? That’s where verification comes in.

What is Verification?

Verification goes beyond the basic question, “Who are you?” Identity verification goes the extra mile and asks, “Are you really who you say you are?” the response needs to provide, a high degree of confidence that, the answer is accurate.

The most accurate way to verify someone's identity is to request and validate more than one form of identification against the person standing in front of you, with at least one of them being a photo ID. A driver's license, a Social Security card, a valid passport, or military photo identification are some forms of identification. Verifying someone’s identity to a high degree of certainty takes effort. At a time when service providers want to provide a “frictionless” onboarding process, some may cut corners and require a low barrier to entry. Typical social media accounts, for example, only ask new users to provide a name, email address, username, and password. A phone number may be thrown in there for good measure.

Depending on the organization and the level of assurance needed, a university ID or other non-government issued identification card may suffice for one form of ID. Identity verification in the electronic sense also called identity "proofing" or "vetting", is used to confirm an identity where the individual is not standing before you to show some sort of picture ID. In these cases, most organizations require a real-time process that validates the personal information provided by the individual.

Apply for an online bank account, though, and you may be expected to provide a social security number, photo ID or passport, and proof of your current address. The stakes associated with a bank account are much greater than those with a TikTok account, therefore the verification requirements are more stringent. In fact, in the financial sector alone, there are numerous regulatory acts to prevent fraudsters from setting up false bank accounts, laundering money, and other unseemly criminal activities. The compliance mandates associated with these regulations are not satisfied by traditional verification methods, which is why businesses are beginning to make a shift to pairing a customer's identity information with one of their biometric markers at the point of onboarding.

The Information Security Triad: Confidentiality, Integrity, Availability (CIA)

The Information Security Triad, also known as the CIA triad, is a guide for organizations. This guide helps them make policies to protect information security.

In this context, the CIA means the following:

- Confidentiality – set of rules that limit access to information
- Integrity – the assurance that the information is reliable and correct
- Availability – a guarantee of reliable access to the information

Confidentiality

When protecting information, we want to be able to restrict access to those who are allowed to see it; everyone else should be disallowed from learning anything about its contents. This is the essence of confidentiality. For example, federal law requires that universities restrict access to private student information. The university must be sure that only those who are authorized have access to view the grade records.

Integrity

Integrity is the assurance that the information being accessed has not been altered and truly represents what is intended. Just as a person with integrity means what he or she says and can be trusted to consistently represent the truth, information integrity means information truly represents its intended meaning. Information can lose its integrity through malicious intent, such as when someone who is not authorized makes a change to intentionally misrepresent something. An example of this would be when a hacker is hired to go into the university's system and change a grade.

Integrity can also be lost unintentionally, such as when a computer power surge corrupts a file or someone authorized to make a change accidentally deletes a file or enters incorrect information.

Availability

Information availability is the third part of the CIA triad. *Availability* means that information can be accessed and modified by anyone authorized to do so in an appropriate timeframe. Depending on the type of information, an *appropriate timeframe* can mean different things. For example, a stock trader needs information to be available immediately, while a salesperson may be happy to get sales numbers for the day in a report the next morning. Companies such as Amazon.com will require their servers to be available twenty-four hours a day, seven days a week. Other companies may not suffer if their web servers are down for a few minutes once in a while.

Tools for Information Security

In order to ensure the confidentiality, integrity, and availability of information, organizations can choose from a variety of tools. Each of these tools can be utilized as a part of an overall information-security policy.

Access Control

Once a user has been authenticated, the next step is to ensure that they can only access the information resources that are appropriate. This is done through the use of access control. Access control determines which users are authorized to read, modify, add, and/or delete information. Several different access control models exist. Here we will discuss two: the access control list (ACL) and role-based access control (RBAC).

For each information resource that an organization wishes to manage, a list of users who have the ability to take specific actions can be created. This is an access control list or ACL. For each user, specific capabilities are assigned, such as *read*, *write*, *delete*, or *add*. Only users with those capabilities are allowed to perform those functions. If a user is not on the list, they have no ability to even know that the information resource exists.

ACLs are simple to understand and maintain. However, they have several drawbacks. The primary drawback is that each information resource is managed separately, so if a security administrator wanted to add or remove a user to a large set of information resources, it would be quite difficult. And as the number of users and resources increases, ACLs become harder to maintain. This has led to an improved method of access control, called role-based access control, or RBAC. With RBAC, instead of giving specific users access rights to an information resource, users are assigned to roles and then those roles are assigned access. This allows the administrators to manage users and roles separately, simplifying administration and, by extension, improving security.

Encryption

Many times, an organization needs to transmit information over the Internet or transfer it on external media such as a CD or flash drive. In these cases, even with proper authentication and access control, it is possible for an unauthorized person to get access to the data. Encryption is a process of encoding data upon its transmission or storage so that only authorized individuals can read it. This encoding is accomplished by a computer program, which encodes the plain text that needs to be transmitted; then the recipient receives the ciphertext and decodes it (decryption). In order for this to work, the sender and receiver need to agree on the method of encoding so that both parties can communicate properly. Both parties share the encryption key, enabling them to encode and decode each other's messages. This is called symmetric key encryption. This type of encryption is problematic because the key is available in two different places.

An alternative to symmetric key encryption is public-key encryption. In public-key encryption, two keys are used: a public key and a private key. To send an encrypted message, you obtain the public key, encode the message, and send it. The recipient then uses the private key to decode it. The public key can be given to anyone who wishes to send the recipient a message. Each user simply needs one private key and one public key in order to secure messages. The private key is necessary in order to decrypt something sent with the public key.

Backups

Another essential tool for information security is a comprehensive backup plan for the entire organization. Not only should the data on the corporate servers be backed up, but individual computers used throughout the organization should also be backed up. A good backup plan should consist of several components.

- A full understanding of the organizational information resources. What information does the organization actually have? Where is it stored? Some data may be stored on the organization's servers, other data on users' hard drives, some in the cloud, and some on third-party sites. An organization should make a full inventory of all of the information that needs to be backed up and determine the best way to back it up.
- Regular backups of all data. The frequency of backups should be based on how important the data is to the company, combined with the ability of the company to replace any data that is lost. Critical data should be backed up daily, while less critical data could be backed up weekly.
- Offsite storage of backup data sets. If all of the backup data is being stored in the same facility as the original copies of the data, then a single event, such as an earthquake, fire, or tornado, would take out both the original data and the backup! It is essential that part of the backup plan is to store the data in an offsite location.
- Test of data restoration. On a regular basis, the backups should be put to the test by having some of the data restored. This will ensure that the process is working and will give the organization confidence in the backup plan.

Besides these considerations, organizations should also examine their operations to determine what effect downtime would have on their business. If their information technology were to be unavailable for any sustained period of time, how would it impact the business?

Additional concepts related to backup include the following:

- Universal Power Supply (UPS). A UPS is a device that provides battery backup to critical components of the system, allowing them to stay online longer and/or allowing the IT staff to shut them down using proper procedures in order to prevent the data loss that might occur from a power failure.
- Alternate, or "hot" sites. Some organizations choose to have an alternate site where an exact replica of their critical data is always kept up to date. When the primary site goes down, the alternate site is immediately brought online so that little or no downtime is experienced.

As information has become a strategic asset, a whole industry has sprung up around the technologies necessary for implementing a proper backup strategy. A company can contract with a service provider to back up all of their data or they can purchase large amounts of online storage space and do it themselves. Technologies such as storage area networks and archival systems are now used by most large businesses.

Firewalls

Network configuration with firewalls, IDS, and a DMZ. Click to enlarge.

Another method that an organization should use to increase security on its network is a firewall. A firewall can exist as hardware or software (or both). A hardware firewall is a device that is connected to the network and filters the packets based on a set of rules. A software firewall runs on the operating system and intercepts packets as they arrive at a computer. A firewall protects all company servers and computers by stopping packets from outside the organization's network that does not meet a strict set of criteria. A firewall may also be configured to restrict the flow of packets leaving the organization. This may be done to eliminate the possibility of employees watching YouTube videos or using Facebook from a company computer.

Some organizations may choose to implement multiple firewalls as part of their network security configuration, creating one or more sections of their network that are partially secured. This segment of the network is referred to as a DMZ, borrowing the term *demilitarized zone* from the military, and it is where an organization may place resources that need broader access but still need to be secured.

Intrusion Detection Systems

Another device that can be placed on the network for security purposes is an intrusion detection system or IDS. An IDS does not add any additional security; instead, it provides the functionality to identify if the network is being attacked. An IDS can be configured to watch for specific types of activities and then alert security personnel if that activity occurs. An IDS also can log various types of traffic on the network for analysis later. An IDS is an essential part of any good security setup.

Physical Security

An organization can implement the best authentication scheme in the world, develop the best access control, and install firewalls and intrusion prevention, but its security cannot be complete without the implementation of physical security. Physical security is the protection of the actual hardware and networking components that store and transmit information resources. To implement physical security, an organization must identify all of the vulnerable resources and take measures to ensure that these resources cannot be physically tampered with or stolen. These measures include the following.

- Locked doors: It may seem obvious, but all the security in the world is useless if an intruder can simply walk in and physically remove a computing device. High-value information assets should be secured in a location with limited access.
- Physical intrusion detection: High-value information assets should be monitored through the use of security cameras and other means to detect unauthorized access to the physical locations where they exist.
- Secured equipment: Devices should be locked down to prevent them from being stolen. One employee's hard drive could contain all of your customer information, so it is essential that it be secured.
- Environmental monitoring: An organization's servers and other high-value equipment should always be kept in a room that is monitored for temperature, humidity, and airflow. The risk of a server failure rises when these factors go out of a specified range.
- Employee training: One of the most common ways thieves steal corporate information is to steal employee laptops while employees are traveling. Employees should be trained to secure their equipment whenever they are away from the office.

Security Policies

Besides the technical controls listed above, organizations also need to implement security policies as a form of administrative control. In fact, these policies should really be a starting point in developing an overall security plan. A good information-security policy lays out the guidelines for employee use of the information resources of the company and provides the company recourse in the case that an employee violates a policy.

According to the SANS Institute, a good policy is "a formal, brief, and high-level statement or plan that embraces an organization's general beliefs, goals, objectives, and acceptable procedures for a specified subject area." Policies require compliance; failure to comply with a policy will result in disciplinary action. A policy does not lay out the specific technical details, instead it focuses on the desired results. A security policy should be based on the guiding principles of confidentiality, integrity, and availability.[2]

A good example of a security policy that many will be familiar with is a web use policy. A web use policy lays out the responsibilities of company employees as they use company resources to access the Internet.

A security policy should also address any governmental or industry regulations that apply to the organization. For example, if the organization is a university, it must be aware of the Family Educational Rights and Privacy Act (FERPA), which restricts who has access to student information. Health care organizations are obligated to follow several regulations, such as the Health Insurance Portability and Accountability Act (HIPAA).

A good resource for learning more about security policies is the SANS Institute's Information Security Policy Page.

Usability

When looking to secure information resources, organizations must balance the need for security with users' need to effectively access and use these resources. If a system's security measures make it difficult to use, then users will find ways around the security, which may make the system more vulnerable than it would have been without the security measures! Take, for example, password policies. If the organization requires an extremely long password with several special characters, an employee may resort to writing it down and putting it in a drawer since it will be impossible to memorize.

Personal Information Security

Poster from Stop. Think. Connect. Click to enlarge. (Copyright: Stop. Think. Connect. http://stopthinkconnect.org/resources)

We will end this chapter with a discussion of what measures each of us, as individual users, can take to secure our computing technologies. There is no way to have 100% security, but there are several simple steps we, as individuals, can take to make ourselves more secure.

- Keep your software up to date. Whenever a software vendor determines that a security flaw has been found in their software, they will release an update to the software that you can download to fix the problem. Turn on automatic updating on your computer to automate this process.
- Install antivirus software and keep it up to date. There are many good antivirus software packages on the market today, including free ones.
- Be smart about your connections. You should be aware of your surroundings. When connecting to a Wi-Fi network in a public place, be aware that you could be at risk of being spied on by others sharing that network. It is advisable not to access your financial or personal data while attached to a Wi-Fi hotspot. You should also be aware that connecting USB flash drives to your device could also put you at risk. Do not attach an unfamiliar flash drive to your device unless you can scan it first with your security software.
- Back up your data. Just as organizations need to back up their data, individuals need to as well. And the same rules apply: do it regularly and keep a copy of it in another location. One simple solution for this is to set up an account with an online backup service, such as Mozy or Carbonite, to automate your backups.
- Secure your accounts with two-factor authentication. Most e-mail and social media providers now have a two-factor authentication option. The way this works is simple: when you log in to your account from an unfamiliar computer for the first time, it sends you a text message with a code that you must enter to confirm that you are really you. This means that no one else can log in to your accounts without knowing your password *and* having your mobile phone with them.
- Make your passwords long, strong, and unique. For your personal passwords, you should follow the same rules that are recommended for organizations. Your passwords should be long (eight or more characters) and contain at least two of the following: upper-case letters, numbers, and special characters. You also should use different passwords for different accounts, so that if someone steals your password for one account, they still are locked out of your other accounts.
- Be suspicious of strange links and attachments. When you receive an e-mail, tweet, or Facebook post, be suspicious of any links or attachments included there. Do not click on the link directly if you are at all suspicious. Instead, if you want to access the website, find it yourself and navigate to it directly.

You can find more about these steps and many other ways to be secure with your computing by going to Stop. Think. Connect. This website is part of a campaign that was launched in October of 2010 by STOP. THINK. CONNECT. Messaging Convention in partnership with the U.S. government, including the White House.

Sidebar: Mobile Security

As the use of mobile devices such as smartphones and tablets proliferates, organizations must be ready to address the unique security concerns that the use of these devices brings. One of the first questions an organization must consider is whether to allow mobile devices in the workplace at all. Many employees already have these devices, so the question becomes: Should we allow employees to bring their own devices and use them as part of their employment activities? Or should we provide the devices to our employees? Creating a BYOD ("Bring Your Own Device") policy allows employees to integrate themselves more fully into their job and can bring higher employee satisfaction and productivity. In many cases, it may be virtually impossible to prevent employees from having their own smartphones or iPads in the workplace. If the organization provides the devices to its employees, it gains more control over use of the devices, but it also exposes itself to the possibility of an administrative (and costly) mess.

Mobile devices can pose many unique security challenges to an organization. Probably one of the biggest concerns is the theft of intellectual property. For an employee with malicious intent, it would be a very simple process to connect a mobile device either to a computer via the USB port, or wirelessly to the corporate network, and download confidential data. It would also be easy to secretly take a high-quality picture using a built-in camera.

When an employee does have permission to access and save company data on his or her device, a different security threat emerges: that device now becomes a target for thieves. Theft of mobile devices (in this case, including laptops) is one of the primary methods that data thieves use.

So what can be done to secure mobile devices? It will start with a good policy regarding their use. According to a 2013 SANS study, organizations should consider developing a mobile device policy that addresses the following issues: use of the camera, use of voice recording, application purchases, encryption at rest, Wi-Fi auto-connect settings, Bluetooth settings, VPN use, password settings, lost or stolen device reporting, and backup. [3]

Besides policies, there are several different tools that an organization can use to mitigate some of these risks. For example, if a device is stolen or lost, geolocation software can help the organization find it. In some cases, it may even make sense to install remote data-removal software, which will remove data from a device if it becomes a security risk.

Study Questions

1. Briefly define each of the three members of the information security triad.
2. What does the term *authentication* mean?
3. What is multi-factor authentication?
4. What is role-based access control?
5. What is the purpose of encryption?
6. What are two good examples of a complex password?
7. What is pretexting?
8. What are the components of a good backup plan?
9. What is a firewall?
10. What does the term *physical security* mean?

Exercises

1. Describe one method of multi-factor authentication that you have experienced and discuss the pros and cons of using multi-factor authentication.
2. What are some of the latest advances in encryption technologies? Conduct some independent research on encryption using scholarly or practitioner resources, then write a two- to three-page paper that describes at least two new advances in encryption technology.
3. What is the password policy at your place of employment or study? Do you have to change passwords every so often? What are the minimum requirements for a password?
4. When was the last time you backed up your data? What method did you use? In one to two pages, describe a method for backing up your data. Ask your instructor if you can get extra credit for backing up your data.
5. Find the information security policy at your place of employment or study. Is it a good policy? Does it meet the standards outlined in the chapter?
6. How are you doing on keeping your own information secure? Review the steps listed in the chapter and comment on how well you are doing.

5.2: The Ethical and Legal Implications of Information Systems

The rapid changes in information technology in the past few decades have brought a broad array of new capabilities and powers to governments, organizations, and individuals alike. These new capabilities have required thoughtful analysis and the creation of new norms, regulations, and laws. In this chapter, we have seen how the areas of intellectual property and privacy have been affected by these new capabilities and how the regulatory environment has been changed to address them.

Information Systems Ethics

The term *ethics* is defined as "a set of moral principles" or "the principles of conduct governing an individual or a group." Since the dawn of civilization, the study of ethics and its impact has fascinated mankind. But what do ethics have to do with information systems?

The introduction of new technology can have a profound effect on human behavior. New technologies give us capabilities that we did not have before, which in turn create environments and situations that have not been specifically addressed in ethical terms. Those who master new technologies gain new power; those who cannot or do not master them may lose power. In 1913, Henry Ford implemented the first moving assembly line to create his Model T cars. While this was a great step forward technologically (and economically), the assembly line reduced the value of human beings in the production process. The development of the atomic bomb concentrated unimaginable power in the hands of one government, which then had to wrestle with the decision to use it. Today's digital technologies have created new categories of ethical dilemmas.

For example, the ability to anonymously make perfect copies of digital music has tempted many music fans to download copyrighted music for their own use without making payment to the music's owner. Many of those who would never have walked into a music store and stolen a CD find themselves with dozens of illegally downloaded albums.

Digital technologies have given us the ability to aggregate information from multiple sources to create profiles of people. What would have taken weeks of work in the past can now be done in seconds, allowing private organizations and governments to know more about individuals than at any time in history. This information has value but also chips away at the privacy of consumers and citizens.

Code of Ethics

One method for navigating new ethical waters is a code of ethics. A code of ethics is a document that outlines a set of acceptable behaviors for a professional or social group; generally, it is agreed to by all members of the group. The document details different actions that are considered appropriate and inappropriate.

A good example of a code of ethics is the *Code of Ethics and Professional Conduct* of the Association for Computing Machinery, an organization of computing professionals that includes academics, researchers, and practitioners. Here is a quote from the preamble:

Commitment to ethical professional conduct is expected of every member (voting members, associate members, and student members) of the Association for Computing Machinery (ACM).

This Code, consisting of 24 imperatives formulated as statements of personal responsibility, identifies the elements of such a commitment. It contains many, but not all, issues professionals are likely to face. Section 1 outlines fundamental ethical considerations, while Section 2 addresses additional, more specific considerations of professional conduct. Statements in Section 3 pertain more specifically to individuals who have a leadership role, whether in the workplace or in a volunteer capacity such as with organizations like ACM. Principles involving compliance with this Code are given in Section 4.

In the ACM's code, you will find many straightforward ethical instructions, such as the admonition to be honest and trustworthy. But because this is also an organization of professionals that focuses on computing, there are more specific admonitions that relate directly to information technology:

- No one should enter or use another's computer system, software, or data files without permission. One must always have appropriate approval before using system resources, including communication ports, file space, other system peripherals, and computer time.
- Designing or implementing systems that deliberately or inadvertently demean individuals or groups is ethically unacceptable.

- Organizational leaders are responsible for ensuring that computer systems enhance, not degrade, the quality of working life. When implementing a computer system, organizations must consider the personal and professional development, physical safety, and human dignity of all workers. Appropriate human-computer ergonomic standards should be considered in system design and in the workplace.

One of the major advantages of creating a code of ethics is that it clarifies the acceptable standards of behavior for a professional group. The varied backgrounds and experiences of the members of a group lead to a variety of ideas regarding what is acceptable behavior. While to many the guidelines may seem obvious, having these items detailed provides clarity and consistency. Explicitly stating standards communicates the common guidelines to everyone in a clear manner.

Having a code of ethics can also have some drawbacks. First of all, a code of ethics does not have legal authority; in other words, breaking a code of ethics is not a crime in itself. So what happens if someone violates one of the guidelines? Many codes of ethics include a section that describes how such situations will be handled. In many cases, repeated violations of the code result in expulsion from the group.

In the case of ACM: "Adherence of professionals to a code of ethics is largely a voluntary matter. However, if a member does not follow this code by engaging in gross misconduct, membership in ACM may be terminated." Expulsion from ACM may not have much of an impact on many individuals, since membership in ACM is usually not a requirement for employment. However, expulsion from other organizations, such as a state bar organization or medical board, could carry a huge impact.

Another possible disadvantage of a code of ethics is that there is always a chance that important issues will arise that are not specifically addressed in the code. Technology is quickly changing, and a code of ethics might not be updated often enough to keep up with all of the changes. A good code of ethics, however, is written in a broad enough fashion that it can address the ethical issues of potential changes to technology while the organization behind the code makes revisions.

Finally, a code of ethics could have also been a disadvantage in that it may not entirely reflect the ethics or morals of every member of the group. Organizations with a diverse membership may have internal conflicts as to what is acceptable behavior. For example, there may be a difference of opinion on the consumption of alcoholic beverages at company events. In such cases, the organization must make a choice about the importance of addressing a specific behavior in the code.

Acceptable Use Policy

Many organizations that provide technology services to a group of constituents or the public require agreement to an acceptable use policy (AUP) before those services can be accessed. Similar to a code of ethics, this policy outlines what is allowed and what is not allowed while someone is using the organization's services. An everyday example of this is the terms of service that must be agreed to before using the public Wi-Fi at Starbucks, McDonald's, or even a university. Here is an example of an acceptable use policy from Virginia Tech.

Just as with a code of ethics, these acceptable use policies specify what is allowed and what is not allowed. Again, while some of the items listed are obvious to most, others are not so obvious:

- "Borrowing" someone else's login ID and password is prohibited.
- Using the provided access for commercial purposes, such as hosting your own business website, is not allowed.
- Sending out unsolicited emails to a large group of people is prohibited.

Also as with codes of ethics, violations of these policies have various consequences. In most cases, such as with Wi-Fi, violating the acceptable use policy will mean that you will lose your access to the resource. While losing access to Wi-Fi at Starbucks may not have a lasting impact, a university student getting banned from the university's Wi-Fi (or possibly all network resources) could have a large impact.

Intellectual Property

One of the domains that have been deeply impacted by digital technologies is the domain of intellectual property. Digital technologies have driven a rise in new intellectual property claims and made it much more difficult to defend intellectual property.

Intellectual property is defined as "property (as an idea, invention, or process) that derives from the work of the mind or intellect."[3] This could include creations such as song lyrics, a computer program, a new type of toaster, or even a sculpture.

Practically speaking, it is very difficult to protect an idea. Instead, intellectual property laws are written to protect the tangible results of an idea. In other words, just coming up with a song in your head is not protected, but if you write it down it can be protected.

Protection of intellectual property is important because it gives people an incentive to be creative. Innovators with great ideas will be more likely to pursue those ideas if they have a clear understanding of how they will benefit. In the US Constitution, Article 8, Section 8, the authors saw fit to recognize the importance of protecting creative works:

Congress shall have the power . . . To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.

An important point to note here is the "limited time" qualification. While protecting intellectual property is important because of the incentives it provides, it is also necessary to limit the amount of benefit that can be received and allow the results of ideas to become part of the public domain.

Outside of the US, intellectual property protections vary. You can find out more about a specific country's intellectual property laws by visiting the World Intellectual Property Organization.

Copyright

Copyright is the protection given to songs, computer programs, books, and other creative works; any work that has an "author" can be copyrighted. Under the terms of copyright, the author of work controls what can be done with the work, including:

- Who can make copies of the work.
- Who can make derivative works from the original work.
- Who can perform the work publicly.
- Who can display the work publicly.
- Who can distribute the work.

Many times, work is not owned by an individual but is instead owned by a publisher with whom the original author has an agreement. In return for the rights to the work, the publisher will market and distribute the work and then pay the original author a portion of the proceeds.

Copyright protection lasts for the life of the original author plus seventy years. In the case of a copyrighted work owned by a publisher or another third party, the protection lasts for ninety-five years from the original creation date. For works created before 1978, the protections vary slightly. You can see the full details on copyright protections by reviewing the Copyright Basics document available at the US Copyright Office's website.

Obtaining Copyright Protection

In the United States, copyright is obtained by the simple act of creating the original work. In other words, when an author writes down that song, makes that film, or designs that program, he or she automatically has the copyright. However, for a work that will be used commercially, it is advisable to register for a copyright with the US Copyright Office. A registered copyright is needed in order to bring legal action against someone who has used a work without permission.

First Sale Doctrine

If an artist creates a painting and sells it to a collector who then, for whatever reason, proceeds to destroy it, does the original artist have any recourse? What if the collector, instead of destroying it, begins making copies of it and sells them? Is this allowed? The first sale doctrine is a part of copyright law that addresses this, as shown below[4]:

The first sale doctrine, codified at 17 U.S.C. § 109, provides that an individual who knowingly purchases a copy of a copyrighted work from the copyright holder receives the right to sell, display or otherwise dispose of *that particular copy*, notwithstanding the interests of the copyright owner.

So, in our examples, the copyright owner has no recourse if the collector destroys her artwork. But the collector does not have the right to make copies of the artwork.

Fair Use

Another important provision within copyright law is that of fair use. Fair use is a limitation on copyright law that allows for the use of protected works without prior authorization in specific cases. For example, if a teacher wanted to discuss a current event in her class, she could pass out copies of a copyrighted news story to her students without first getting permission. Fair use is also what allows a student to quote a small portion of a copyrighted work in a research paper.

Unfortunately, the specific guidelines for what is considered fair use and what constitutes copyright violation are not well defined. Fair use is a well-known and respected concept and will only be challenged when copyright holders feel that the integrity or market value of their work is being threatened. The following four factors are considered when determining if something constitutes fair use: [5]

1. The purpose and character of the use, including whether such use is of commercial nature or is for nonprofit educational purposes;
2. The nature of the copyrighted work;
3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole;
4. The effect of the use upon the potential market for, or value of, the copyrighted work.

If you are ever considering using a copyrighted work as part of something you are creating, you may be able to do so under fair use. However, it is always best to check with the copyright owner to be sure you are staying within your rights and not infringing upon theirs.

Patent

Another important form of intellectual property protection is the patent. A patent creates protection for someone who invents a new product or process. The definition of invention is quite broad and covers many different fields. Here are some examples of items receiving patents:

- circuit designs in semiconductors;
- prescription drug formulas;
- firearms;
- locks;
- plumbing;
- engines;
- coating processes; and
- business methods.

Once a patent is granted, it provides the inventor with protection from others infringing on his or her patent. A patent holder has the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States for a limited time in exchange for public disclosure of the invention when the patent is granted.

As with copyright, patent protection lasts for a limited period of time before the invention or process enters the public domain. In the US, a patent lasts twenty years. This is why generic drugs are available to replace brand-name drugs after twenty years.

Business Method Patents

Most patents protect a physical object. However, a business method patent is a type of utility patent that protects a process such as one created in a software program. It has proved especially useful to online companies that use software for their business. Protection through a business method patent, also known as an Internet patent, allows the owner to control rights to that business method during the life of the patent. This means that they can prevent other companies from using the process, or they can license the process to other companies for a fee.

Requirements for Business Method Patents

There are four main requirements for patenting a business method. First, the business method must be patentable subject matter rather than an abstract idea. It also must meet the usefulness requirement, but this is a low threshold to meet. The method only must produce a concrete result. The two remaining requirements are more complicated. These involve showing that the method is novel and that it is not obvious.

A novel business method must be different from any previous method or invention, and it must not have been exposed to the public. A method is exposed to the public if it was used publicly or described in a publication before the application was filed. (An exception applies if the applicant described the method in a publication no more than one year before the filing date.)

A business method that is non-obvious must produce a result that is new or unexpected to someone in the relevant field. This is often determined by assessing whether someone in the field would have seen the new method as a natural evolution of previous methods.

The Digital Millennium Copyright Act

As digital technologies have changed what it means to create, copy, and distribute media, a policy vacuum has been created. In 1998, the US Congress passed the Digital Millennium Copyright Act (DMCA), which extended copyright law to take into consideration digital technologies. Two of the best-known provisions from the DMCA are the anti-circumvention provision and the "safe harbor" provision.

- The anti-circumvention provision makes it illegal to create technology to circumvent technology that has been put in place to protect a copyrighted work. This provision includes not just the creation of the technology but also the publishing of information that describes how to do it. While this provision does allow for some exceptions, it has become quite controversial and has led to a movement to have it modified.
- The "safe harbor" provision limits the liability of online service providers when someone using their services commits copyright infringement. This is the provision that allows YouTube, for example, not to be held liable when someone posts a clip from a copyrighted movie. The provision does require the online service provider to take action when they are notified of the violation (a "takedown" notice).

Many think that the DMCA goes too far and ends up limiting our freedom of speech. The Electronic Frontier Foundation (EFF) is at the forefront of this battle. For example, in discussing the anti-circumvention provision, the EFF states:

Yet the DMCA has become a serious threat that jeopardizes fair use, impedes competition and innovation, chills free expression and scientific research, and interferes with computer intrusion laws. If you circumvent DRM [digital rights management] locks for non-infringing fair uses or create the tools to do so you might be on the receiving end of a lawsuit.

Privacy

The term *privacy* has many definitions, but for our purposes, privacy will mean the ability to control information about oneself. Our ability to maintain our privacy has eroded substantially in the past decades, due to information systems.

Personally Identifiable Information

Information about a person that can be used to uniquely establish that person's identify is called personally identifiable information, or PII. This is a broad category that includes information such as:

- name;
- social security number;
- date of birth;
- place of birth;
- mother's maiden name;
- biometric records (fingerprint, face, etc.);
- medical records;
- educational records;
- financial information; and
- employment information.

Organizations that collect PII are responsible to protect it. The Department of Commerce recommends that "organizations minimize the use, collection, and retention of PII to what is strictly necessary to accomplish their business purpose and mission." They go on to state that "the likelihood of harm caused by a breach involving PII is greatly reduced if an organization minimizes the amount of PII it uses, collects, and stores."[7] Organizations that do not protect PII can face penalties, lawsuits, and loss of business. In the US, most states now have laws in place requiring organizations that have had security breaches related to PII to notify potential victims, as does the European Union.

Just because companies are required to protect your information does not mean they are restricted from sharing it. In the US, companies can share your information without your explicit consent (see sidebar below), though not all do so. Companies that collect PII are urged by the FTC to create a privacy policy and post it on their website.

While the privacy laws in the US seek to balance consumer protection with promoting commerce, in the European Union privacy is considered a fundamental right that outweighs the interests of commerce. This has led to much stricter privacy protection in the EU, but also makes commerce more difficult between the US and the EU.

Other Privacy Laws

Restrictions on Record Collecting

In the US, the government has strict guidelines on how much information can be collected about its citizens. Certain classes of information have been restricted by laws over time, and the advent of digital tools has made these restrictions more important than ever.

Children's Online Privacy Protection Act

Websites that are collecting information from children under the age of thirteen are required to comply with the Children's Online Privacy Protection Act (COPPA), which is enforced by the Federal Trade Commission (FTC). To comply with COPPA, organizations must make a good-faith effort to determine the age of those accessing their websites and, if users are under thirteen years old, must obtain parental consent before collecting any information.

Family Educational Rights and Privacy Act

The Family Educational Rights and Privacy Act (FERPA) is a US law that protects the privacy of student education records. In brief, this law specifies that parents have a right to their child's educational information until the child reaches either the age of eighteen or begins attending school beyond the high school level. At that point, control of the information is given to the child. While this law is not specifically about the digital collection of information on the Internet, the educational institutions that are collecting student information are at a higher risk for disclosing it improperly because of digital technologies.

Health Insurance Portability and Accountability Act

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is the law that specifically singles out records related to health care as a special class of personally identifiable information. This law gives patients specific rights to control their medical records, requires health care providers and others who maintain this information to get specific permission in order to share it, and imposes penalties on the institutions that breach this trust. Since much of this information is now shared via electronic medical records, the protection of those systems becomes paramount.

Non-Obvious Relationship Awareness

Digital technologies have given us many new capabilities that simplify and expedite the collection of personal information. Every time we come into contact with digital technologies, information about us is being made available. From our location to our web-surfing habits, our criminal record to our credit report, we are constantly being monitored. This information can then be aggregated to create profiles of each and every one of us. While much of the information collected was available in the past, collecting it and combining it took time and effort. Today, detailed information about us is available for purchase from different companies. Even information not categorized as PII can be aggregated in such a way that an individual can be identified.

This process of collecting large quantities of a variety of information and then combining it to create profiles of individuals is known as non-obvious relationship awareness or NORA. First commercialized by big casinos looking to find cheaters, NORA is used by both government agencies and private organizations, and it is big business.

In some settings, NORA can bring many benefits, such as in law enforcement. By being able to identify potential criminals more quickly, crimes can be solved more quickly or even prevented before they happen. But these advantages come at a price: our privacy.

Study Questions

1. What does the term *information systems ethics* mean?
2. What is a code of ethics? What are one advantage and one disadvantage of a code of ethics?
3. What does the term *intellectual property* mean? Give an example.
4. What protections are provided by copyright? How do you obtain one?
5. What is fair use?
6. What protections are provided by a patent? How do you obtain one?
7. What does a trademark protect? How do you obtain one?
8. What does the term *personally identifiable information* mean?
9. What protections are provided by HIPAA, COPPA, and FERPA?
10. How would you explain the concept of NORA?

Exercises

1. Provide one example of how information technology has created an ethical dilemma that would not have existed before the advent of information technology.
2. Find an example of a code of ethics or acceptable use policy related to information technology and highlight five points that you think are important.
3. Do some original research on the effort to combat patent trolls. Write a two-page paper that discusses this legislation.
4. Give an example of how NORA could be used to identify an individual.
5. How are intellectual property protections different across the world? Pick two countries and do some original research, then compare the patent and copyright protections offered in those countries to those in the US. Write a two to three pages paper describing the differences.

5.3: Windows Security

Computer Viruses

A computer virus, much like a flu virus, is designed to spread from host to host and can replicate itself. In more technical terms, a computer virus is a type of malicious code or program written to alter the way a computer operates and is designed to spread from one computer to another. A virus operates by inserting or attaching itself to a legitimate program or document that supports macros in order to execute its code. In the process, a virus has the potential to cause unexpected or damaging effects, such as harming the system software by corrupting or destroying data (Norton.com, n.d.).

How does a computer virus attack?

Once a virus has successfully attached to a program, file, or document, the virus will lie dormant until circumstances cause the computer or device to execute its code. For a virus to infect your computer, you must run the infected program, which in turn causes the virus code to be executed. This means that a virus can remain dormant on your computer, without showing major signs or symptoms. However, once the virus infects your computer, the virus can infect other computers on the same network. Stealing passwords or data, logging keystrokes, corrupting files, spamming your email contacts, and even taking over your machine are just some of the devastating and irritating things a virus can do. While some viruses can be playful in intent and effect, others can have profound and damaging effects. This includes erasing data or causing permanent damage to your hard disk. Worse yet, some viruses are designed with financial gains in mind. (Norton.com, n.d.)

How do computer viruses spread?

In a constantly connected world, you can contract a computer virus in many ways, some more obvious than others. Viruses can be spread through email and text message attachments, Internet file downloads, and social media scam links. Your mobile devices and smartphones can become infected with mobile viruses through shady App downloads. Viruses can hide disguised as attachments of socially shareable content such as funny images, greeting cards, or audio and video files. To avoid contact with a virus, it's important to exercise caution when surfing the web, downloading files, and opening links or attachments. To help stay safe, never download text or email attachments that you're not expecting, or files from websites you don't trust. (Norton.com, n.d.)

What are the signs of a computer virus?

A computer virus attack can produce a variety of symptoms. Here are some of them: (Norton.com, n.d.)

- Frequent pop-up windows. Pop-ups might encourage you to visit unusual sites. Or they might prod you to download antivirus or other software programs.
- Changes to your homepage. Your usual homepage may change to another website, for instance. Plus, you may be unable to reset it.
- Mass emails being sent from your email account. A criminal may take control of your account or send emails in your name from another infected computer.
- Frequent crashes. A virus can inflict major damage on your hard drive. This may cause your device to freeze or crash. It may also prevent your device from coming back on.
- Unusually slow computer performance. A sudden change in processing speed could signal that your computer has a virus.
- Unknown programs that startup when you turn on your computer. You may become aware of the unfamiliar program when you start your computer. Or you might notice it by checking your computer's list of active applications.
- Unusual activities like password changes. This could prevent you from logging into your computer.

Computer Malware

Malware is an abbreviated form of "malicious software." This is software that is specifically designed to gain access to or damage a computer, usually without the knowledge of the owner. There are various types of malware, including spyware, ransomware, viruses, worms, Trojan horses, adware, or any type of malicious code that infiltrates a computer. Generally, the software is considered malware based on the intent of the creator rather than its actual features. Malware creation is on the rise due to money that can be made through organized Internet crime. Originally malware was created for experiments and pranks, but eventually, it was used for vandalism and destruction of targeted machines. Today, much of malware is created to make a profit from forced advertising (adware), stealing sensitive information (spyware), spreading email spam or child pornography (zombie computers), or extorting money (ransomware). Various factors can make computers more vulnerable to malware attacks, including defects in the operating system (OS) design, all the computers on a network running the same OS, giving users too many permissions, or just because a computer runs on a particular operating system, such as Windows, for example. The best protection from malware — whether ransomware, bots, browser hijackers, or other malicious software — continues to be the usual, preventive advice: be careful about what email attachments you open, be cautious when surfing by staying away from suspicious websites, and install and maintain an updated, quality antivirus program. (Norton.com, n.d.)

Vulnerabilities and attacks

A vulnerability is a system susceptibility or flaw. Many vulnerabilities are documented in the Common Vulnerabilities and Exposures (CVE) database. An exploitable vulnerability is one for which at least one working attack or "exploit" exists.

To secure a computer system, it is important to understand the attacks that can be made against it, and these threats can typically be classified into one of the categories below:

Backdoors

A backdoor in a computer system, a cryptosystem, or an algorithm, is any secret method of bypassing normal authentication or security controls. They may exist for a number of reasons, including by original design or poor configuration. They may have been added by an authorized party to allow some legitimate access, or by an attacker for malicious reasons; but regardless of the motives for their existence, they create a vulnerability.

Denial-of-service attack

Denial of service attacks are designed to make a machine or network resource unavailable to its intended users. Attackers can deny service to individual victims, such as by deliberately entering a wrong password enough consecutive times to cause the victim account to be locked, or they may overload the capabilities of a machine or network and block all users at once. While a network attack from a single IP address can be blocked by adding a new firewall rule, many forms of Distributed denial of service (DDoS) attacks are possible, where the attack comes from a large number of points – and defending is much more difficult. Such attacks can originate from the zombie computers of a botnet, but a range of other techniques are possible including reflection and amplification attacks, where innocent systems are fooled into sending traffic to the victim.

Direct-access attacks

An unauthorized user gaining physical access to a computer is most likely able to directly copy data from it. They may also compromise security by making operating system modifications, installing software worms, keyloggers, covert listening devices, or using wireless mice. Even when the system is protected by standard security measures, these may be able to be bypassed by booting another operating system or tool from a CD-ROM or other bootable media. Disk encryption and Trusted Platform Module are designed to prevent these attacks.

Eavesdropping

Eavesdropping is the act of surreptitiously listening to a private conversation, typically between hosts on a network. For instance, programs such as Carnivore and NarusInsight have been used by the FBI and NSA to eavesdrop on the systems of internet service providers. Even machines that operate as a closed system (i.e., with no contact to the outside world) can be eavesdropped upon via monitoring the faint electro-magnetic transmissions generated by the hardware; TEMPEST is a specification by the NSA referring to these attacks.

Tampering

Tampering describes a malicious modification of products. So-called "Evil Maid" attacks and security services planting surveillance capability into routers are examples.

Spoofing

Spoofing, in general, is a fraudulent or malicious practice in which communication is sent from an unknown source disguised as a source known to the receiver. Spoofing is most prevalent in communication mechanisms that lack a high level of security.

Clickjacking

Clickjacking, also known as “UI redress attack” or “User Interface redress attack”, is a malicious technique in which an attacker tricks a user into clicking on a button or link on another webpage while the user intended to click on the top-level page. This is done using multiple transparent or opaque layers. The attacker is basically “hijacking” the clicks meant for the top-level page and routing them to some other irrelevant page, most likely owned by someone else. A similar technique can be used to hijack keystrokes. Carefully drafting a combination of stylesheets, iframes, buttons, and text boxes, a user can be led into believing that they are typing the password or other information on some authentic webpage while it is being channeled into an invisible frame controlled by the attacker.

Spam

One of the most annoying emails to receive is junk email. This is also referred to as spam, unsolicited bulk email from cyber criminals or unethical companies. It comes in the form of images and videos.

Phishing

Phishing is a cyber-attack that uses disguised email as a weapon. The goal is to trick the email recipient into believing that the message is something they want or need — a request from their bank, for instance, or a note from someone in their company — and to click a link or download an attachment. What really distinguishes phishing is the form the message takes: the attackers masquerade as a trusted entity of some kind, often a real or plausibly real person, or a company the victim might do business with. It's one of the oldest types of cyberattacks, dating back to the 1990s, and it's still one of the most widespread and pernicious, with phishing messages and techniques becoming increasingly sophisticated. (Norton.com, n.d.)

While there are several paid-for and free applications that protect against viruses and malware, Windows has built-in applications to monitor the health of your PC and protect it against hostile threats.

Internet of Things and physical vulnerabilities

The Internet of Things (IoT) is the network of physical objects such as devices, vehicles, and buildings that are embedded with electronics, software, sensors, and network connectivity that enables them to collect and exchange data – and concerns have been raised that this is being developed without appropriate consideration of the security challenges involved.

While the IoT creates opportunities for more direct integration of the physical world into computer-based systems, it also provides opportunities for misuse. In particular, as the Internet of Things spreads widely, cyber-attacks are likely to become an increasingly physical (rather than simply virtual) threat. If a front door’s lock is connected to the Internet and can be locked/unlocked from a phone, then a criminal could enter the home at the press of a button from a stolen or hacked phone. People could stand to lose much more than their credit card numbers in a world controlled by IoT-enabled devices. Thieves have also used electronic means to circumvent non-Internet-connected hotel door locks.

Medical devices have either been successfully attacked or had potentially deadly vulnerabilities demonstrated, including both in-hospital diagnostic equipment and implanted devices including pacemakers and insulin pumps.

5.4: Summary

Summary

Information security, sometimes shortened to infosec, is the practice of protecting information by mitigating information risks. It is part of information risk management.

Authentication is the process of verification.

Authentication is done in different ways, Multi-factor authentication, and Mutual authentication.

A password is usually used to confirm a user's identity.

Identification is basically the process of someone claiming to be a specific person.

The Information Security Triad, CIA is Confidentiality, Integrity, and Availability.

Use simple steps to make computers more secure.

There are ethical and legal implications while using information systems.

Information about a person that can be used to uniquely establish that person’s identity is called personally identifiable information.

Computer Virus designed to spread from host to host and can replicate themselves.

Malware is software that is specifically designed to gain access to or damage a computer, usually without the knowledge of the owner.

CHAPTER OVERVIEW

6: Networking and Communication

Today's computing and smart devices are expected to be always-connected devices to support the way we learn, communicate, do business, work, and play, in any place, on any devices, and at any time. In this chapter, we review the history of networking, how the Internet works, and the use of multiple networks in organizations today.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Understand how multiple networks are used in everyday life.
- Define how topologies and devices are connected in a small to medium-sized business network.
- Understand the basic characteristics of a network that supports communication in a small to medium-sized business.
- Describe trends in networking that will affect the use of networks in small to medium-sized businesses.

Thumbnail: CCC (Unsplash License; author via source)

6.1: Introduction to Networking and Communication

We are at a basic turning point with many innovations to expand and engage our capacity to communicate. The globalization of the Web has succeeded faster than anybody has envisioned. The way social, commercial, political, and individual motivation happens is quickly changing to keep up with the advancement of this worldwide network. Within our improvement network, innovators will utilize the Web as a beginning point for their efforts, creating modern items and administrations particularly planned to require advantage of the network capabilities. As designers thrust the limits of what is conceivable, the capabilities of the interconnected systems that shape the Web will expand part within these projects' victory.

This chapter presents a brief history of the Internet and the stage of information systems upon which our social and commerce connections progressively depend. The fabric lays the foundation for investigating the administrations, innovations, and issues experienced by network experts as they plan, construct, and keep up the present-day network.

In the Beginning: ARPANET

The story of the Internet and networking can be traced back to the late 1950s. The US was in the Cold War's depths with the USSR, and each nation closely watched the other to determine which would gain a military or intelligence advantage. In 1957, the Soviets surprised the US with the launch of Sputni, propelling us into the space age. In response to Sputnik, the US Government created the Advanced Research Projects Agency (ARPA), whose initial role was to ensure that the US was not surprised again. From ARPA, now called DARPA (Defense Advanced Research Projects Agency), the Internet first sprang. ARPA was the center of computing research in the 1960s, but there was just one problem: many computers could not talk to each other. In 1968, ARPA sent out a request for a communication technology proposal that would allow different computers located around the country to be integrated into one network. Twelve companies responded to the request, and a company named Bolt, Beranek, and Newman (BBN) won the contract and developed the first protocol for the network (Roberts, 1978). They began work right away and completed the job just one year later: in September 1969, the ARPANET was turned on. The first four nodes were at UCLA, Stanford, MIT, and the University of Utah.

The Internet and the World Wide Web

Over the next decade, the ARPANET grew and gained popularity. During this time, other networks also came into existence. Different organizations were connected to different networks. This led to a problem: the networks could not talk to each other. Each network used its own proprietary language or protocol (see sidebar for the definition of protocol) to send information back and forth. This problem was solved using the transmission control protocol/Internet protocol (TCP/IP). TCP/IP was designed to allow networks running on different protocols to have an intermediary protocol that would allow them to communicate. So as long as a network supporting TCP/IP, users could communicate with all other networks running TCP/IP. TCP/IP quickly became the standard protocol and allowed networks to communicate with each other. We first got the term Internet from this breakthrough, which means “an interconnected network of networks.”

As we moved into the 1980s, computers were added to the Internet at an increasing rate. These computers were primarily from government, academic, and research organizations. Much to the engineers' surprise, the early popularity of the Internet was driven by the use of electronic mail (see sidebar below). Using the Internet in these early days was not easy. To access information on another server, you had to know how to type in the commands necessary to access it and know the name of that device. That all changed in 1990 when Tim Berners-Lee introduced his World Wide Web project, which provided an easy way to navigate the Internet through the use of linked text (hypertext). The World Wide Web gained even more steam with the release of the Mosaic browser in 1993, which allowed graphics and text to be combined to present information and navigate the Internet. The Mosaic browser took off in popularity and was soon superseded by Netscape Navigator, the first commercial web browser, in 1994. The chart below shows the growth in internet users globally.

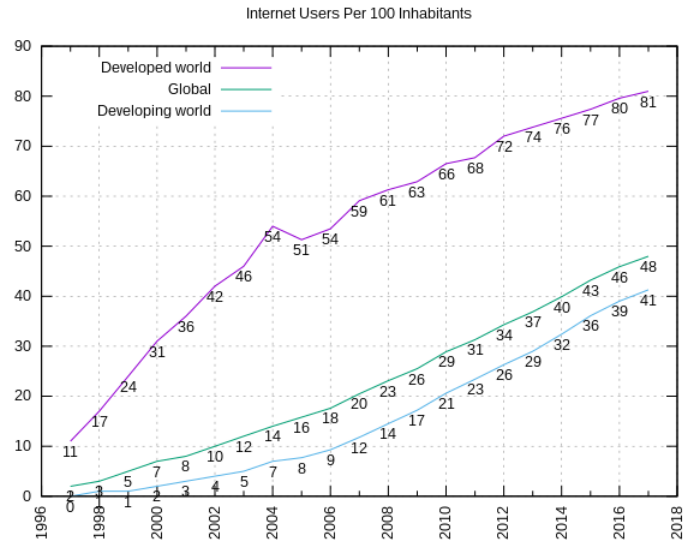


Figure 6.1.1: Graph of "Internet users per 100 inhabitants 1997 to 2017", years on the x-axis, number of users on the y-axis, according to the International Telecommunication Union (ITU). Image by Jeff Ogden (W163) and Jim Scarborough (Ke4roh) is licensed CC BY-SA

According to the International Telecommunications Union (ITU, 2020), over 53.6% or 4.1 billion people worldwide are using the internet, by the end of 2019.

The Internet has evolved from Web 1.0 to 2.0 (discussed in Chapter 1) to the many popular social media websites today.

Networks in Our Daily Lives

Among all of the fundamentals for human presence, the need to interact with others ranks underneath our need to maintain life. Communication is nearly as imperative to us as our dependence on air, water, nourishment, and shelter.

Today, networking systems have enabled people to connect from anywhere. Individuals can communicate and collaborate immediately with others. News ideas and discoveries are shared with the world in seconds. People can indeed interface and play with others without the physical barriers of seas and landmasses from wherever they locate.

Figure 7.3.17.3.1: Global Networking. Image by Gerd Altmann from Pixabay is licensed CC BY 2.0

Technology Then and Now

Envision a world without the Internet. No more Google, YouTube, texting, Facebook, Wikipedia, web-based gaming, Netflix, iTunes, and simple access to current data. No more social media, staying away from lines by shopping on the web, or rapidly looking into telephone numbers and guide headings to different areas at the snap of a finger. How unique would our lives be without the entirety of this? That was the world we lived in only 15 to 20 years back. Throughout the years, information systems have gradually extended and been repurposed to improve personal satisfaction for individuals all over the place.

No Boundaries

Progressions in systems administration advancements are maybe the most noteworthy changes on the planet today. They assist with making a world where national fringes, geographic separations, and physical confinements become less important, introducing ever-lessening obstacles.

CISCO

Figure 7.3.27.3.2: Registered trademark of Cisco Systems, Inc.

Cisco Systems Inc. alludes to this as the human network. The human network fixates on the effect of the Internet and networks on individuals and organizations.

AutoAttribution

- **5.3: Networking Today** by Ly-Huong T. Pham, Tejal Desai-Naik, Laurie Hammond, & Wael Abdeljabbar, is licensed CC BY-NC-SA

References

ITU estimate of global population using the internet. Retrieved September 6, 2020, from https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx

Roberts, Lawrence G., *The Evolution of Packet Switching, (1978, November)*. Retrieved on September 6, 2020, from www.ismlab.usf.edu/dcom/Ch10_Roberts_EvolutionPacketSwitching_IEEE_1978.pdf

6.2: Network Basics

IP Addresses

An IP (Internet Protocol) address is a unique code that identifies a piece of equipment connected to a network. These addresses are used in messages between network devices like the network or wireless card in your computer, the equipment from your ISP (internet service provider), and all pieces of equipment between your machine and the one your computer needs to talk to.

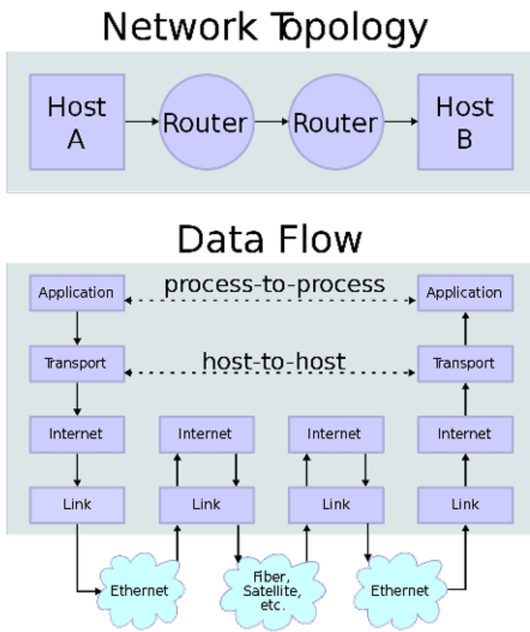


Figure 6.2.1: *Network Topology*. (*By User:Kbrose [CC-BY-SA-3.0] via Wikimedia*)

IP Addresses live in the network layer, which is one of seven layers in the protocol suite defined in the OSI Model. The OSI model stands for Open Systems Interconnection, and was created by the International Organization for Standardization, an international non-governmental group of professionals who strive to establish standards and best practices in a variety of fields. The OSI Model for networking breaks the system of transmitting data into the layers show below in an attempt to delineate where certain actions should take place.

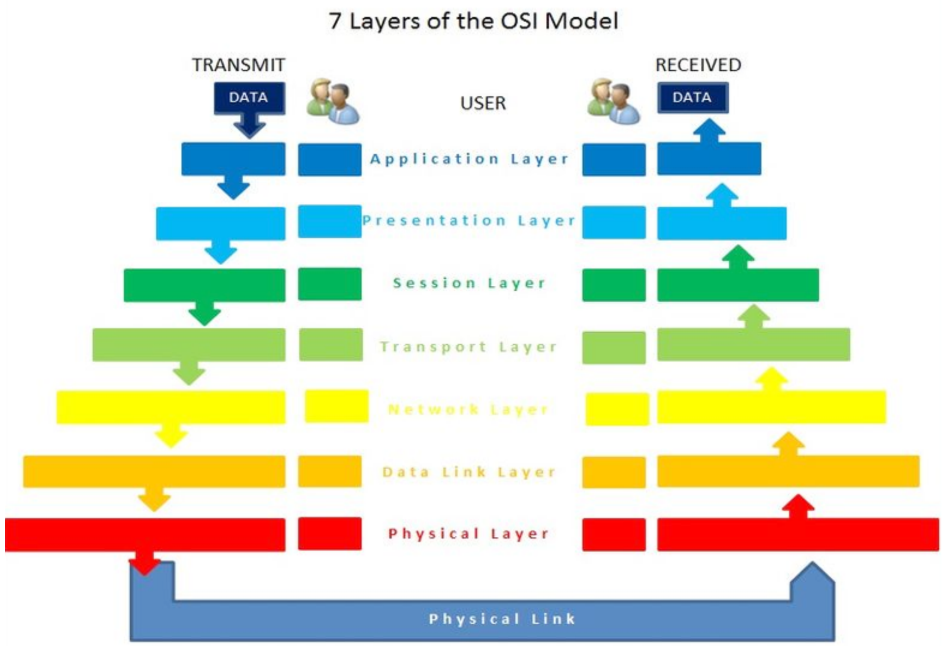


Figure 6.2.2: *OSI 7 Layer Model*. (By MrsValdry [CC-By-SA 3.0] via Wikimedia)

The seven layers depicted above make up the OSI body's recommended protocol suite. In the diagram, transmission of data crosses two routers and over the Internet to reach its destination. By following the data along the arrows, we see it pass through various layers of communication and processing as it crosses the internal network, through the first router, across the public network (internet connection), into the recipient's router, and then is reassembled into its original form.

Until recently, most network equipment has operated on IPv4, the fourth standard released for IP addresses, which has been in place for about thirty years. Addresses in this format are typically represented as a pattern of four blocks of up to three digits separated by periods, with no block of numbers exceeding 255 such as 127.0.0.1 or 24.38.1.251. This is referred to as dot-decimal representation, and although it is not the only way to express an IPv4 address it is the most recognized form. Segments of the addresses within the ranges of 192.168.xxx.xxx, 172.16.xxx.xxx to 172.31.xxx.xxx, and 10.0.xxx.xxx to 10.255.xxx.xxx are reserved for private networks, meaning they are used within a network in your house, at work, or anywhere else where a group of computers share a connection to the internet.

Each of these networks uses one or more of these blocks of numbers for devices on that network. Only the equipment connecting that local network to the Internet needs a unique address from the rest of the world. That equipment will track which computer inside the network to send data to and from by reading packets—the individual pieces of messages that are sent across networks. This means your computer might be 192.168.1.25 at home, and so might your computer at work, according to your home and work networks. The connection between your house and office thought still have a different, unique number assigned to them.

This separation of networks was done to reduce the speed at which unique addresses were consumed. Although this scheme allows for almost 4.3 billion (accurately, 232) addresses, the last one was officially assigned on February 4th, 2012. To sustain today's growing number of devices, IPv6 was created, which is depicted as eight blocks of four hexadecimal digits now separated by colons. These new addresses might look like 2001:0db8:85a3:0042:1000:8a2e:0370:7334, and can support roughly 4 billion unique addresses. Since the new range is so staggeringly large, additional protocols were created that specify when certain values or ranges are used in addresses. This allows additional information about the device to be conveyed just from the address.

The actual messages sent between machines are broken down into multiple pieces. These pieces, called packets, are sent piece by piece from sender to recipient. Each packet is sent the fastest way possible, which means some packets may take different routes—picture a short cut, or getting off a congested road to take a different one. This helps to ensure that the message gets from sender to receiver as fast as possible, but also means packets may arrive in a different order than they were sent.

Additional Notes

Hexadecimal is a number scheme that allows 0 through 9 and A through F as unique values, which means we can count to 15 with one character.

To account for this, each piece of the message, or payload, is wrapped in a header—additional information that describes how many other pieces there are, what protocol is being used, where the packet came from and is headed to, along with some other related information.

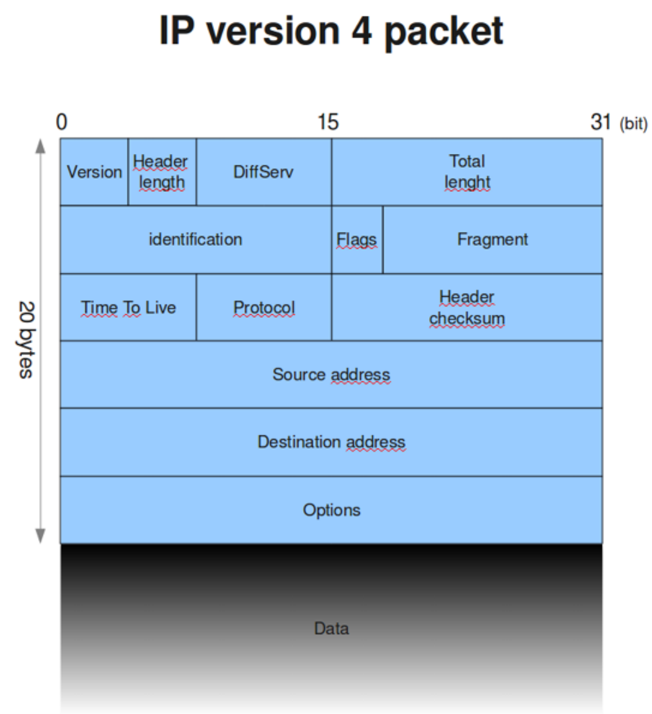


Figure 6.2.3: *IP 4 Packet*. (*Nicolargo [CC-BY-SA-3.0-2.5-2.0-1.0] via Wikimedia Commons*)

After the packets are reassembled, the receiving computer sends any necessary responses, and the process repeats. This all takes place in fractions of a second, beginning with the "hello," or handshake packet to announce a communication request, to the last piece of the packet.

URL

Seeing as most of us would have a hard time remembering what IP address is needed to get to, say, Facebook (173.252.100.16) or the Weather Channel (96.8.80.132) we instead use URLs, universal resource locators. This allows us to use www.facebook.com and www.weather.com to get to where we want to go without referring to a long list of IP addresses. Specialized servers (called name servers) all around the world are responsible for responding to requests from computers for this information. When you type facebook.com into your address bar, if your router does not have a note of its own as to where that is, it will "ask" a name server, which will look it up in its records and reply.

There are three parts to a network address: the protocol, name, and resource id. The protocol represents how we want to send and receive messages, for example we can use http:// for accessing websites and ftp:// for moving files. The name is what we associate with the site, like www.facebook.com, and the resource id, or URI, is everything after that, which points to the particular file we want to see.

Ports

While an IP address and a URL will bring you to a particular web server, there may be more than one way you want to interact with it, or more than one thing you want it to do. Maybe you also want the server to provide email services, or you want to use FTP to update your files. These ports act as different doors into your server, so different applications can communicate without getting in each other's way. Certain ports are typically used for certain activities, for example port 80 is the standard port for web traffic (your browser viewing a page), as opposed to ftp, which typically uses port 21. Using standard ports is not a rule, as applications can be configured to use any available port number, but it is recommended in most cases as firewalls and other security devices may require additional configuring to keep them from blocking wanted traffic because it is arriving at an unusual, fire walled, or "locked" port.

Hosting Facilities

If you are using a server that is not under your physical care, and is managed by an off-site third party, then you likely have an agreement with a hosting facility. Hosting facilities are typically for-profit companies that manage the physical equipment necessary to provide access to websites for a number of clients. Many offer web development and management services as well, but if you are still reading, then that tidbit is probably of little interest as you are here to build it yourself.

Additional Notes

Up Time is the average amount of time that all services on a server are operational and accessible to end users. It is a typical measurement of a hosting company's ability to provide the services they promise.

The benefit of using a hosting service falls under the same principles as other cloud computing services. You are paying to rent equipment and/or services in place of investing in equipment and managing the server and Internet connection yourself. Additionally, hosting facilities are equipped with backup power sources as well as redundant connections to the internet, and may even have multiple facilities that are physically dispersed, ensuring their clients have the best up time as possible. Ads like the one below are common to these services and often emphasize their best features. Price competition makes for relatively affordable hosting for those who are not looking for dedicated servers and are comfortable with sharing their (virtual) server resources with other customers.

Domain Registrar

Domain registrars coordinate the name servers that turn URLs into the IP addresses that get us to our destinations. These companies are where you register available names in order to allow others to find your site. One of the most recognized registrars right now is GoDaddy—you may know them from their ads, which feature racecar driver Danica Patrick. Like many registrars, GoDaddy also offers other services like web and email hosting as well as web development in an effort to solve all of your website needs.

Learn more

Keywords, search terms: Networking, network topology, OSI, network architecture

Cisco Networking Example: http://docwiki.cisco.com/wiki/Internetworking_Basics

List and description of all top level domains: http://www.icann.org/en/resources/registries/tlds

Ongoing comparison of hosting providers: http://www.findmyhosting.com/

6.3: Providing Resources in a Network

Networks of Many Sizes

Networks come in all sizes. They can go from basic networks consisting of two PCs to networks interfacing with many gadgets.

Basic networks introduced in homes empower sharing of assets, for example, printers, archives, pictures, and music between a couple of nearby PCs.

Worldwide internet users expect always to stay connected to the internet. They expect their connected devices to do the following:

- Stay connected to the internet to complete their work.
- Have the ability to send and receive data fast.
- Have the ability to send small and large quantities of data globally via any device connected to the internet.

Home office networks and small office networks are regularly set up by people who work from home or remote offices. They need to associate with a corporate network or other concentrated assets. Moreover, numerous independently employed business people utilize home office and little office networks to publicize and sell items, request supplies and speak with clients.

The Internet is the biggest network presently. Indeed, the term Internet implies a network of networks. The internet is the global worldwide network that connects millions of computers around the world. A computer can connect to another computer in a different country via the internet.

Clients and Servers

All PCs associated with a network are named hosts. Hosts are also called end devices.

Servers are PCs with programming that empower them to give data, similar to emails or website pages, to other network devices called clients. Each assistance requires separate server programming. For instance, a server requires web server programming to give web administrations to the network. A PC with server programming can offer types of assistance at the same time to one or numerous customers. Furthermore, a solitary PC can run numerous sorts of server programming. It might be vital for one PC to go about as a document server, a web server, and an email server in a home or private company.

Clients are PCs with programming introduced that empower them to ask for and show the server's data. A case of client programming is an internet browser, similar to Chrome or Firefox. A solitary PC can likewise run different kinds of custom programming. For instance, a client can browse email and view a site page while texting and tuning in to Internet radio.

Peer-to-Peer

Client and server programming ordinarily run on discrete PCs, yet it is also feasible for one PC to simultaneously complete the two jobs. In private companies and homes, hosts work as servers or clients on the network. This sort of system is known as a shared network. An example of that would be several users connected to the same printer from their individual devices.

Overview of Network Components

The link between the sender and the receiver can be as simple as a single cable connection between these two devices or more sophisticated as a set of switches and routers between them.

Figure 6.3.1: Lan-wan Networks. Image by Stuart Gray is licensed CC BY-SA

The network framework contains three classes of network segments:

- Devices
- Media
- Services

Devices and media are the physical components, or equipment, of the network. Equipment is regularly the noticeable segment of the network stage, for example, a PC, switch, remote passageway, or the cabling used to associate the devices.

Administrations incorporate a significant number of the basic network applications individuals utilize each day, similar to email facilitating administrations and web facilitating administrations. Procedures give the usefulness that coordinates and moves the messages through the network. Procedures are more subtle to us yet are basic to the activity of networks.

End Devices

An end device is either the source or destination of a message transmitted over the network. Each end device is identified by an IP address and a physical address. Both addresses are needed to communicate over a network. IP addresses are unique logical IP addresses that are assigned to every device within a network. If a device moves from one network to another, then the IP address has to be modified.

Physical addresses, also known as MAC (Media Access Control) addresses, are unique addresses assigned by the device manufacturers. These addresses are permanently burned into the hardware.

Intermediary Network Devices

Some devices act as intermediaries between devices. They are called delegated devices. These delegate devices give availability and guarantee that information streams over the network.

Routers utilize the destination end device address, related to data about the network interconnections, to decide how messages should take through the network.

Network Media

A medium called network media carries the act of transport data. The medium gives the channel over which the message makes a trip from source to destination.

Present-day organizations basically utilize three sorts of media to interconnect devices and give the pathway over which information can be transmitted.

These media are:

- Metallic wires within cables (Copper) - information is encoded into electrical driving forces.
- Glass or plastic fibers (fiber optic cable) - information is encoded as beats of light.
- Wireless transmission - information is encoded utilizing frequencies from the electromagnetic range.

Various sorts of network media have various highlights and advantages. Not all network media have similar qualities, nor are they all appropriate for the same purpose.

Figure 6.3.2: Network Cables. Image by blickpixel from Pixabay is licensed CC BY SA

Figure 6.3.3: Fiber Optic Cable. Image by blickpixel from Pixabay is licensed CC BY SA

Bluetooth

While Bluetooth is not generally used to connect a device to the Internet, it is an important wireless technology that has enabled many functionalities that are used every day. When created in 1994 by Ericsson, it was intended to replace wired connections between devices. Today, it is the standard method for connecting nearby devices wirelessly. Bluetooth has a range of approximately 300 feet and consumes very little power, making it an excellent choice for various purposes.

Figure 6.3.4: Bluetooth combo wordmark 2011. Image by House is licensed under Public Domain

Some applications of Bluetooth include: connecting a printer to a personal computer, connecting a mobile phone and headset, connecting a wireless keyboard and mouse to a computer, and connecting a remote for a presentation made on a personal computer.

Network Representations

To draw a diagram of a network, symbols are utilized by network professionals to represent the different devices and connections which make up a network.

A diagram gives a simple method to see how devices in a huge network are associated. This kind of "picture" of a network is known as a topology diagram. The capacity to perceive the legitimate portrayals of the physical systems administration segments is basic to have the option to imagine the association and activity of a network.

Notwithstanding these portrayals, particular phrasing is utilized while discussing how every one of these devices and media interfaces with one another. Significant terms to recall are:

- **Network Interface Card**: A NIC or LAN connector gives the physical association with the PC or opposite end device's network. The media that are associating the PC to the network administration device plug legitimately into the NIC.
- **Physical Port**: A connector or outlet on a network administration device where the media is associated with an end device or another network administration device.

- **Interface**: Specialized ports on a network administration device that associate with singular networks. Since switches are utilized to interconnect networks, the ports on a network allude to network interfaces.

Topology Diagrams

Understanding topology diagrams are required for anybody working with a network. They give a visual guide of how the network is associated.

There are two sorts of Topology diagrams:

- Physical topology and Logical topology diagrams. The physical topology diagrams identify the physical location of intermediary devices and cable installation.
- The Logical topology diagrams identify devices, addressing schemes, and ports.

With physical topology, it is quite self-explanatory. It is how they are interconnected with cables and wires physically. The logical topology is how connected devices are seen to the user.

Types of Networks

Networks foundations can fluctuate extraordinarily regarding:

- Size of the territory secured
- Number of users connected
- Number and kinds of administrations accessible
- Territory of obligation

The two most normal sorts of system frameworks:

- **Local Area Network (LAN)**: A network framework that gives access to clients and end devices in a little topographical zone, commonly an enterprise, small business, home, or small business network possessed and oversaw by an individual or IT department.
- **Wide Area Network (WAN)**: A network foundation that gives access to different networks over a wide topographical region, commonly possessed and overseen by a broadcast communications specialist co-op.

Different kinds of networks include:

- **Metropolitan Area Network (MAN)**: A network foundation that traverses a physical region bigger than a LAN yet littler than a WAN (e.g., a city). Keep an eye on are ordinarily worked by a solitary substance, for example, a huge association.
- **Wireless LAN (WLAN)**: Like a LAN, it remotely interconnects clients and focuses on a little geological region.
- **Storage Area Network (SAN)**: A network foundation intended to help record servers and give information stockpiling, recovery, and replication.

Local Area Networks

LANs are a network foundation that traverses a little topographical territory. Explicit highlights of LANs include:

- LANs interconnect end devices in a restricted region, for example, a home, school, place of business, or grounds.
- A solitary association or person normally directs a LAN. The managerial control that oversees the security and access control arrangements is upheld on the network level.
- LANs give rapid data transfer capacity to inward end gadgets and delegate devices.

Figure 6.3.5: Local Area Network. Image by T.seppelt, derivative work from File:Ethernet.png, including content of the Open Clip Art Library, by © 2007 Nuno Pinheiro & David Vignoni & David Miller & Johann Ollivier Lapeyre & Kenneth Wimer & Riccardo Iaconelli / KDE / LGPL 3, User:George Shuklin and the Tango Project! is licensed CC BY-SA

Wide Area Networks

WANs are a network foundation that traverses a wide topographical zone. WANs are ordinarily overseen by specialist organizations (SP) or Internet Service Providers (ISP).

Explicit highlights of WANs include:

- WANs interconnect LANs over wide geological zones, for example, between urban areas, states, territories, nations, or the mainland.

- Numerous specialist organizations typically manage WANs.
- WANs ordinarily give more slow speed joins between LANs

The Internet

The Internet is an overall assortment of interconnected networks (internetworks or web for short).

A portion of the LAN models is associated with one another through a WAN association. WANs are then associated with one another. The WAN association lines speak to all the assortments of ways we interface networks. WANs can connect through copper wires, fiber optic cables, and wireless transmissions.

No individual or group doesn't own the Internet. Guaranteeing compelling correspondence over this various framework requires the use of steady and generally perceived advances and norms, just as the collaboration of many network organization offices. Some associations have been produced to keep up the structure and normalization of Internet conventions and procedures. These organizations incorporate the Internet Engineering Task Force (IETF), Internet Corporation for Assigned Names and Numbers (ICANN), and the Internet Architecture Board (IAB), in addition to numerous others.

Have you ever wondered how your smartphone can function the way it does? Have you ever wondered how you can search for information on the web and find it within milliseconds? The world's largest implementation of client/server computing and internetworking is the Internet.

The world's largest implementation of client/server computing and internetworking is the Internet. The internet is also a system, which is the most extensive public way of communicating. The internet began in the 20th century; it initially started as a network for the U.S Department of Defense to globally connect university professors and scientists. Most small businesses and homes have access to the internet by subscribing to an internet service provider (ISP), a commercial organization with a permanent connection to the internet, which sells temporary connections to retail subscribers. For example, AT&T, NetZero, and T-Mobile. A DSL (Digital subscriber line) operates over existing telephone lines to carry data, voice, and video transmission rates. The base of the internet is TCP/IP networking protocol suite. When two users on the internet exchange messages, each message is decomposed into packets using the TCP/IP protocol.

Have you ever wondered what happens when you type a URL in the browser and press enter? The browser checks a DNS record in the cache to find the corresponding IP address to the domain. First, you type in a specific URL into your browser. The browser then checks the cache for a DNS record to find the website's corresponding IP address. If the URL is not in the cache, ISP's (Internet Service Provider)'s DNS server starts a DNS query to find the server's IP address that hosts the website. The browser then starts a TCP connection with the server. Then, the browser sends an HTTP request to the webserver. After that, the server handles the request and sends an HTTP response back. Finally, the browser shows the HTML content. For example, www.Wikipedia.org/ has an IP address, that specific IP address could be searched starting with http:// on a browser/ The DNS contains a list of URLs, including their IP addresses.

The DNS (Domain Name System) changes domain names into IP addresses. The domain name is the English name of the thing, and that has 32-bits which are unique and numeric to that English name. To access a computer on the internet, they only need to specify the domain name.

Intranets and Extranets

There are two different terms which are like the term Internet: Intranets and Extranets.

Intranet is a term frequently used to describe a private association of LANs and WANs that has a place with an association. It is intended to be available only for approved individuals, workers, or others of an organization.

An extranet is a term used to describe the case when an organization wants to give secure and safe access to people who work for another organization yet expect access to the association's information. Examples of extranets include:

- An organization that is giving access to outside providers and temporary workers.
- An emergency clinic gives a booking system to specialists so they can make arrangements for their patients.
- A nearby office of training gives spending plans and staff data to the schools in its region.

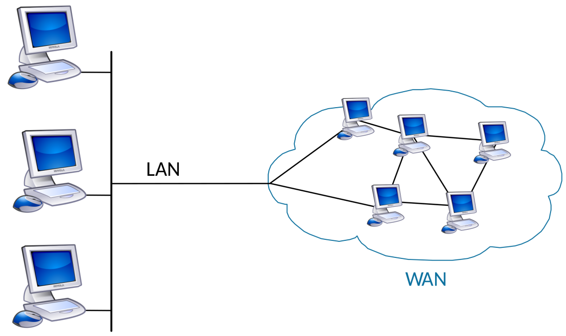


Figure 6.3.2: *LAN WAN scheme. Image by Gateway_firewall.svg: Harald Mühlböck derivative work: Ggia is licensed CC BY-SA*

6.4: Internet Connections

Internet Access Technologies

There is a wide range of approaches to associate users and associations with the Internet.

Home clients (telecommuters) and workplaces regularly require an association with an Internet Service Provider (ISP) to access the Internet. Association alternatives change significantly among ISP and topographical areas. Notwithstanding, companies incorporate a broadband link, broadband computerized endorser line (DSL), remote WANs, and versatile administrations.

Associations commonly expect access to other corporate destinations and the Internet. Quick associations are required to help business administrations, including IP telephones, video conferencing, and server farm stockpiling.

Business-class interconnections are normally given by specialist organizations (SP). Well-known business-class administrations incorporate business DSL, rented lines, and Metro Ethernet.

Home and Small Office Internet Connections

Regular connection choices for little office and home office users:

- **Cable**: Typically offered by digital TV specialist co-ops, the Internet information signal is carried on a similar link that conveys satellite TV. It gives a high transmission capacity, consistently on, association with the Internet.
- **DSL**: Digital Subscriber Lines gives a high data transmission, consistently on, association with the Internet. DSL runs over a phone line when all is said in done, small office and home office clients associate utilizing Asymmetrical DSL (ADSL), which implies that the download speed is quicker than the upload speed.
- **Cellular**: For a Cell phone network to connect, it utilizes cellular internet access. Any place you can get a phone signal, you can get cell Internet. Execution will be restricted by the telephone's abilities and the cell tower to which it is associated. The fourth generation of broadband cellular network technology is 4G, which most people are familiar with because it is on smartphones. 5G is upcoming and expected to be faster than and succeed 4G by 100 times, which will have the ability to transmit a lot more data at a much faster pace than 4G.
- **Satellite**: Internet access through satellite is a genuine advantage in those territories that would somehow or another have no Internet availability by any means. Satellite dishes require a clear line of sight to the satellite.
- **Dial-up telephone**: An economical choice that utilizes any telephone line and a modem. The low transmission capacity supported by a dial-up modem association is normally not adequate for huge information transfer. However, it is still a valuable choice wherever other options are not available such as in rural areas or remote locations where phones are the only means of communication.

Fiber optic links are increasingly becoming more available to home and small businesses. This empowers an ISP to give higher data transmission speeds and bolster more administrations, for example, Internet, telephone, and TV.

Business Internet Connections

Corporate connection choices contrast from home client alternatives. Organizations may require higher transmission capacity, devoted data transmission, and oversaw administrations. Business connection options include:

- **Dedicated Leased Line**: Leased lines are really saved circuits inside the specialist organization's system that interface geologically isolated workplaces for private voice or potentially information organizing. The circuits are ordinarily leased at a month-to-month or yearly rate. They can be costly.
- **Ethernet WAN**: Ethernet WANs broaden LAN access into the WAN. Ethernet is a LAN innovation you will find out about in a later section. The advantages of Ethernet are currently being reached out into the WAN.
- **DSL**: Business DSL is accessible in different organizations. A famous decision is Symmetric Digital Subscriber Lines (SDSL) which is like the purchaser rendition of DSL. However, it gives transfers and downloads at similar paces.
- **Satellite**: Like small office and home office clients, satellite help can give an association when a wired arrangement isn't accessible.

The decision of connection shifts relying upon topographical area and specialist organization accessibility.

Sidebar: An Internet Vocabulary Lesson

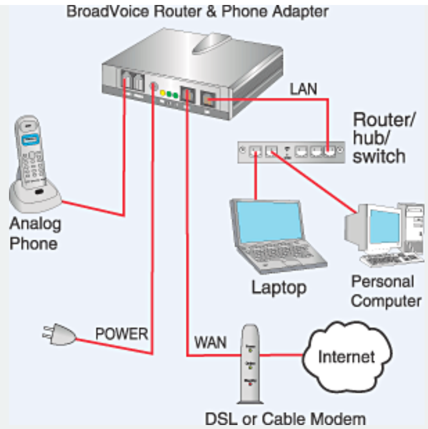


Figure 6.4.1: Devices connection. Image by BroadVoice is licensed CC BY 1.0

Networking communication is full of some very technical concepts based on some simple principles. Learn the terms below, and you will be able to hold your own in a conversation about the Internet.

- **Packet**: The fundamental unit of data transmitted over the Internet. When a device intends to send a message to another device (for example, your PC sends a request to YouTube to open a video), it breaks the message down into smaller pieces, called packets. Each packet has the sender's address, the destination address, a sequence number, and a piece of the overall message to be sent.
- **Hub**: A simple network device connects other devices to the network and sends packets to all the devices connected to it.
- **Bridge**: A network device that connects two networks and only allows packets through that are needed.
- **Switch**: A network device that connects multiple devices and filters packets based on their destination within the connected devices.
- **Router**: A device that receives and analyzes packets and then routes them towards their destination. In some cases, a router will send a packet to another router; it will send it directly to its destination in other cases.
- **IP Address**: Every device that communicates on the Internet, whether it be a personal computer, a tablet, a smartphone, or anything else, is assigned a unique identifying number called an IP (Internet Protocol) address. Historically, the IP-address standard used has been IPv4 (version 4), which has the format of four numbers between 0 and 255 separated by a period. For example, the domain Saylor.org has an IP address of 107.23.196.166. The IPv4 standard has a limit of 4,294,967,296 possible addresses. As the use of the Internet has proliferated, the number of IP addresses needed has grown to the point where IPv4 addresses will be exhausted. This has led to the new IPv6 standard, which is currently being phased in. The IPv6 standard is formatted as eight groups of four hexadecimal digits, such as 2001:0db8:85a3:0042:1000:8a2e:0370:7334. The IPv6 standard has a limit of 3.4×1038 possible addresses. For more detail about the new IPv6 standard, see this Wikipedia article.
- **Domain name**: If you had to try to remember the IP address of every web server you wanted to access, the Internet would not be nearly as easy to use. A domain name is a human-friendly name for a device on the Internet. These names generally consist of a descriptive text followed by the top-level domain (TLD). For example, Wikipedia's domain name is Wikipedia.org; Wikipedia describes the organization, and .org is the top-level domain. In this case, the .org TLD is designed for nonprofit organizations. Other well-known TLDs include .com , .net , and .gov . For a complete list and description of domain names, see this Wikipedia article.
- **DNS**: DNS stands for "domain name system," which acts as the directory on the Internet. A DNS server is queried when a request to access a device with a domain name is given. It returns the IP address of the device requested, allowing for proper routing.
- **Packet-switching**: When a packet is sent from one device out over the Internet, it does not follow a straight path to its destination. Instead, it is passed from one router to another across the Internet until it reaches its destination. In fact, sometimes, two packets from the same message will take different routes! Sometimes, packets will arrive at their destination out of order. When this happens, the receiving device restores them to their proper order. For more details on packet switching, see this interactive web page.
- **Protocol**: In computer networking, a protocol is the set of rules that allow two (or more) devices to exchange information back and forth across the network.

6.5: The Changing Network Environment Network Trends

New Trends

As new technologies and end-user devices come to market, businesses and purchasers must keep on acclimating to this ever-evolving condition. The job of the network is changing to empower the connections between individuals, devices, and data. There are a few new networking trends that will impact organizations and purchasers. A portion of the top trends include:

- Bring Your Own Device (BYOD)
- Video communications
- Online collaboration
- Cloud computing

Bring Your Own Device

The idea of any device, to any content, in any way, is a significant worldwide trend that requires huge changes to the manner in which devices are utilized. This trend is known as Bring Your Own Device (BYOD).

BYOD is about end users having the opportunity to utilize individual tools in order to get to data and convey over a business or campus network. With the development of customer devices and the related drop in cost, representatives and students can be relied upon to have probably the most progressive computing and networking tools for individual use. These individual tools can be laptops, e-books, tablets, cell phones, and tablets. These can be devices bought by the organization or school, bought by the individual, or both.

BYOD implies any device, with any possession, utilized anyplace. For instance, previously, a student who expected to get access to the campus network or the Internet needed to utilize one of the school's PCs. These devices were commonly constrained and seen as instruments just for work done in the study hall or in the library. Expanded availability through portable and remote access to the campus network gives students a lot of adaptability and opens doors of learning for the student.

Online Collaboration

People want to connect with the network, for access to data applications, in addition to team up with each other.

Collaboration is characterized as "the demonstration of working with another or others on a joint venture." Collaboration tools, give representatives, students, instructors, clients, and accomplices an approach to quickly interface, connect, and accomplish their targets.

For businesses, collaboration is a basic and vital need that associations are utilizing to sustain their competition. Collaboration is additionally a need in training. Students need to work together to help each other in learning, to create group abilities utilized in the workplace, and to cooperate on group based projects.

Video Communication

Another trend in networking that is basic to the correspondence and joint effort exertion is video. Video is being utilized for interchanges, cooperation, and amusement. Video calls can be made to and from anyplace with an Internet connection.

Figure 6.5.1: A video call showing a group of people on the screen. Image by photo by Chris Montgomery on Unsplash is licensed under CC BY SA 2.0

Video conferencing is an incredible asset for speaking with others from a distance, both locally and worldwide. Video is turning into a basic necessity for successful joint effort as associations stretch out across geographic and social limits.

Cloud Computing

Cloud computing is another worldwide trend changing how we access and store information. Cloud computing permits us to store individual files, even backup our whole hard disk drive on servers over the Internet. Applications, for example, word processing, and photograph editing, can be accessed utilizing the Cloud.

When it comes to businesses, cloud computing expands IT's capabilities without requiring interest in new infrastructure, preparing new faculty, or permitting new software. These services are accessible on request and conveyed economically to any device on the planet without trading off security or capacity.

There are four essential Clouds: Public Clouds, Private Clouds, Hybrid Clouds, and Custom Clouds. Snap each Cloud to find out additional.

Cloud computing is conceivable because of data centers. A data center is an office used to house PC frameworks and related parts. A data center can consume one room of a building, at least one story, or the whole thing. Data centers are commonly over the top expensive to manufacture and keep up. Therefore, just huge associations utilize secretly fabricated data centers to house their information and offer users assistance. Smaller associations that can't afford to keep up their own private data center can lessen the general expense of ownership by renting server and capacity services from a bigger data center association in the Cloud.

Technology Trends in the Home

Networking trends are not just influencing how we work or study, and they are also changing pretty much every part of the home.

The most up-to-date home trends incorporate smart home technology, a technology that is coordinated into habitual appliances, permitting them to interconnect with different devices, making them progressively 'smart' or automated. For instance, envision having the option to set up a dish and spot it in the broiler for cooking before going out for the afternoon. Envision if the stove knew of the dish it was cooking and was associated with your 'schedule of occasions' so it could figure out what time you will be eating and change start times and length of cooking accordingly. It could even modify cooking times and temperatures dependent on plan changes. Furthermore, a cell phone or tablet connection permits the user to interface with the broiler straightforwardly to make any ideal changes. When the dish is "accessible," the stove sends an alarm message to a predefined end-user device that the dish is done and warming.

This situation isn't long-off. Actually, smart home technology is being created for all rooms inside a house. It will turn out to be a greater degree of reality as home networking and high-speed Internet technology become progressively far-reaching. New home networking technologies are being grown day by day to meet these sorts of developing technology needs.

Powerline Networking

Powerline networking is a rising trend for home networking that utilizes existing electrical wiring to connect devices.

The idea of "no new wires" signifies the capacity to connect a device to the network where there is an electrical outlet. This spares the expense of introducing data cables and with no extra expense to the electrical bill. Utilizing similar wiring that conveys power, powerline networking sends information by sending data on specific frequencies.

Utilizing a standard powerline adapter, devices can connect with the LAN any place there is an electrical outlet. Powerline networking is beneficial when wireless access points can't be utilized or can't arrive at all to the devices in the home. Powerline networking isn't intended to fill in for committed cabling in data networks. But it is an alternative when data network cables or wireless communications are not a reasonable choice.

Wireless Broadband

Connecting with the Internet is indispensable in savvy home innovation. DSL and cable are basic advances used to connect homes and private companies to the Internet. Nonetheless, remote access might be another choice in numerous regions.

Another remote answer for home and independent companies is wireless broadband. This uses the equivalent cell innovation to get to the Internet with an advanced mobile phone or tablet. A radio wire is introduced outside the house, giving either remote or wired availability for home devices. In numerous zones, home wireless broadband is contending legitimately with DSL and cable services.

Wireless Internet Service Provider (WISP)

Wireless Internet Service Provider (WISP) is an ISP that connects subscribers of an assigned passage or problem area utilizing comparable remote innovations found in-home wireless local area networks (WLANs). WISPs are all the more usually found in provincial situations where DSL or cable services are not accessible.

Though a different transmission tower might be introduced for the antenna, the antenna is usually connected to a current raised structure, such as a water tower or a radio pinnacle. A little dish or radio wire is introduced on the subscriber's rooftop in the WISP transmitter's scope. The subscriber's entrance unit is associated with the wired system inside the home. From the home user's point of view, the arrangement isn't vastly different from DSL or cable service. The principle distinction is that the home's connection to the ISP is remote rather than a physical link.

Sidebar: Why Doesn’t My Cell Phone Work When I Travel Abroad?

As mobile phone technologies have evolved, providers in different countries have chosen different communication standards for their mobile phone networks. In the US, both of the two competing standards exist GSM (used by AT&T and T-Mobile) and CDMA (used by the other major carriers). Each standard has its pros and cons, but the bottom line is that phones using one standard cannot easily switch to the other.

In the US, this is not a big deal because mobile networks exist to support both standards. But when you travel to other countries, you will find that most of them use GSM networks, with the one big exception being Japan, which has standardized on CDMA. It is possible for a mobile phone using one type of network to switch to the other type of network by switching out the SIM card, which controls your access to the mobile network. However, this will not work in all cases. If you are traveling abroad, it is always best to consult with your mobile provider to determine the best way to access a mobile network.

6.6: Network Security

Security Threats

Network security is an indispensable piece of computer networking today, whether or not the network is restricted to a home domain with a solitary connection with the Internet or as extensive as an organization with many users. The network security that is executed must consider the environment, just as the system's devices and prerequisites. It must have the option to keep the data secure while considering the quality of service anticipated from the network.

Ensuring a network is secure includes technologies, protocols, devices, tools, and techniques to keep data secure and moderate threat vectors. Threat vectors might be external or internal. Numerous external network security threats today are spread over the Internet.

The most widely recognized external threats to networks include:

- Viruses, worms, and Trojan horses- malignant programming and subjective code running on a client device
- Spyware and adware - software installed on a user device that covertly gathers data about the user Zero-day attacks, likewise called zero-hour attacks - an assault that happens on a principal day that a defenselessness gets known
- Hacker attacks- an assault by an educated individual to user devices or network assets
- Denial of service attacks- assaults intended to slow or crash applications and procedures on a network device
- Data interception and theft - an assault to catch private data from an association's network
- Identity theft- an assault to take the login qualifications of a user to get to private information

It is similarly critical to think about internal threats. There have been numerous examinations showing that the most well-known data breaches happen due to the network's internal users. This can be credited to lost or taken devices, inadvertent abuse by workers, and in the business condition, even malignant representatives. With the advancing BYOD systems, corporate information is considerably more powerless. Accordingly, it is critical to address both outside and interior security dangers when building up a security strategy.

Security Solutions

No single arrangement can shield the network from the many threats that exist. Consequently, security ought to be implemented in various layers, utilizing more than one security arrangement. If one part of the security fails to recognize and shield the network, others will stand.

A home network security execution is typically rather essential. It is commonly executed on the interfacing end devices, just as connected with the Internet, and can even depend on contracted services from the ISP.

Conversely, the network security implementation for a corporate network, for the most part, comprises numerous segments incorporated with the network to screen and channel traffic.

In a perfect world, all segments cooperate, which limits maintenance and improves overall security.

Network security parts for a home or little office network should at least incorporate the following:

- Antivirus and antispyware: These are utilized to shield end devices from getting contaminated with vindictive software.
- Firewall filtering: This is utilized to prevent unapproved access to the network. This may incorporate a host-based firewall system that is actualized to forestall unapproved access to the end device or an essential separating service on the home router to keep unapproved access from the outside world into the network.

Bigger networks and corporate networks frequently have other security necessities:

- Dedicated firewall systems: These are utilized to develop further firewall abilities that can channel a lot of traffic with greater granularity.
- Access control lists (ACL): These are utilized to channel access and traffic sending additionally.
- Intrusion prevention systems (IPS): These are utilized to distinguish quick-spreading dangers, for example, zero-day or zero-hour assaults.
- Virtual Private Networks (VPN): These are utilized to give secure access to telecommuters.

Networks security necessities must consider the network condition, just like the different applications and processing prerequisites. Both home situations and organizations must have the option to secure their data yet consider the quality of service that is anticipated from every innovation. Furthermore, the security arrangement executed must be versatile to the developing and changing trends of the network.

The study of network security dangers and relief strategies begins with a concise understanding of the underlying switching and routing infrastructure utilized to organize network services.

6.7: Summary

Summary

Networks and the Internet have changed how we impart, learn, work, and even play.

Networks come in all sizes. They can run from basic networks consisting of two PCs to networks connecting a large number of devices.

The Internet is the biggest network presently. Truth be told, the term Internet implies a 'network of networks.'

The Internet offers the types of assistance that empower us to interface and speak with our families, companions, work, and interests.

The network foundation is the stage that underpins the network. It gives the steady and dependable channel over which correspondence can happen. It comprises network parts, including end devices, halfway devices, and network media.

Networks must be dependable. This implies the network must be tolerant to flaws, adaptable, give quality of service, and guarantee the network's data and assets. Network security is a basic piece of PC networking, whether or not the network is restricted to a home situation with a solitary connection with the Internet or as extensive as an enterprise with many users. No single arrangement can shield the network from the assortment of dangers that exist. Consequently, security ought to be executed in numerous layers, utilizing more than one security arrangement.

The network infrastructure can change significantly based on size, many users, and the sorts of upheld administrations. The network infrastructure must develop and change by how the network is utilized. The routing and switching stage is the establishment of any networked framework.

CHAPTER OVERVIEW

7: Databases

Introduction

You have already been introduced to the first two components of information systems: hardware and software. However, those two components by themselves do not make a computer useful. Imagine if you turned on a computer, started the word processor, but could not save a document. Imagine if you opened a music player but there was no music to play. Imagine opening a web browser but there were no web pages. Without data, hardware and software are not very useful! Data is the third component of an information system.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Describe the differences between data, information, and knowledge
- Describe Database, Database Management System, and Database Types
- Characterize the contribution of databases to websites
- Identify relational database elements
- Identify Fields, Records, and Tables
- Recognize that tables can have relationships
- Identify Data Warehouse and Data Mining
- Recognize Meta Data and their use
- Recognize the need for database security

Thumbnail: panumas nikhomkhai https://www.pexels.com/photo/bandwid...ction-1148820/

7.1: Data and Databases

Introduction

You have already been introduced to hardware and software. However, those two components by themselves do not make a computer useful. Imagine if you turned on a computer, started the word processor, but could not save a document. Imagine if you opened a music player but there was no music to play. Imagine opening a web browser but there were no web pages. Without data, hardware and software are not very useful! Data is the third component of an information system.

Data, Information, and Knowledge

Data are the raw bits and pieces of information with no context. If I told you, "15, 23, 14, 85," you would not have learned anything. But I would have given you data.

Data can be quantitative or qualitative. Quantitative data is numeric, the result of a measurement, count, or some other mathematical calculation. Qualitative data is descriptive. "Ruby Red," the color of a 2013 Ford Focus, is an example of qualitative data. A number can be qualitative too: if I tell you my favorite number is 5, that is qualitative data because it is descriptive, not the result of a measurement or mathematical calculation.

By itself, data is not that useful. To be useful, it needs to be given context. Returning to the example above, if I told you that "15, 23, 14, and 85″ are the numbers of students that had registered for upcoming classes, that would be *information*. By adding the context – that the numbers represent the count of students registering for specific classes – I have converted data into information.

Once we have put our data into context, aggregated and analyzed it, we can use it to make decisions for our organization. We can say that this consumption of information produces *knowledge*. This knowledge can be used to make decisions, set policies, and even spark innovation.

The final step up the information ladder is the step from knowledge (knowing a lot about a topic) to *wisdom*. We can say that someone has wisdom when they can combine their knowledge and experience to produce a deeper understanding of a topic. It often takes many years to develop wisdom on a particular topic and requires patience.

Examples of Data

Almost all software programs require data to do anything useful. For example, if you are editing a document in a word processor such as Microsoft Word, the document you are working on is the data. The word-processing software can manipulate the data: create a new document, duplicate a document, or modify a document. Some other examples of data are: an MP3 music file, a video file, a spreadsheet, a web page, and an e-book. In some cases, such as with an e-book, you may only have the ability to read the data.

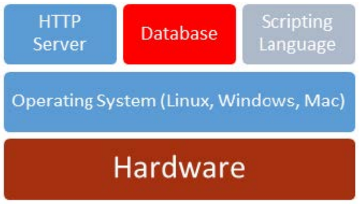
Databases

The goal of many information systems is to transform data into information in order to generate knowledge that can be used for decision making. In order to do this, the system must be able to take data, put the data into context, and provide tools for aggregation and analysis. A database is designed for just such a purpose.

A database is an organized collection of related information. It is an *organized* collection, because in a database, all data is described and associated with other data. All information in a database should be *related* as well; separate databases should be created to manage unrelated information. For example, a database that contains information about students should not also hold information about company stock prices. Databases are not always digital – a filing cabinet, for instance, might be considered a form of a database. For the purposes of this text, we will only consider digital databases.

Database Types

While there are a number of databases available for use like MySQL, node.js, and Access, there is an additional list of the types of database structures each of these belongs to. These types each represent a different organizational method of storing the information and denoting how elements within the data relate to each other. We will look at the three you are most likely to come across in the web development world, but this is still not an exhaustive representation of every approach available.



Flat File

Flat files are flat in that the entire database is stored in one file, usually separating data by one or more delimiters, or separating markers, that identify where elements and records start and end. If you have ever used Excel or another spreadsheet program, then you have already interacted with a flat-file database. This structure is useful for smaller amounts of data, such are spreadsheets and personal collections of movies. Typically these are comma-separated value files, .csv being a common extension. This refers to the fact that values in the file are separated by commas to identify where they start and end, with records marked by terminating characters like a new line, or by a different symbol like a colon or semi-colon.

The nature of all data being in one file makes the database easily portable, and somewhat human-readable. However, information that would be repeated in the data would be written out fully in each record. Following our movie example, this could be the producer, studio, or actors. They have what we call a one-to-many relationship with other data in the database that cannot be tracked in this format.

Drawbacks to this format can affect your system in several ways. First, in our example, we would enter the studio name into each record for a movie made by that studio. If the person typing miss-typed an entry, it may not be found when users search the database, skewing search results through missing information. This often results in users creating new entries for a record that appears not to exist, causing duplication. Beyond search issues, every repetition is more space the database requires, especially when the value repeated is large. This was more of an issue when data was growing faster than storage capacity. Now, with the exception of big data storage systems, the average user's storage space on an entry-level PC typically surpasses the user's needs unless they are avid music or movie collectors. It is still important to consider when you are dealing with limited resources, such as mobile applications that are stored on smartphones that have memory limitations lower than desktops and servers.

Another issue with these files is the computational effort required to search the file, edit records, and insert new ones that are placed somewhere within the body of data as opposed to the start or end of the file.

Finally, flat files are not well suited for multiple, concurrent use by more than one party. Since all of the data is in one file, you are faced with two methods of interaction. The first approach is allowing anyone to access the file at the same time, usually by creating a local temporary copy on their system. While this allows multiple people the ability to use the file, if more than one party is allowed to make changes, we risk losing data. Say User 1 creates two new records while User 2 is editing one. If User 2 finished first and saves their changes, they are written back to the server. Then, User 1 sends their changes back, but their system is unaware of the changes made by User 2. When their changes are saved, User 2's changes are lost as User 1 overwrites the file. This can be prevented by checking the last modified timestamps before allowing data to be written, but the user "refreshing" may have conflicts between their edits and the edits from another user when the same record is changed.

The alternate approach to allowing multiple users is to block multiple users from making changes by only allowing one of them to have the file open at a time. This is done by creating a file lock, a marker on the file the operating system would see, that would block other users from using an open file. This approach does not support concurrent access to the data, and again even allowing read rights to other users would not show them changes in progress that another user has not completed and submitted. Another downside to this approach is what is called a race condition—where multiple systems are trying to access a file, but are unable to do so because another has the file locked, stalling all of the programs trying to access the data.

This was a key element in a large-scale blackout of 2003 that took place in the Northeast United States and part of Canada. A summer heatwave created a significant strain on the power system as demand for air conditioning increased, resulting in the emergency shutdown of a power station. This station happened to be editing its health status in a shared file between itself and other stations, a method used to allow other stations to adjust their operational levels in response to their neighbors. The purpose of this file was to act as a protection method, warning of potential spikes or drops in power at individual facilities. When the plant using the file shut down, the file remained locked as the computer using it did not have time to send a close file command. Unable to properly close the file with the systems down, other stations were unaware of the problem until power demand at their facilities rapidly increased. As these stations could not access the file due to the lock, a warning could not be passed along. Since the power stations were under increasing strain with each failure, a cascading effect occurred throughout the entire system. Admittedly an extreme result of the file lock failure, it is a very real-world example of the results of using the wrong tools when designing a system.

Structured Query/Relational Database

Structured query databases can be viewed similar to flat files in that the presentation of a section of data can be viewed as its own table, similar to a single spreadsheet. The difference is that instead of one large file, the data is broken up based on user needs and by grouping related data together into different tables. You could picture this as a multi-page spreadsheet, with each page containing different information. For example, continuing with our movie example, one table would contain everything about the studio—name, opening date, tax code, and so on. The next table would contain everything about the movies—name, release date, description, production cost, etc. Finally, we might have a table for actors, producers, and everyone else involved. This table would have their information like birthday, hometown, and more.

What we do not have yet is a way to link these elements together. There is also a lot of information we *do not* want to include, because we can determine it from something else. For example, we do not want to store the actor's age, or we would have to update the table every year on their birthday. Since we already have their birth date, we can have the server do the math based on the current date and their birth date to determine how old they are each time it is asked.

To address relating an actor in our people table to a movie they were in from the movie table, as well as to the studio that made the movie in the studio table, we use a structured query. The structured query is a human-readable (relatively) sentence-style language that uses a fixed vocabulary to describe what data we want and how to manipulate it. Part of this comes from adding extra tables. Since one actor can be in many movies, and each movie can have many actors, we have a many-to-many relationship between them. Due to this, we create an extra table where each row represents a record of an actor and a movie they were in. Instead of putting their full names into this table, we put the row number that identifies their information from their respective tables. This gives us a long, skinny table that is all numbers, called an “all-reference table,” as it refers to other tables and does not contain any new information of its own. We will see this in action soon.

We can use our query language to ask the database to find all records in this skinny table where the movie ID matches the movie ID in the movie table, and also where the movie name is “Die Hard.” The query will come back with a list of rows from our skinny table that has the value in the movie ID column. We can also match the actor IDs from a table that pairs actors with movies to records in the actor table in order to get their names. We could do this as two different steps or in one larger query. In using the query, we recreate what would have been one very long record in our flat file. The difference is we have done it with a smaller footprint, reduced mistyping errors, and only see exactly what we need to. We can also “lock” data in a query database at the record level, or a particular row in a database when editing data, allowing other users access to the rest of the database.

While this style can be very fast and efficient in terms of storage size, interacting with the data through queries can be difficult as both one-to-many and many-to-many relationships are best represented through intermediary tables as we described above (one-to-one relationships are typically found within the same table, or as a value in one table directly referencing another table). In order to piece our records together, we need an understanding of the relationships between the data.

MySQL

Structured query language databases are a very popular data storage method in web development. MySQL, commonly pronounced as “my seequl” or “my s q l,” is a relational database structure that is an open source implementation of the structured query language. The relational element arises from the file structure, which in this case refers to the fact that data is separated into multiple files based on how the elements of the data relate to one another in order to create a more efficient storage pattern that takes up less space.

MySQL plays the role of our data server, where we will store information that we want to be able to manipulate based on user interaction. Contents are records, like all the items available in Amazon’s store. User searches and filters affect how much of the data is sent from the database to the web server, allowing the page to change at the user’s request.

Structure

We organize data in MySQL by breaking it into different groups, called tables. Within these tables are rows and columns, in which each row is a record of data and each column identifies the nature of the information in that position. The intersection of a row and column is a cell or one piece of information. Databases are collections of tables that represent a system. You can imagine a database server like a file cabinet. Each drawer represents a database on our server. Inside those drawers are folders that hold files.

The folders are like our tables, each of which holds multiple records. In a file cabinet, our folders hold pieces of paper or records, just like the individual rows in a table. While this may seem confusing now, we will see it in action soon; this is the approach we will focus on for this section of the text.

Unstructured

Unstructured data, typically categorized as qualitative data, cannot be processed and analyzed via conventional data tools and methods. Since unstructured data does not have a predefined data model, it is best managed in non-relational (NoSQL) databases. Another way to manage unstructured data is to use data lakes to preserve it in raw form.

The importance of unstructured data is rapidly increasing. Recent projections indicate that unstructured data is over 80% of all enterprise data, while 95% of businesses prioritize unstructured data management.

NoSQL

NoSQL databases represent systems that maintain collections of information that do not specify relationships within or between each other. In reality, a more appropriate name would be NoRel or NoRelation as the focus is on allowing data to be more free form.

Most NoSQL systems follow a key-value pairing system where each element of data is identified by a label. These labels are used as consistently as possible to establish common points of reference from file to file, but may not be present in each record. Records in these systems can be represented by individual files. In MongoDB, the file structure is a single XML formatted record in each file, or it can be ALL records as a single XML file. Searching for matches in a situation like this involves analyzing each record, or the whole file, for certain values.

These systems excel when high numbers of static records need to be stored. The more frequently data needs to be changed, the more you may find performance loss here. However, searching these static records can be significantly faster than relational systems, especially when the relational system is not properly normalized. This is actually an older approach to data storage that has been resurrected by modern technology's ability to capitalize on its benefits, and there are dozens of solutions vying for market dominance in this sector. Unless you are constructing a system with big data considerations or high volumes of static records, relational systems are still the better starting place for most systems.

Semi-structured

Semi-structured data (e.g., JSON, CSV, XML) is the "bridge" between structured and unstructured data. It does not have a predefined data model and is more complex than structured data, yet easier to store than unstructured data.

Semi-structured data uses "metadata" (e.g., tags and semantic markers) to identify specific data characteristics and scale data into records and preset fields. Metadata ultimately enables semi-structured data to be better cataloged, searched, and analyzed than unstructured data.

- **Example of metadata usage:** An online article displays a headline, a snippet, a featured image, image alt-text, slug, etc., which helps differentiate one piece of web content from similar pieces.
- **Example of semi-structured data vs. structured data:** A tab-delimited file containing customer data versus a database containing CRM tables.
- **Example of semi-structured data vs. unstructured data:** A tab-delimited file versus a list of comments from a customer's Instagram.

Exercises

1. Discuss each of the following terms:
   1. data
   2. Information
   3. Knowledge
2. Explain the difference between data and information.

7.2: Before Database Systems

Before the Advent of Database Systems

The way in which computers manage data has come a long way over the last few decades. Today's users take for granted the many benefits found in a database system. However, it wasn't that long ago that computers relied on a much less elegant and costly approach to data management called the file-based system.

File-based System

One way to keep information on a computer is to store it in permanent files. A company system has a number of application programs; each of them is designed to manipulate data files. These application programs have been written at the request of the users in the organization. New applications are added to the system as the need arises. The system just described is called the *file-based system*.

Consider a traditional banking system that uses the file-based system to manage the organization's data shown in Figure 1.1. As we can see, there are different departments in the bank. Each has its own applications that manage and manipulate different data files. For banking systems, the programs may be used to debit or credit an account, find the balance of an account, add a new mortgage loan, and generate monthly statements.

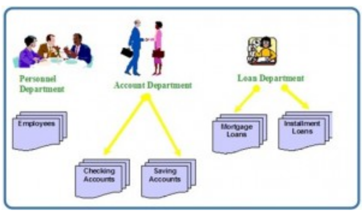


Figure 1.1. Example of a file-based system used by banks to manage data.

Disadvantages of the file-based approach

Using the file-based system to keep organizational information has a number of disadvantages. Listed below are five examples.

Data redundancy

Often, within an organization, files and applications are created by different programmers from various departments over long periods of time. This can lead to *data redundancy*, a situation that occurs in a database when a field needs to be updated in more than one table. This practice can lead to several problems such as:

- Inconsistency in data format
- The same information being kept in several different places (files)
- *Data inconsistency*, a situation where various copies of the same data are conflicting, wastes storage space and duplicates effort

Data isolation

Data isolation is a property that determines when and how changes made by one operation become visible to other concurrent users and systems. This issue occurs in a concurrency situation. This is a problem because:

- It is difficult for new applications to retrieve the appropriate data, which might be stored in various files.

Integrity problems

Problems with *data integrity* is another disadvantage of using a file-based system. It refers to the maintenance and assurance that the data in a database are correct and consistent. Factors to consider when addressing this issue are:

- Data values must satisfy certain consistency constraints that are specified in the application programs.
- It is difficult to make changes to the application programs in order to enforce new constraints.

Security problems

Security can be a problem with a file-based approach because:

- There are constraints regarding accessing privileges.
- Application requirements are added to the system in an ad-hoc manner so it is difficult to enforce constraints.

Concurrency access

Concurrency is the ability of the database to allow multiple users access to the same record without adversely affecting transaction processing. A file-based system must manage, or prevent, concurrency by the application programs. Typically, in a file-based system, when an application opens a file, that file is locked. This means that no one else has access to the file at the same time.

In database systems, concurrency is managed thus allowing multiple users access to the same record. This is an important difference between database and file-based systems.

Database Approach

The difficulties that arise from using the file-based system have prompted the development of a new approach in managing large amounts of organizational information called the *database approach*.

Databases and database technology play an important role in most areas where computers are used, including business, education and medicine. To understand the fundamentals of database systems, we will start by introducing some basic concepts in this area.

Role of databases in business

Everybody uses a database in some way, even if it is just to store information about their friends and family. That data might be written down or stored in a computer by using a word-processing program or it could be saved in a spreadsheet. However, the best way to store data is by using *database management software*. This is a powerful software tool that allows you to store, manipulate and retrieve data in a variety of different ways.

Most companies keep track of customer information by storing it in a database. This data may include customers, employees, products, orders or anything else that assists the business with its operations.

The meaning of data

Data are factual information such as measurements or statistics about objects and concepts. We use data for discussions or as part of a calculation. Data can be a person, a place, an event, an action or any one of a number of things. A single fact is an element of data, or a *data element*.

If data are information and information is what we are in the business of working with, you can start to see where you might be storing it. Data can be stored in:

- Filing cabinets
- Spreadsheets
- Folders
- Ledgers
- Lists
- Piles of papers on your desk

All of these items store information, and so too does a database. Because of the mechanical nature of databases, they have terrific power to manage and process the information they hold. This can make the information they house much more useful for your work.

With this understanding of data, we can start to see how a tool with the capacity to store a collection of data and organize it, conduct a rapid search, retrieve and process, might make a difference to how we can use data. This book and the chapters that follow are all about managing information.

Key Terms

concurrency: the ability of the database to allow multiple users access to the same record without adversely affecting transaction processing

data element: a single fact or piece of information

data inconsistency: a situation where various copies of the same data are conflicting

data isolation: a property that determines when and how changes made by one operation become visible to other concurrent users and systems

data integrity: refers to the maintenance and assurance that the data in a database are correct and consistent

data redundancy: a situation that occurs in a database when a field needs to be updated in more than one table

database approach: allows the management of large amounts of organizational information

file-based system: an application program designed to manipulate data files

Exercises

1. Discuss each of the following terms:
    1. field
    2. record
    3. file
2. Discuss the disadvantages of file-based systems.

7.3: The Relational Data Model and others

The Relational Data Model

The relational data model was introduced by C. F. Codd in 1970. Currently, it is the most widely used data model.

The relational model has provided the basis for:

- Research on the theory of data/relationship/constraint
- Numerous database design methodologies
- The standard database access language called s*tructured query language (SQL)*
- Almost all modern commercial database management systems

The relational data model describes the world as “a collection of inter-related relations (or tables).”

Fundamental Concepts in the Relational Data Model

Relation

A *relation*, also known as a *table* or *file*, is a subset of the Cartesian product of a list of domains characterized by a name. And within a table, each row represents a group of related data values. A *row*, or record, is also known as a *tuple*. The columns in a table is a field and is also referred to as an attribute. You can also think of it this way: an attribute is used to define the record and a record contains a set of attributes.

Table

A database is composed of multiple tables and each table holds the data. Figure 7.1 shows a database that contains three tables.

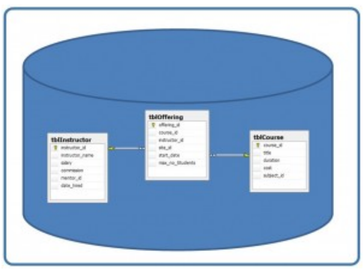


Figure 7.1. Database with three tables.

Column

A database stores pieces of information or facts in an organized way. Understanding how to use and get the most out of databases requires us to understand that method of organization.

The principal storage units are called *columns* or *fields* or *attributes*. These house the basic components of data into which your content can be broken down. When deciding which fields to create, you need to think generically about your information, for example, drawing out the common components of the information that you will store in the database and avoiding the specifics that distinguish one item from another.

Look at the example of an ID card in Figure 7.2 to see the relationship between fields and their data.

Field Name	Data
First Name	Isabelle
Family Name	Whelan
Nationality	British
Salary	109,900
Date of Birth	15 September 1983
Marital Status	Single
Shift	Mon, Wed
Place of issue	Addis Ababa
Valid until	17 December 2003

Figure 7.2. Example of an ID card by A. Watt.

Domain

A *domain* is the original sets of atomic values used to model data. By *atomic value*, we mean that each value in the domain is indivisible as far as the relational model is concerned. For example:

- The domain of Marital Status has a set of possibilities: Married, Single, Divorced.
- The domain of Shift has the set of all possible days: {Mon, Tue, Wed…}.
- The domain of Salary is the set of all floating-point numbers greater than 0 and less than 200,000.
- The domain of First Name is the set of character strings that represents names of people.

In summary, a domain is a set of acceptable values that a column is allowed to contain. This is based on various properties and the data type for the column. We will discuss data types in another chapter.

Records

Just as the content of any one document or item needs to be broken down into its constituent bits of data for storage in the fields, the link between them also needs to be available so that they can be reconstituted into their whole form. Records allow us to do this. *Records* contain fields that are related, such as a customer or an employee. As noted earlier, a tuple is another term used for record.

Records and fields form the basis of all databases. A simple table gives us the clearest picture of how records and fields work together in a database storage project.

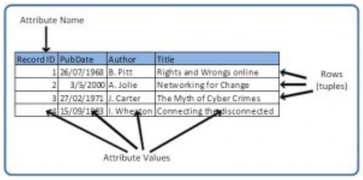


Figure 7.3. Example of a simple table by A. Watt.

The simple table example in Figure 7.3 shows us how fields can hold a range of different sorts of data. This one has:

- A Record ID field: this is an ordinal number; its data type is an integer.
- A PubDate field: this is displayed as day/month/year; its data type is date.
- An Author field: this is displayed as Initial. Surname; its data type is text.
- A Title field text: free text can be entered here.

You can command the database to sift through its data and organize it in a particular way. For example, you can request that a selection of records be limited by date: 1. all before a given date, 2. all after a given date or 3. all between two given dates. Similarly, you can choose to have records sorted by date. Because the field, or record, containing the data is set up as a Date field, the database reads the information in the Date field not just as numbers separated by slashes, but rather, as dates that must be ordered according to a calendar system.

Degree

The *degree* is the number of attributes in a table. In our example in Figure 7.3, the degree is 4.

Properties of a Table

- A table has a name that is distinct from all other tables in the database.
- There are no duplicate rows; each row is distinct.
- Entries in columns are atomic. The table does not contain repeating groups or multivalued attributes.
- Entries from columns are from the same domain based on their data type including:
  - number (numeric, integer, float, smallint,…)
  - character (string)
  - date
  - logical (true or false)
- Operations combining different data types are disallowed.
- Each attribute has a distinct name.
- The sequence of columns is insignificant.

- The sequence of rows is insignificant.

Data Warehouse

As organizations have begun to utilize databases as the centerpiece of their operations, the need to fully understand and leverage the data they are collecting has become more and more apparent. However, directly analyzing the data that is needed for day-to-day operations is not a good idea; we do not want to tax the operations of the company more than we need to. Further, organizations also want to analyze data in a historical sense: How does the data we have today compare with the same set of data this time last month, or last year? From these needs arose the concept of the data warehouse.

The concept of the data warehouse is simple: extract data from one or more of the organization's databases and load it into the data warehouse (which is itself another database) for storage and analysis. However, the execution of this concept is not that simple. A data warehouse should be designed so that it meets the following criteria:

- It uses non-operational data. This means that the data warehouse is using a copy of data from the active databases that the company uses in its day-to-day operations, so the data warehouse must pull data from the existing databases on a regular, scheduled basis.
- The data is time-variant. This means that whenever data is loaded into the data warehouse, it receives a time stamp, which allows for comparisons between different time periods.
- The data is standardized. Because the data in a data warehouse usually comes from several different sources, it is possible that the data does not use the same definitions or units. For example, our Events table in our Student Clubs database lists the event dates using the mm/dd/yyyy format (e.g., 01/10/2013). A table in another database might use the format yy/mm/dd (e.g., 13/01/10) for dates. In order for the data warehouse to match up dates, a standard date format would have to be agreed upon and all data loaded into the data warehouse would have to be converted to use this standard format. This process is called extraction-transformation-load (ETL).

There are two primary schools of thought when designing a data warehouse: bottom-up and top-down. The bottom-up approach starts by creating small data warehouses, called data marts, to solve specific business problems. As these data marts are created, they can be combined into a larger data warehouse. The top-down approach suggests that we should start by creating an enterprise-wide data warehouse and then, as specific business needs are identified, create smaller data marts from the data warehouse.

Data warehouse process (top-down)

Benefits of Data Warehouses

Organizations find data warehouses quite beneficial for a number of reasons:

- The process of developing a data warehouse forces an organization to better understand the data that it is currently collecting and, equally important, what data is not being collected.
- A data warehouse provides a centralized view of all data being collected across the enterprise and provides a means for determining data that is inconsistent.
- Once all data is identified as consistent, an organization can generate one version of the truth. This is important when the company wants to report consistent statistics about itself, such as revenue or number of employees.
- By having a data warehouse, snapshots of data can be taken over time. This creates a historical record of data, which allows for an analysis of trends.
- A data warehouse provides tools to combine data, which can provide new information and analysis.

Data Mining and Stuff

Data Mining

Data mining is the process of analyzing data to find previously unknown trends, patterns, and associations in order to make decisions. Generally, data mining is accomplished through automated means against extremely large data sets, such as a data warehouse. Some examples of data mining include:

- An analysis of sales from a large grocery chain might determine that milk is purchased more frequently the day after it rains in cities with a population of less than 50,000.

- A bank may find that loan applicants whose bank accounts show particular deposit and withdrawal patterns are not good credit risks.
- A baseball team may find that collegiate baseball players with specific statistics in hitting, pitching, and fielding make for more successful major league players.

In some cases, a data-mining project is begun with a hypothetical result in mind. For example, a grocery chain may already have some idea that buying patterns change after it rains and want to get a deeper understanding of exactly what is happening. In other cases, there are no presuppositions and a data-mining program is run against large data sets in order to find patterns and associations.

Privacy Concerns

The increasing power of data mining has caused concerns for many, especially in the area of privacy. In today's digital world, it is becoming easier than ever to take data from disparate sources and combine them to do new forms of analysis. In fact, a whole industry has sprung up around this technology: data brokers. These firms combine publicly accessible data with information obtained from the government and other sources to create vast warehouses of data about people and companies that they can then sell. This subject will be covered in much more detail in chapter 12 – the chapter on the ethical concerns of information systems.

Business Intelligence and Business Analytics

With tools such as data warehousing and data mining at their disposal, businesses are learning how to use information to their advantage. The term *business intelligence* is used to describe the process that organizations use to take data they are collecting and analyze it in the hopes of obtaining a competitive advantage. Besides using data from their internal databases, firms often purchase information from data brokers to get a big-picture understanding of their industries. *Business analytics* is the term used to describe the use of internal company data to improve business processes and practices.

Knowledge Management

All companies accumulate knowledge over the course of their existence. Some of this knowledge is written down or saved, but not in an organized fashion. Much of this knowledge is not written down; instead, it is stored inside the heads of its employees. Knowledge management is the process of formalizing the capture, indexing, and storing of the company's knowledge in order to benefit from the experiences and insights that the company has captured during its existence.

Key Terms

column: see *attribute*

degree: number of attributes in a table

domain: the original sets of atomic values used to model data; a set of acceptable values that a column is allowed to contain

field: see *attribute*

file: see *relation*

record: contains fields that are related; *see tuple*

relation: a subset of the Cartesian product of a list of domains characterized by a name; the technical term for table or file

row: *see tuple*

structured query language (SQL): the standard database access language

table: see *relation*

tuple: a technical term for row or record

Terminology Key

Several of the terms used in this chapter are synonymous. In addition to the Key Terms above, please refer to Table 8.1 below. The terms in the Alternative 1 column are most commonly used.

Formal Terms (Codd)	Alternative 1	Alternative 2
Relation	Table	File
Tuple	Row	Record
Attribute	Column	Field

Table 8.1. Terms and their synonyms by A. Watt.

7.4: Databases and Security Issues

Why Data Security is Important

It is the confidentiality, integrity, and availability (CIA) of the data in a database that need to be protected. Confidentiality can be lost if an unauthorized person gains entry or access to a database, or if a person who is authorized to view selected records in a database accesses other records he or she should not be able to view. If the data is altered by someone who is unauthorized to do so, the result is a loss of data integrity. And if those who need to have access to the database and its services are blocked from doing so, there is a resulting loss of availability. Security of any database is significantly impacted by any one or more of these basic components of CIA being violated (Nuramn, 2011).

There are various reasons for spending money, time, and effort on data protection. The main reason is reducing financial loss, followed by compliance with regulatory requirements, maintaining high levels of productivity, and meeting customer expectations (Petrocelli, 2005).

Both businesses and home computer users should be concerned about data security. The information stored in databases—client information, payment information, personal files, bank account details, and more—can be hard to replace, whether the loss results from

- physical threats such as a fire or a significant power outage
- human error that results in errors in the processing of information or unintended deletion of data, or from erroneous input
- corporate espionage, theft, or malicious activity.

Loss of this data is potentially dangerous if it falls into the wrong hands (*Why Data Security*, n.d.).

It is in these three areas that a risk assessment of the database's security and protection of the data should focus. Is there a backup procedure that would allow access to the data if the primary database is destroyed by a physical threat? That same backup procedure might be important in case the CIA of the database is inadvertently affected by human error. And what safeguards can/should be put in place to prevent incidents of espionage, theft, or other malicious activity? We will look again at risk assessments later on this page.

How Common Are Database Breaches?

Just how prevalent are the threats against databases? Is it worth the time, money, and personnel effort to ensure that the database is safeguarded? Remember the Target and Neiman Marcus problems that surfaced in late 2013? And the continuing saga of Edward Snowden and the NSA leaks? These may have been the most widely publicized data breaches of 2013. But they were definitely just two of many such database breaches. ***Database breaches*** *are the exposure of database records containing personally identifiable information (PII) or other sensitive information to unauthorized viewers.* Risk-Based Security (RBS), a group of consultants and founders of the Open Security Foundation (OSF), report that 2013 saw a record number of data records exposed via data breaches. Over 822 million such records were made available to persons who had no authority to view these records (*Risk Based Security*, 2014). But remember, the number of reported database breaches does not reflect the total number of breaches that occurred. Some companies do not report breaches in order to protect their reputations or to prevent customers from abandoning the company. The following is a shortlist of what RBS discovered.

- The business sector accounted for 53.4% of reported incidents, followed by government (19.3%), medical (11.5%), education (8.2%), and unknown (7.6%).
- Hacking was the cause of 59.8% of reported incidents, accounting for 72.0% of exposed records.
- Of the reported incidents, 4.8% were the result of web-related attacks, which amounted to 16.9% of exposed records.
- Four incidents ***in 2013 alone*** secured a place on the Top 10 All-Time Breaches list:
  - Adobe—152 million records. Customer IDs, encrypted passwords, debit or credit card numbers, and other information relating to customer orders was compromised.
  - Unknown organizations—140 million records. North Korean hackers exposed e-mail addresses and identification numbers of South Korean individuals.
  - Target—110 million records. The information included customer names, addresses, phone numbers, e-mail addresses, credit/debit card numbers, PINs, and security codes.

- Pinterest—70 million records. A flaw in the site's application programming interface (API) exposed users' e-mail addresses.

Even if you were not impacted by any of the above data breaches, if you have used a credit card, made an airline reservation, subscribed to a magazine, been a patient in a hospital, or shopped at a chain store (supermarket or department store), or if you are a member of an online social media site, your personally identifiable information (PII) is stored in a database. How vulnerable is your PII?

What Are the Most Common Causes of Database Breaches?

As evidenced by the NSA Snowden leaks and the Target breach, no database, and no government agency, company, or business is as secure as the owners of that database think. It is difficult for database administrators and security managers to keep pace with the new threats and vulnerabilities that continually emerge. And to compound the issues, every company/business/government has different security issues, making it a particularly hard challenge to standardize any one solution that fits all. However, there are some common threats and vulnerabilities that seem to occur repeatedly.

Threats

Unauthorized Access by Insiders

The malicious insider with approved access to the system is one of the greatest threats to database security.

> People attack computers because that's where the information is, and in our hyper-competitive, hi-tech business and international environment, information increasingly has great value. Some alienated individuals also gain a sense of power, control, and self-importance through successful penetration of computer systems to steal or destroy the information or disrupt an organization's activities. (*Threats to Computer Systems*, n.d.)

Another scenario might involve employees affected by a workforce reduction who take customer account lists, financial data, or strategic plans with them when they leave. Proprietary information could end up in the hands of competitors or be widely disseminated online (*Data Loss Prevention*, n.d.).

Insiders may also be a threat to database security if they are granted database access privileges that go beyond the requirements of their job function, abuse legitimate database privileges for unauthorized purposes, or convert access privileges from those of an ordinary user to those of an administrator.

Accidental Breaches Resulting from Incorrect—but Not Malicious—Usage

The data breach is not always the result of a deliberate attempt to subvert data security; sometimes it is an unintended consequence. For example, employees might export data from the parent database system at work and send it, typically unencrypted, to personal e-mail addresses so they can work from home. The data then might be subsequently compromised on someone's home computer. Or a data mining application might contain flaws that allow a user without the correct access credentials to stumble upon database records inadvertently. (Note: If the user deliberately continues to access the data without permission, this situation becomes a malicious insider threat.)

Unprotected Personal Hardware Collection

It is becoming increasingly common for data to be transferred to other personal mobile devices—USB flash drives, smartphones, tablets, and the like. It is rare now to find an employee who never uses a mobile device—personal or company-supplied—for business purposes. However, mobile devices continue to be a significant source of data breaches, stemming from a range of circumstances, including loss or theft of the devices, failure to install antimalware tools on the devices, or failing to password-protect a device being used for business purposes. Data is at risk if an employee stores any proprietary information on such a device or if that device is used to access a company's network and/or database (Bruemmer, 2014).

Stolen Laptops

Forgetful or careless laptop owners whose equipment is taken expose data on that laptop to persons not authorized to have access to the data. This can also happen if a laptop is replaced and the hard drive on the original machine is not properly erased or destroyed.

Weak Authentication

A legitimate database user typically is required to submit an ID and password in order to gain access to a protected database. Authentication is the process (internal to the database program itself) by which the credentials of the user are verified and access may be granted. If the process of authentication is weak, an attacker can assume the identity of a legitimate user by stealing or obtaining login credentials. Credentials may be illegitimately obtained by various means:

- Credential theft. The attacker accesses password files or finds a paper on which the legitimate user has written down the ID and password.
- Social engineering. The attacker deceives someone into providing the login ID and password by posing as a supervisor, IT maintenance personnel, or other authority.
- Brute-force attacks. Have you ever been locked out of an account after attempting to log in more than three times with an incorrect password? If so, this is the simplest (and perhaps least effective) means of blocking a brute force attack, whether it is an attempt to access files on your machine or to access a database. However, not all password-protected systems, databases, or files block you from access after three attempts. For example, if you have put a lock on a file on your computer, you most likely have not set a limit on the number of attempts on that file. A brute-force attack is a password-guessing approach in which the attacker attempts to discover a password by systematically testing every combination of letters, numbers, and symbols until the correct combination is found. Depending upon the password's length and complexity, this can be a very difficult task to complete. However, there are widely available tools that hackers can use to find the password, and it can be difficult to block all the means by which hacker will try to find the password (*Blocking Brute Force Attacks*, 2007).

Exploiting Weaknesses in an Operating System or Network

Worms, viruses, or Trojan horses could be introduced into an unprotected or poorly protected operating system or computer network that supports the database, leading to potential unauthorized database access (loss of confidentiality), data corruption (loss of integrity), or denial of service (DOS), a loss of access to legitimate users. A DOS may be achieved by causing a server to stop functioning, or “crash,” flooding a network with message traffic or overloading resources on the computer, forcing it to stop handling additional tasks or processing.

Theft of Database Backup Tapes or Hard Drives

Database backups typically do not have the same security measures in place that the primary database employs. These backups may not be encrypted, and the media on which backups are stored are also unprotected. Theft of the backup media may allow the attacker full access to the data stored within the backup (Schulman, 2007).

Vulnerabilities

There are other means by which databases are exposed to security breaches, and these are considered vulnerabilities that may subject a database to a security breach. These are more passive, but they can do as much harm as direct threats:

- *Data at rest* (unencrypted information) that is passively residing in storage within the boundaries of company computers, perhaps waiting to be moved to a secure database. Data at rest typically is not as well protected as data that has been entered into the database and enjoys the database security measures.
- *Data in motion* is information that is being electronically transmitted outside the company’s protected network via e-mail or other communication mediums. For example, the data might be transferred to a backup facility that is not part of the internal storage media used for daily work. Or if the company uses the cloud for data storage backups, the transfer might take place outside of the company’s protected network. This can lead to a loss of sensitive data if there is a malicious attack via malware during the transfer process or during the execution of a flawed business process that allows unauthorized persons to view or obtain the data. (This is not the same as the accidental breach resulting from incorrect but not malicious usage noted above, where the home computer to which the data has been transferred is attacked or breached. That accidental breach occurred without any intention of harm by the employee.)
- *Poor architecture*, in which security was not adequately factored into the design and development of the database structure. This vulnerability may not be discovered until there is an attempted or successful data breach.
- *Vendor bugs*, particularly programming flaws that allow actions to take place within the database and with the data that were not intended or planned. Much like poor application architecture, this vulnerability may not be uncovered until there is an attempted or successful data breach.
- *An unlocked database* is one that has no security measures in place to control access or auditing. This seems counterintuitive, but many home users employing a database for personal needs, or even for working on company data while at home, maybe working with an unlocked database (Nichols, n.d.;*Data Loss Prevention*, n.d.)

Risk Assessments

In the business environment, it is critical that a thorough risk assessment takes place and be periodically reviewed. The assessment should address:

- who has access to what data
- the circumstances under which access to the database may need to change
- who maintains the passwords needed to access the database
- who uses the company's computers for access to the internet, e-mail programs, etc., and how employees access those resources
- what type of firewalls and anti-malware solutions to put in place
- the training of the staff
- who has responsibility for enforcement procedures related to data security (*Why Data Security*, n.d.)

There are identified solutions for each of the threats and vulnerabilities discussed here, including well-defined and enforced access policies, use of strong data encryption, vulnerability assessments, policies related to strong passwords, and installation of firewalls. There are companies that specialize in designing plans, procedures, and software to prevent data loss or data leakage. With ***data loss****, the data is lost forever, either by deletion, theft, or data corruption.* ***Data leakage*** *allows unauthorized people to get access to the data, either by intentional action or by mistake.* So data loss and data leakage can be intentional or unintentional, and both can be malicious or just human errors (VJ, 2013).

How Can You Protect Your PII?

Protecting databases and the data contained within can be a costly and all-consuming activity. But what does this mean for you, the individual who uses that credit card, makes airline reservations, files taxes online, subscribes to a magazine, has been a patient in a hospital, shops at a chain store, or is a member of an online social media site? Your PII is out there, stored in multiple databases. Obviously, you cannot implement security measures for the company, business, or government agency that holds your PII. But are there many measures you can take to better protect yourself? Here are a few rules of thumb that you can implement:

Keep your passwords to yourself.	Do not leave a slip with a list of passwords under your computer, or anywhere where it can be viewed or taken by someone. Just giving your password to a friend is not a good idea, either.
Use different passwords for different accounts.	Remembering multiple passwords can be a challenge, and it’s often convenient to use the same password for multiple accounts, ranging from Facebook and your bank account to your Twitter page. The danger here is that a compromise of any one of these accounts could also result in the compromise of others if the same password is used for multiple accounts.
Use strong passwords.	Many of your user IDs must have strong passwords to gain entry into one or more systems. In those instances when you can choose any password configuration, pick a strong password to protect your information.
Check your credit reports annually.	Sometimes people don’t learn that they’re victims of identity theft until their credit rating and identity are destroyed. It’s proactive to get copies of your credit reports from the credit bureaus and carefully review them for any errors. Be sure to follow-up with the credit bureaus to make any corrections to your reports, if needed. By law, you can get one free credit report from each of the three credit bureaus every year.
Google yourself.	Enter your own name in Google, Yahoo or other search engine and see what data comes up. Investigate any postings about yourself in the information that you find. Look for any suggestions that your PII may be compromised.
Remember that people can be a very weak link in security.	No matter how secure you make your passwords and how careful you are with your technology, there is always a human element to protecting your information.
Control physical access to your devices.	It’s important not leave laptops and other mobile devices unattended in public locations, like a coffee shop or other location with free WiFi. An unattended machine is at risk, for both theft and other security threats. When you aren't controlling physical access to your machine, you shouldn’t let it out of your sight.

Remember to logout of a website when you are finished using it.	Whether it's your email, bank account, retail store shopping account or library account, always remember to logout when you leave the website.
Remember to lock your computer with a password when you are finished using it.	By requiring a password to access your computer (or other electronic device) you are protecting your information. You are also making your computer useless to a thief who cannot break password locks.

References

1. Blocking Brute Force Attacks. (2007). ***Retrieved March 17, 2014, from System Administration Database UVA Computer Science: http://www.cs.virginia.edu/~csadmin/...rute_force.php.***
2. Bruemmer, M. (2014, January 21). ***How Mobile Devices Can Imperil Your Organization's Cyber Security.*** Retrieved March 17, 2014, from Experian Information Solutions.com: http://www.experian.com/blogs/data-b...yber-security/.
3. Dartnell, J. (2014, April 20). ***EMC: Digital universe data to grow tenfold by 2020.*** Retrieved from cnme: computer news middle east: http://www.cnmeonline.com/news/emc-d...nfold-by-2020/.
4. Data Loss Prevention: Keeping sensitive data out of the wrong hands. (n.d.). ***Retrieved March 17, 2014, from Price Waterhouse Cooper Advisory Services/security: http://www.pwc.com/en_US/us/increasi...prevention.pdf.***
5. Nichols, E. (n.d.). ***Eleven specific solutions to today's most common database security threats and vulnerabilities.***Retrieved March 16, 2014, from University of Oregon: http://aimdegree.com/research/ebrief...eb-nichols.php.
6. Nuramn, A. (2011, August 10). ***Database Security.*** (L. Stonecypher, Editor) Retrieved April 29, 2014, from Bright Hub: http://www.brighthub.com/computing/s...les/61400.aspx.
7. Risk Based Security. (2014, February 18). ***Data Breach QuickView.*** Retrieved from Risk Based Security: https://www.riskbasedsecurity.com/re...hQuickView.pdf.
8. Threats to Computer Systems. (n.d.). ***Retrieved March 16, 2014, from USDA.gov: http://www.dm.usda.gov/ocpm/Security...ut/Threats.htm.***
9. Vishen, N. (2013, April 20). ***Largest Databases of the World.*** Retrieved from Neeraj0dba.blogspot: http://neeraj-dba.blogspot.com/2013/...-of-world.html.
10. VJ (2013, April 2). ***Is There a Difference Between Data LOSS and Data LEAKAGE Prevention?.*** Retrieved March 28, 2014, from Rational Survivability: http://www.rationalsurvivability.com...ge-prevention/.
11. Why Data Security is of Paramount Importance. (n.d.). ***Retrieved March 21, 2014, from SpamLaws.com: http://www.spamlaws.com/data-security-importance.html.***
12. Petrocelli, T. (2005). "The Changing Face of Data Protection." InformIT. Retrieved May 1, 2019, from http://www.informit.com/articles/art...22303&seqNum=3.

7.5: Fundamental Database Concepts

What Is a Database?

A *database* is a shared collection of related data used to support the activities of a particular organization. A database can be viewed as a repository of data that is defined once and then accessed by various users as shown in Figure 2.1.

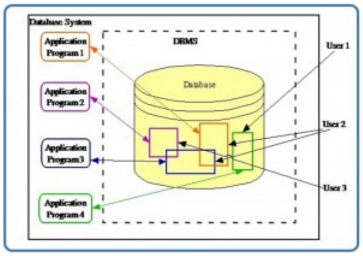


Figure 2.1. A database is a repository of data.

Database Properties

A database has the following properties:

- It is a representation of some aspect of the real world or a collection of *data element*s (facts) representing real-world information.
- A database is logical, coherent and internally consistent.
- A database is designed, built and populated with data for a specific purpose.
- Each data item is stored in a field.
- A combination of fields makes up a *table*. For example, each field in an employee table contains data about an individual employee.

A database can contain many tables. For example, a membership system may contain an address table and an individual member table as shown in Figure 2.2. Members of Science World are individuals, group homes, businesses and corporations who have an active membership to Science World. Memberships can be purchased for a one- or two-year period, and then renewed for another one- or two-year period.

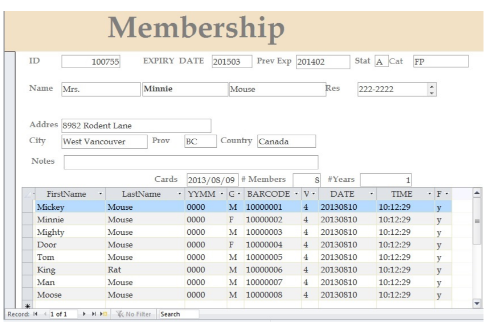


Figure 2.2. Membership system at Science World by N. Eng.

In Figure 2.2, Minnie Mouse renewed the family membership with Science World. Everyone with membership ID#100755 lives at 8932 Rodent Lane. The individual members are Mickey Mouse, Minnie Mouse, Mighty Mouse, Door Mouse, Tom Mouse, King Rat, Man Mouse and Moose Mouse.

Database Management System

A *database management system (DBMS)* is a collection of programs that enables users to create and maintain databases and control all access to them. The primary goal of a DBMS is to provide an environment that is both convenient and efficient for users to retrieve and store information.

With the database approach, we can have the traditional banking system as shown in Figure 2.3. In this bank example, a DBMS is used by the Personnel Department, the Account Department and the Loan Department to access the shared corporate database.

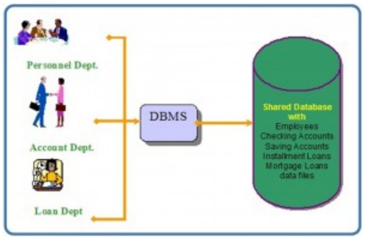


Figure 2.3. A bank database management system (DBMS).

Characteristics and Benefits of a Database

Managing information means taking care of it so that it works for us and is useful for the tasks we perform. By using a DBMS, the information we collect and add to its database is no longer subject to accidental disorganization. It becomes more accessible and integrated with the rest of our work. Managing information using a database allows us to become strategic users of the data we have.

We often need to access and re-sort data for various uses. These may include:

- Creating mailing lists
- Writing management reports
- Generating lists of selected news stories
- Identifying various client needs

The processing power of a database allows it to manipulate the data it houses, so it can:

- Sort
- Match
- Link
- Aggregate
- Skip fields
- Calculate
- Arrange

Because of the versatility of databases, we find them powering all sorts of projects. A database can be linked to:

- A website that is capturing registered users
- A client-tracking application for social service organizations
- A medical record system for a health care facility
- Your personal address book in your email client
- A collection of word-processed documents
- A system that issues airline reservations

Characteristics and Benefits of a Database

There are a number of characteristics that distinguish the database approach from the file-based system or approach. This chapter describes the benefits (and features) of the database system.

Self-describing nature of a database system

A database system is referred to as *self-describing* because it not only contains the database itself, but also *metadata* which defines and describes the data and relationships between tables in the database. This information is used by the DBMS software or database users if needed. This separation of data and information about the data makes a database system totally different from the traditional file-based system in which the data definition is part of the application programs.

Insulation between program and data

In the file-based system, the structure of the data files is defined in the application programs so if a user wants to change the structure of a file, all the programs that access that file might need to be changed as well.

On the other hand, in the database approach, the data structure is stored in the system catalogue and not in the programs. Therefore, one change is all that is needed to change the structure of a file. This insulation between the programs and data is also called program-data independence.

Support for multiple views of data

A database supports multiple views of data. A *view* is a subset of the database, which is defined and dedicated for particular users of the system. Multiple users in the system might have different views of the system. Each view might contain only the data of interest to a user or group of users.

Sharing of data and multi-user system

Current database systems are designed for multiple users. That is, they allow many users to access the same database at the same time. This access is achieved through features called *concurrency control strategies*. These strategies ensure that the data accessed are always correct and that data integrity is maintained.

The design of modern multiuser database systems is a great improvement from those in the past which restricted usage to one person at a time.

Control of data redundancy

In the database approach, ideally, each data item is stored in only one place in the database. In some cases, data redundancy still exists to improve system performance, but such redundancy is controlled by application programming and kept to minimum by introducing as little redudancy as possible when designing the database.

Data sharing

The integration of all the data, for an organization, within a database system has many advantages. First, it allows for data sharing among employees and others who have access to the system. Second, it gives users the ability to generate more information from a given amount of data than would be possible without the integration.

Enforcement of integrity constraints

Database management systems must provide the ability to define and enforce certain constraints to ensure that users enter valid information and maintain data integrity. A *database constraint* is a restriction or rule that dictates what can be entered or edited in a table such as a postal code using a certain format or adding a valid city in the City field.

There are many types of database constraints. *Data type*, for example, determines the sort of data permitted in a field, for example numbers only. *Data uniqueness* such as the primary key ensures that no duplicates are entered. Constraints can be simple (field based) or complex (programming).

Restriction of unauthorized access

Not all users of a database system will have the same accessing privileges. For example, one user might have *read-only access* (i.e., the ability to read a file but not make changes), while another might have *read and write privileges*, which is the ability to both read and modify a file. For this reason, a database management system should provide a security subsystem to create and control different types of user accounts and restrict unauthorized access.

Data independence

Another advantage of a database management system is how it allows for data independence. In other words, the system data descriptions or data describing data (metadata) are separated from the application programs. This is possible because changes to the data structure are handled by the database management system and are not embedded in the program itself.

Transaction processing

A database management system must include concurrency control subsystems. This feature ensures that data remains consistent and valid during transaction processing even if several users update the same information.

Provision for multiple views of data

By its very nature, a DBMS permits many users to have access to its database either individually or simultaneously. It is not important for users to be aware of how and where the data they access is stored

Backup and recovery facilities

Backup and recovery are methods that allow you to protect your data from loss. The database system provides a separate process, from that of a network backup, for backing up and recovering data. If a hard drive fails and the database stored on the hard drive is not accessible, the only way to recover the database is from a backup.
If a computer system fails in the middle of a complex update process, the recovery subsystem is responsible for making sure that the database is restored to its original state. These are two more benefits of a database management system.

Classification of Database Management Systems

Database management systems can be classified based on several criteria, such as the data model, user numbers and database distribution, all described below.

Classification Based on Data Model

The most popular data model in use today is the relational data model. Well-known DBMSs like Oracle, MS SQL Server, DB2 and MySQL support this model. Other traditional models, such as hierarchical data models and network data models, are still used in industry mainly on mainframe platforms. However, they are not commonly used due to their complexity. These are all referred to as *traditional models* because they preceded the relational model.

In recent years, the newer *object-oriented data models* were introduced. This model is a database management system in which information is represented in the form of objects as used in object-oriented programming. Object-oriented databases are different from relational databases, which are table-oriented. Object-oriented database management systems (OODBMS) combine database capabilities with object-oriented programming language capabilities.

The object-oriented models have not caught on as expected so are not in widespread use. Some examples of object-oriented DBMSs are O2, ObjectStore and Jasmine.

Classification Based on User Numbers

A DBMS can be classification based on the number of users it supports. It can be a *single-user database system*, which supports one user at a time, or a *multiuser database system*, which supports multiple users concurrently.

Classification Based on Database Distribution

There are four main distribution systems for database systems and these, in turn, can be used to classify the DBMS.

Centralized systems

With a *centralized database system*, the DBMS and database are stored at a single site that is used by several other systems too. This is illustrated in Figure 6.1.

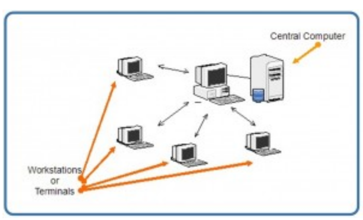


Figure 6.1. Example of a centralized database system.

In the early 1980s, many Canadian libraries used the GEAC 8000 to convert their manual card catalogues to machine-readable centralized catalogue systems. Each book catalogue had a barcode field similar to those on supermarket products.

Distributed database system

In a *distributed database system*, the actual database and the DBMS software are distributed from various sites that are connected by a computer network, as shown in Figure 6.2.

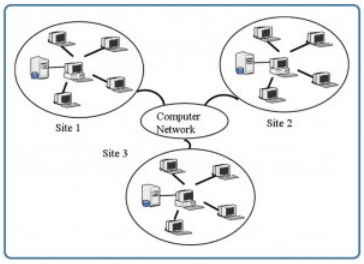


Figure 6.2. Example of a distributed database system.

Homogeneous distributed database systems

Homogeneous distributed database systems use the same DBMS software from multiple sites. Data exchange between these various sites can be handled easily. For example, library information systems by the same vendor, such as Geac Computer Corporation, use the same DBMS software which allows easy data exchange between the various Geac library sites.

Heterogeneous distributed database systems

In a *heterogeneous distributed database system*, different sites might use different DBMS software, but there is additional common software to support data exchange between these sites. For example, the various library database systems use the same machine-readable cataloguing (MARC) format to support library record data exchange.

Database Users

End Users

End users are the people whose jobs require access to a database for querying, updating and generating reports.

Application user

The application user is someone who accesses an existing application program to perform daily tasks.

Sophisticated user

Sophisticated users are those who have their own way of accessing the database. This means they do not use the application program provided in the system. Instead, they might define their own application or describe their need directly by using query languages. These specialized users maintain their personal databases by using ready-made program packages that provide easy-to-use menu driven commands, such as MS Access.

Application Programmers

These users implement specific application programs to access the stored data. They must be familiar with the DBMSs to accomplish their task.

Database Administrators (DBA)

This may be one person or a group of people in an organization responsible for authorizing access to the database, monitoring its use and managing all of the resources to support the use of the entire database system.

Key Terms

application programmer: user who implements specific application programs to access the stored data

application user: accesses an existing application program to perform daily tasks.

centralized database system: the DBMS and database are stored at a single site that is used by several other systems too

data elements: facts that represent real-world information

database: a shared collection of related data used to support the activities of a particular organization

database administrator (DBA): responsible for authorizing access to the database, monitoring its use, and managing all the resources to support the use of the entire database system

database management system (DBMS): a collection of programs that enables users to create and maintain databases and control all access to them

datatype: determines the sort of data permitted in a field, for example, numbers only

distributed database system: the actual database and the DBMS software are distributed from various sites that are connected by a computer network

end user: people whose jobs require access to a database for querying, updating, and generating reports

heterogeneous distributed database system: different sites might use different DBMS software, but there is additional common software to support data exchange between these sites

homogeneous distributed database systems: use the same DBMS software at multiple sites

metadata: defines and describes the data and relationships between tables in the database

multiuser database system: a database management system that supports multiple users concurrently

object-oriented data model: a database management system in which information is represented in the form of objects as used in object-oriented programming

single-user database system: a database management system that supports one user at a time

sophisticated user: those who use other methods, other than the application program, to access the database

Exercises

1. What is a database management system (DBMS)?
2. How is a DBMS distinguished from a file-based system?
3. What is metadata?
4. What are the properties of a DBMS?
5. Provide three examples of a real-world database (e.g., the library contains a database of books).
6. Provide three examples of the most popular relational databases used.
7. What is the difference between centralized and distributed database systems?
8. What is the difference between homogenous distributed database systems and heterogeneous distributed database systems?

7.6: Appendix A - Designing a Database

Designing a Database

Suppose a university wants to create an information system to track participation in student clubs. After interviewing several people, the design team learns that the goal of implementing the system is to give better insight into how the university funds clubs. This will be accomplished by tracking how many members each club has and how active the clubs are. From this, the team decides that the system must keep track of the clubs, their members, and their events. Using this information, the design team determines that the following tables need to be created:

- Clubs: this will track the club name, the club president, and a short description of the club.
- Students: student name, e-mail, and year of birth.
- Memberships: this table will correlate students with clubs, allowing us to have any given student join multiple clubs.
- Events: this table will track when the clubs meet and how many students showed up.

Now that the design team has determined which tables to create, they need to define the specific information that each table will hold. This requires identifying the fields that will be in each table. For example, Club Name would be one of the fields in the Clubs table. First Name and Last Name would be fields in the Students table. Finally, since this will be a relational database, every table should have a field in common with at least one other table (in other words: they should have a relationship with each other).

In order to properly create this relationship, a primary key must be selected for each table. This key is a unique identifier for each record in the table. For example, in the Students table, it might be possible to use students' last name as a way to uniquely identify them. However, it is more than likely that some students will share a last name (like Rodriguez, Smith, or Lee), so a different field should be selected. A student's e-mail address might be a good choice for a primary key, since e-mail addresses are unique. However, a primary key cannot change, so this would mean that if students changed their e-mail address we would have to remove them from the database and then re-insert them – not an attractive proposition. Our solution is to create a value for each student — a user ID — that will act as a primary key. We will also do this for each of the student clubs. This solution is quite common and is the reason you have so many user IDs!

You can see the final database design in the figure below:

Student Clubs database diagram

With this design, not only do we have a way to organize all of the information we need to meet the requirements, but we have also successfully related all the tables together. Here's what the database tables might look like with some sample data. Note that the Memberships table has the sole purpose of allowing us to relate multiple students to multiple clubs.

7.7: Summary

Summary

Data are the raw bits and pieces of information with no context.

Information is data with context and has usefulness.

Knowledge is information (Data in context) that is analyzed and aggregated to make decisions.

A database is *a digital collection of related information* to transform data into information in order to generate knowledge that can be used for decision making.

There are three main data types: structured, unstructured, and semi-structured.

Before database systems, computers relied on a much less elegant and costly approach to data management called the file-based system. A file-based system is a collection of application programs that perform services for the users wishing to access information.

A *database management system (DBMS)* is a collection of programs that enables users to create and maintain databases and control all access to them.

There are four main distribution systems for database systems: Centralized systems, Distributed database systems, Homogeneous distributed database systems, and Heterogeneous distributed database systems.

A data warehouse is a centralized view of all data being collected across the enterprise and provides a means for determining data that is inconsistent.

Data mining is the process of analyzing data to find previously unknown trends, patterns, and associations in order to make decisions

It is the confidentiality, integrity, and availability (CIA) of the data in a database that needs to be protected.

The causes of Database Breaches are Threats and Vulnerabilities.

Protect yourself better by Keep passwords to yourself, Use different passwords for different accounts, Use strong passwords, check your credit reports annually, Control physical access to your devices, Remember to log out or lock your computer when you are finished using it, and remember to lock your computer with a password when you are finished using it.

CHAPTER OVERVIEW

8: The People in Information Systems

Introduction

In the opening chapters of this text, we focused on the technology behind information systems: hardware, software, data, and networking. In the last chapter, we discussed business processes and the key role they can play in the success of a business. In this chapter, we will be discussing the last component of an information system: people.

People are involved in information systems in just about every way you can think of: people imagine information systems, people develop information systems, people support information systems, and, perhaps most importantly, people *use* information systems.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe each of the different roles that people play in the design, development, and use of information systems;
- understand the different career paths available to those who work with information systems;
- explain the importance of where the information-systems function is placed in an organization; and
- describe the different types of users of information systems.

8.1: Creators

The Creators of Information Systems

The first group of people we are going to look at play a role in designing, developing, and building information systems. These people are generally very technical and have a background in programming and mathematics. Just about everyone who works in the creation of information systems has a minimum of a bachelor's degree in computer science or information systems, though that is not necessarily a requirement.

Systems Analyst

The role of the systems analyst is to straddle the divide between identifying business needs and imagining a new or redesigned computer-based system to fulfill those needs. This individual will work with a person, team, or department with business requirements and identify the specific details of a system that needs to be built. Generally, this will require the analyst to have a good understanding of the business itself, the business processes involved, and the ability to document them well. The analyst will identify the different stakeholders in the system and work to involve the appropriate individuals in the process.

Once the requirements are determined, the analyst will begin the process of translating these requirements into an information-systems design. A good analyst will understand what different technological solutions will work and provide several different alternatives to the requester, based on the company's budgetary constraints, technology constraints, and culture. Once the solution is selected, the analyst will create a detailed document describing the new system. This new document will require that the analyst understand how to speak in the technical language of systems developers.

A systems analyst generally is not the one who does the actual development of the information system. The design document created by the systems analyst provides the detail needed to create the system and is handed off to a programmer (or team of programmers) to do the actual creation of the system. In some cases, however, a systems analyst may go ahead and create the system that he or she designed. This person is sometimes referred to as a programmer-analyst.

In other cases, the system may be assembled from off-the-shelf components by a person called a systems integrator. This is a specific type of systems analyst that understands how to get different software packages to work with each other.

To become a systems analyst, you should have a background both in the business and in systems design. Many analysts first worked as programmers and/or had to experience in the business before becoming systems analysts.

Programmer

Programmers spend their time writing computer code in a programming language. In the case of systems development, programmers generally attempt to fulfill the design specifications given to them by a systems analyst. Many different styles of programming exist: a programmer may work alone for long stretches of time or may work in a team with other programmers. A programmer needs to be able to understand complex processes and also the intricacies of one or more programming languages. Generally, a programmer is very proficient in mathematics, as mathematical concepts underlie most programming code.

Computer Engineer

Computer engineers design the computing devices that we use every day. There are many types of computer engineers, who work on a variety of different types of devices and systems. Some of the more prominent engineering jobs are as follows:

- Hardware engineer. A hardware engineer designs hardware components, such as microprocessors. Many times, a hardware engineer is at the cutting edge of computing technology, creating something brand new. Other times, the hardware engineer's job is to engineer an existing component to work faster or use less power. Many times, a hardware engineer's job is to write code to create a program that will be implemented directly on a computer chip.
- Software engineer. Software engineers do not actually design devices; instead, they create new programming languages and operating systems, working at the lowest levels of the hardware to develop new kinds of software to run on the hardware.
- Systems engineer. A systems engineer takes the components designed by other engineers and makes them all work together. For example, to build a computer, the motherboard, processor, memory, and hard disk all have to work together. A systems engineer has experience with many different types of hardware and software and knows how to integrate them to create new functionality.
- Network engineer. A network engineer's job is to understand the networking requirements of an organization and then design a communications system to meet those needs, using the networking hardware and software available.

There are many different types of computer engineers, and often the job descriptions overlap. While many may call themselves engineers based on a company job title, there is also a professional designation of “professional engineer,” which has specific requirements behind it. In the US, each state has its own set of requirements for the use of this title, as do different countries around the world. Most often, it involves a professional licensing exam.

8.2: Operations and Administration

Information-Systems Operations and Administration

Another group of information-systems professionals is involved in the day-to-day operations and administration of IT. These people must keep the systems running and up-to-date so that the rest of the organization can make the most effective use of these resources.

Computer Operator

A computer operator is a person who keeps large computers running. This person's job is to oversee the mainframe computers and data centers in organizations. Some of their duties include keeping the operating systems up to date, ensuring available memory and disk storage, and overseeing the physical environment of the computer. Since mainframe computers increasingly have been replaced with servers, storage management systems, and other platforms, computer operators' jobs have grown broader and include working with these specialized systems.

Database Administrator

A database administrator (DBA) is the person who manages the databases for an organization. This person creates and maintains databases that are used as part of applications or the data warehouse. The DBA also consults with systems analysts and programmers on projects that require access to or the creation of databases.

Help-Desk/Support Analyst

Most mid-size to large organizations have their own information-technology help desk. The help desk is the first line of support for computer users in the company. Computer users who are having problems or need information can contact the help desk for assistance. Many times, a help-desk worker is a junior-level employee who does not necessarily know how to answer all of the questions that come his or her way. In these cases, help-desk analysts work with senior-level support analysts or have a computer knowledge base at their disposal to help them investigate the problem at hand. The help desk is a great place to break into working in IT because it exposes you to all of the different technologies within the company. A successful help-desk analyst should have good people and communications skills, as well as at least junior-level IT skills.

Trainer

A computer trainer conducts classes to teach people specific computer skills. For example, if a new ERP system is being installed in an organization, one part of the implementation process is to teach all of the users how to use the new system. A trainer may work for a software company and be contracted to come in to conduct classes when needed; a trainer may work for a company that offers regular training sessions, or a trainer may be employed full time for an organization to handle all of their computer instruction needs. To be successful as a trainer, you need to be able to communicate technical concepts well and also have a lot of patience!

8.3: Managers

Managing Information Systems

The management of information-systems functions is critical to the success of information systems within the organization. Here are some of the jobs associated with the management of information systems.

CIO

The CIO, or chief information officer, is the head of the information-systems function. This person aligns the plans and operations of the information systems with the strategic goals of the organization. This includes tasks such as budgeting, strategic planning, and personnel decisions for the information-systems function. The CIO must also be the face of the IT department within the organization. This involves working with senior leaders in all parts of the organization to ensure good communication and planning.

Interestingly, the CIO position does not necessarily require a lot of technical expertise. While helpful, it is more important for this person to have good management skills and understand the business. Many organizations do not have someone with the title of CIO; instead, the head of the information-systems function is called vice president of information systems or director of information systems.

Functional Manager

As an information-systems organization becomes larger, many of the different functions are grouped together and led by a manager. These functional managers report to the CIO and manage the employees specific to their functions. For example, in a large organization, there is a group of systems analysts who report to a manager of the systems-analysis function. For more insight into how this might look, see the discussion later in the chapter of how information systems are organized.

ERP Management

Organizations using an ERP require one or more individuals to manage these systems. These people make sure that the ERP system is completely up to date, work to implement any changes to the ERP that are needed, and consult with various user departments on needed reports or data extracts.

Project Managers

Information-systems projects are notorious for going over budget and being delivered late. In many cases, a failed IT project can spell doom for a company. A project manager is responsible for keeping projects on time and on budget. This person works with the stakeholders of the project to keep the team organized and communicates the status of the project to management. A project manager does not have authority over the project team; instead, the project manager coordinates schedules and resources in order to maximize the project outcomes. A project manager must be a good communicator and an extremely organized person. A project manager should also have good people skills. Many organizations require each of their project managers to become certified as a project management professional (PMP).

Information-Security Officer

An information security officer is in charge of setting information-security policies for an organization and then overseeing the implementation of those policies. This person may have one or more people reporting to them as part of the information security team. As information has become a critical asset, this position has become highly valued. The information-security officer must ensure that the organization's information remains secure from both internal and external threats.

8.4: Computer and Information Technology Occupations

According to the Bureau of Labor Statistics, employment in computer and information technology occupations is projected to grow 13 percent from 2020 to 2030, faster than the average for all occupations. These occupations are projected to add about 667,600 new jobs. Demand for these workers will stem from greater emphasis on cloud computing, the collection and storage of big data, and information security.

The median annual wage for computer and information technology occupations was $91,250 in May 2020, which was higher than the median annual wage for all occupations of $41,950 (Bureau of Labor Statistics).

	OCCUPATION	JOB SUMMARY	ENTRY-LEVEL EDUCATION	2020 MEDIAN PAY
	Computer and Information Research Scientists	Computer and information research scientists design innovative uses for new and existing computing technology.	Master's degree	$126,830
	Computer Network Architects	Computer network architects design and build data communication networks, including local area networks (LANs), wide area networks (WANs), and Intranets.	Bachelor's degree	$116,780
	Computer Programmers	Computer programmers write and test code that allows computer applications and software programs to function properly.	Bachelor's degree	$89,190
	Computer Support Specialists	Computer support specialists provide help and advice to computer users and organizations.	See How to Become One	$55,510
	Computer Systems Analysts	Computer systems analysts study an organization's current computer systems and find a solution that is more efficient and effective.	Bachelor's degree	$93,730
	Database Administrators and Architects	Database administrators and architects create or organize systems to store and secure data.	Bachelor's degree	$98,860
	Information Security Analysts	Information security analysts plan and carry out security measures to protect an organization's computer networks and systems.	Bachelor's degree	$103,590
	Network and Computer Systems Administrators	Network and computer systems administrators are responsible for the day-to-day operation of computer networks.	Bachelor's degree	$84,810
	Software Developers, Quality Assurance Analysts, and Testers	Software developers design computer applications or programs. Software quality assurance analysts and testers identify problems with applications or programs and report defects.	Bachelor's degree	$110,140
	Web Developers and Digital Designers	Web developers create and maintain websites. Digital designers develop, create, and test website or interface layout, functions, and navigation for usability.	Bachelor's degree	$77,200

https://www.bls.gov/ooh/computer-and...ology/home.htm

Last Modified Date: Wednesday, September 8, 2021

8.5: Summary

Summary

In this chapter, we have reviewed the many different categories of individuals who make up the people component of information systems.

The world of information technology is changing so fast that new roles are being created all the time, and roles that existed for decades are being phased out.

That said, this chapter should have given you a good idea of the importance of the people component of information systems.

Study Questions

1. Describe the role of a systems analyst.
2. What are some of the different roles for a computer engineer?
3. What are the duties of a computer operator?
4. What does the CIO do?
5. Describe the job of a project manager.
6. Explain the point of having two different career paths in information systems.
7. What are the advantages and disadvantages of centralizing the IT function?
8. What impact has information technology had on the way companies are organized?
9. What are the five types of information-systems users?
10. Why would an organization outsource?

Exercises

1. Which IT job would you like to have? Do some original research and write a two-page paper describing the duties of the job you are interested in.
2. Spend a few minutes on Dice or Monster to find IT jobs in your area. What IT jobs are currently available? Write up a two-page paper describing three jobs, their starting salary (if listed), and the skills and education needed for the job.
3. How is the IT function organized in your school or place of employment? Create an organization chart showing how the IT organization fits into your overall organization. Comment on how centralized or decentralized the IT function is.
4. What type of IT user are you? Take a look at the five types of technology adopters and then write a one-page summary of where you think you fit in this model.

1. Rogers, E. M. (1962). Diffusion of innovations. New York: Free Press↵

CHAPTER OVERVIEW

9: Introduction to Web Development

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- the ability to create an HTML document structured to support CSS styling
- the ability to create a CSS file that adapts styling based on device capabilities
- the ability to create basic images using canvas
- the ability to integrate audio and video to a page
- the ability to utilize special device features
- the ability to integrate external font styles

9.1: An Introduction

Web development is an evolving amalgamation of languages that work in concert to receive, modify, and deliver information between parties using the Internet as a mechanism of delivery. While it is easy to describe conceptually, implementation is accompanied by an overwhelming variety of languages, platforms, templates, frameworks, guidelines, and standards. Navigating a project from concept to completion often requires more than mastery of one or two complementing languages, meaning today's developers need both breadths, and depth, of knowledge to be effective.

Brief History of the Internet

The Internet started in the 1960s as a way for government researchers to share information. Computers in the '60s were large and immobile and in order to make use of information stored in any one computer, one had to either travel to the site of the computer or have magnetic computer tapes sent through the conventional postal system.

Another catalyst in the formation of the Internet was the heating up of the Cold War. The concept of a network connecting computers was under development by both government and university researchers looking for a better means to communicate and share research. The military at the time relied in part on microwave transmission technology for communications. An unexpected attack on some of these towers demonstrated how susceptible the technology was to the failure of even small portions of the transmission path. This led the military to seek a method of communicating that could withstand attack. At the same time, university researchers were trying to share their work between campuses and were struggling with similar problems when their transmissions suffered drops in signal. Parties from both groups ended up at the same conference with presentations and decided to collaborate in order to further their work.

As cold war tensions grew and Sputnik was launched, the United States Department of Defense (DoD) began to seek additional methods of transmitting information to supplement existing methods. They sought something that was decentralized, allowing better resiliency in case of attack, where damage at one point would not necessarily disrupt communication. Their network, Arpanet, connected the DoD and participating universities together for the first time. In order to standardize the way networked systems communicated, the Transfer Control Protocol/Internetwork Protocol (TCP/IP) was created. As various network systems migrated to this standard, they could then communicate with any network using the protocol. The Internet was born.

The E-mail was soon to follow, as users of the networks were interested in the timely transmission and notification of messages. This form of messaging fit one of their initial goals. As time progressed, additional protocols were developed to address particular tasks, like FTP for file transfers and UDP for time-sensitive, error-resistant tasks.

Ongoing improvements in our ability to move more information, and move it faster, between systems progressed at a rate similar to the calculative power of the computers we saw earlier. This brings us to where we are today; able to watch full-length movies, streamed in high-quality right to our phones and computers, even while riding in a car.

Website Design

Website design is a topic of study often neglected until after a programming background has been developed. Worse, it may be entirely ignored or missed by computer science students when courses covering the topic are in other programs like graphic arts or media. This results in programmers trying to understand how to write code meant for layout and design elements without understanding design. By studying these elements first, we can develop a better knowledge of the concepts of web design before we write code. Progressing through the topics in this section during your site design will greatly ease your development efforts in the future, allowing stakeholders to understand the project and provide feedback early on, reducing (re)development time.

A number of factors affect design in web development, complicating what would otherwise appear to the end-user to be a relatively simple process of displaying a picture or document. In truth, the development process involves not only the HTML and multimedia that make up the visual aspects of the page but also considerations of software engineering, human-computer interaction, quality assurance and testing, project management, information and requirement engineering, modeling, and system analysis and design.

Today's sites are now becoming more application-centered than traditional sites. This further complicates our projects as we integrate with legacy software and databases, strive to meet real-time data demands, address security vulnerabilities inherent to the environment we are working in, and ongoing support and maintenance typical of robust software applications.

Planning Cycle

Web development is best achieved as a linear process but is usually completed asynchronously. The planning process described is intended to build upon itself to refine project requirements, look and feel, and development plans. However, limitations in timelines, mid-project revisions, and the extensive time that can be invested in the early stages of design lead many programmers to begin development while a project is still in design.

Starting early with programming during design planning can accelerate a project when the elements created early on are unlikely to be affected by later changes in the scope. When done carefully, early programming also allows an opportunity to test concepts before investing time into an idea that may not work. It is important to avoid aspects that are assumed to change, like the visual layout or particular pieces of content, instead focusing on the data structure, frameworks, and other components that are easily adapted to design changes.

While you are planning, keep an eye out for indicators that things are going off-track. Some of the more important flags that should be resolved include:

1. Vaguely defined use cases and inadequate project requirements
2. Overly broad or undefined scope of features
3. Unresolved disputes between stakeholders about project features
4. Unrealistic timetable, budget, or inadequate resources

When considering your milestones, tasks, objectives, or whatever label you or your team place on objectives, a handy acronym to reference is SMART. SMART stands for Specific, Measurable, Attainable, Realistic, and Timely. The idea is to check all of your objectives against these criteria to determine if they are appropriate and well-developed. By ensuring all of your objectives meet the SMART criteria, you will have a better chance of keeping your project on time and well planned

The Fold

As we begin to develop our pages, we need to begin to consider where we want to place pieces of our content. If you look at newspapers, you will find that the most attractive story of the day (as decided, at least, by the publisher) is emblazoned in large letters near the middle or top of the front page, surrounded by the name of the paper, the date, and other pieces of information that quickly lend to your decision of whether or not to purchase a given paper. This is done intentionally, to make the paper attract your attention and get you to buy their edition over their competitors. In the printed news industry, the prime retail space in the paper is the top half of the front page, or what you see when the newspaper is folded normally at a newsstand. This is referred to as “above the fold,” and is crucial to get their audience’s attention. This also applies to websites, except in our case, our “above the fold” is what the user sees on the landing page for our site, without having to scroll down or use any links.

What you typically find here is the name and or logo of the company, and what they feel is most important for you to see first. As you begin to analyze web pages in this light, you will find it very easy to determine what kind of site they are, or what they want or expect from you as their guest. News sites will typically follow a similar setting to a printed paper, leading with headlines and links to other sections. Companies will lead with a featured product or sale to attract your attention, and search engines will make the search bar prominent, usually with ad space close by to increase their revenue streams.

The concept of “The Fold” is another of the many highly argued concepts in web development. Proponents are quick to point out the same example I used of traditional print media methods, while detractors will argue that if it were true, scrolling would never have been created, or users would lose interest in following links. While I endorse “The Fold” as a useful approach to landing pages, I do not mean to imply that all of your pages should fit on only one, non-scrolling screen.

Typography

Typography is the study of the font. While an important topic in media arts, it has until recently received little attention in web development. Utilizing unusual fonts used to be a complicated process that required the end user to have your font(s) installed in order to see the site as you intended. Now, advances in CSS allow us to use unusual fonts by connecting to them through our styling. This allows us to use a tremendous variety of fonts in our sites to add to our look and feel, adding an aspect that has unlocked new approaches to design. Some of the elements of typography include the study of features like readability, conveying meaning or emotion through impression, and the artistic effect of mixing styles.

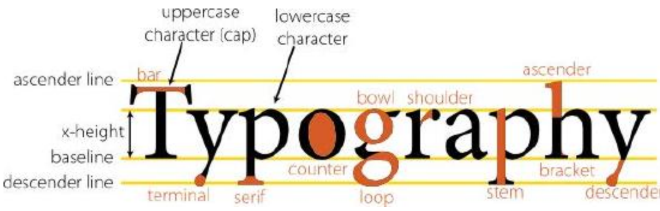


Figure 10 Typography by openclipart.org Public Domain

For ease of reading and to avoid a cluttered appearance, most sites keep to two or three fonts when creating their design. One for text, and one or two for headings, titles, and distinguishing marks. All of these should be kept in the same family for a more congruous experience, and each unusual font defined in your site should include fallback definitions in case there are problems loading your primary style (we will see this in examples later on). You may want to set your regular text as one of the standards supported by all browsers as users are most familiar and usually comfortable with that set.

Web Fonts

To tap into this aspect of design, a great place to start is with the Google Fonts website. This site is a repository of character sets for a great variety of fonts that you can link to or download and include in your own site's files. We will look at connecting to these later, but browsing the site now will give you an initial look at the amount of variety that is available for design.

Site Maps

A sitemap is a file that contains a master list of links to the pages on your site and can provide information about those pages like how often they are updated, how pages connect to each other, and how important it is relative to the other pages. It can be a reference tool to both Bots that index your site for search engines, as well as your visitors trying to find particular content. Sitemap files are XML documents arranged in the hierarchical format that bots read to gain an understanding of your site layout, page relevance, and organization. The file may also be a human-readable page that diagrams how pages relate to one another and serve as a master list of the pages in your site. Sitemaps are best kept in the root of your website, at the same level as your initial index page.

While a complete sitemap cannot be finalized until after your site is ready to be published, I include sitemaps under development methods because laying out your site's organization on paper will help with developing your menu system, logically organizing content, and in defining the scope and purpose of your site. The more content or pages you can define at the beginning of the process will reveal information that will help during your design phase.

To create a sitemap, you can start by creating a running list of all the content you wish to have on your site. Anyone involved in the production or validation of content should be in the room! In each of these steps, it is important to identify your stakeholders. As you are creating your running list, it is often helpful to use index cards so you can determine by card color or pile where a particular piece of content should be. This will help you discover your menu system, as you create names for piles of cards as your menu title. After, as you diagram what cards are with what pile and where that pile is relative to others, your sitemap will begin to take shape.

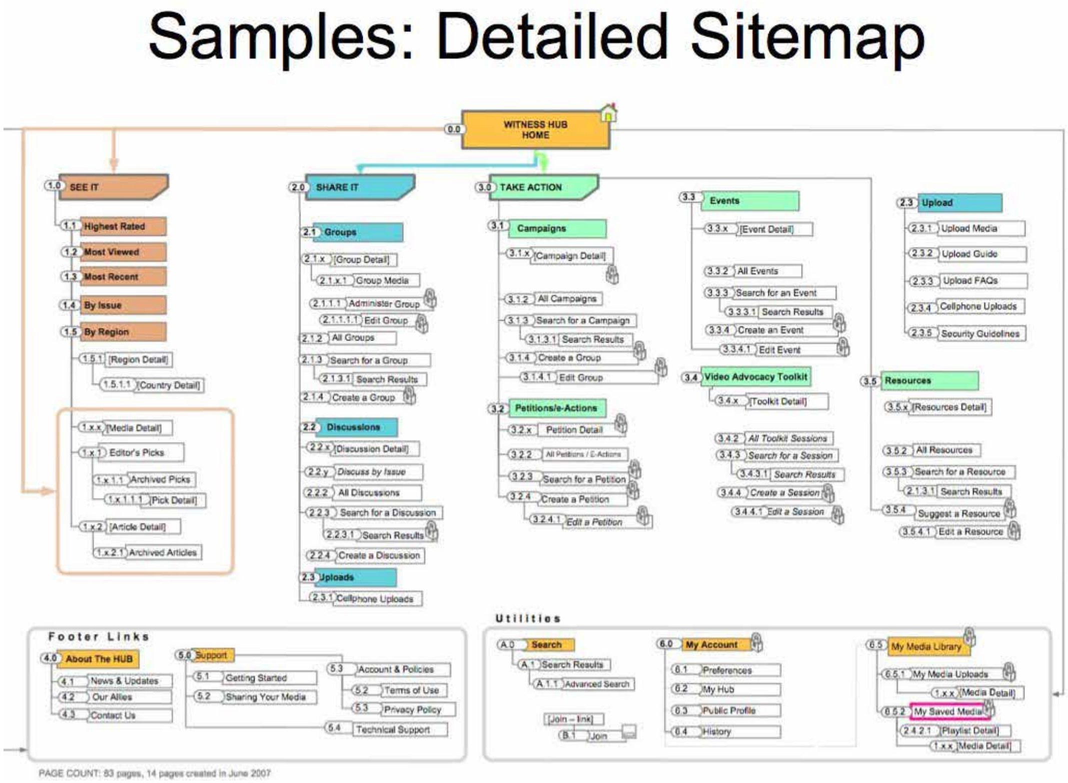


Figure 11 Site Map Creative Commons 2.0 Licensed by Kent Bye

Robots.txt

Robots are automated scripts typically used to index or take inventory, of the content in a website for use in things like web searching sites or collecting statistics. A robots file is a basic text file kept in the root folder of your website that instructs these robots on what sections or types of content in your site you do or do not want them to index. Legitimate robots will read this file when they first arrive on your site to honor your request. Keep in mind this is an enforceable act, and malicious or less-than-reputable robots are still perfectly capable of reading through all non-privileged (i.e. no login required) content on your site.

The simplest robots.txt file involves only two lines:

```
User-agent: *
Disallow: /
```

The first specifies that the rules below apply to all robots that read the file. The second adds that nothing is allowed below (meaning deeper, or all the files and folders inside of) the root folder (/ represents the main folder of the site). If we wanted to be more specific about what sections we want to keep bots out of, we can identify them individually instead of the whole site:

```
User-agent: *
Disallow: /pictures/reserved/
Disallow: /index.php
Disallow: /media
Disallow: /scripts
```

To distinguish cases where a particular bot has a different set of permissions, we can use the bot's name in place of our “all” wildcard:

```
User-agent: BadBot
Allow: /About/robot-policy.html
Disallow: /
User-agent: *
Disallow: /pictures/reserved/
Disallow: /index.php
Disallow: /media
Disallow: /scripts
```

The above settings tell BadBot that it is allowed to see the policy file, but nothing else. It still specifies the blocked paths for the rest of the bots that might visit.

Wireframes

Wireframes in the web development world are not exactly their literal three-dimensional counterparts in the real world, but they bear a similar purpose. A wireframe may include things like location and size of elements such as a login button, where banners and content sections will sit, and provide an overall idea of how a site will operate. When wireframing a website, the idea is to create a mockup of one or more designs that portray how the interface might appear to the user. By the end of your wireframing process, you should have an idea of how the site will operate, and have resolved questions over where users will find particular features and elements.

Wireframes typically do not include color, actual content, or advanced design decisions like typography. Some of these considerations will have been at least partly addressed when creating your sitemap, and the rest will come once we begin storyboarding.

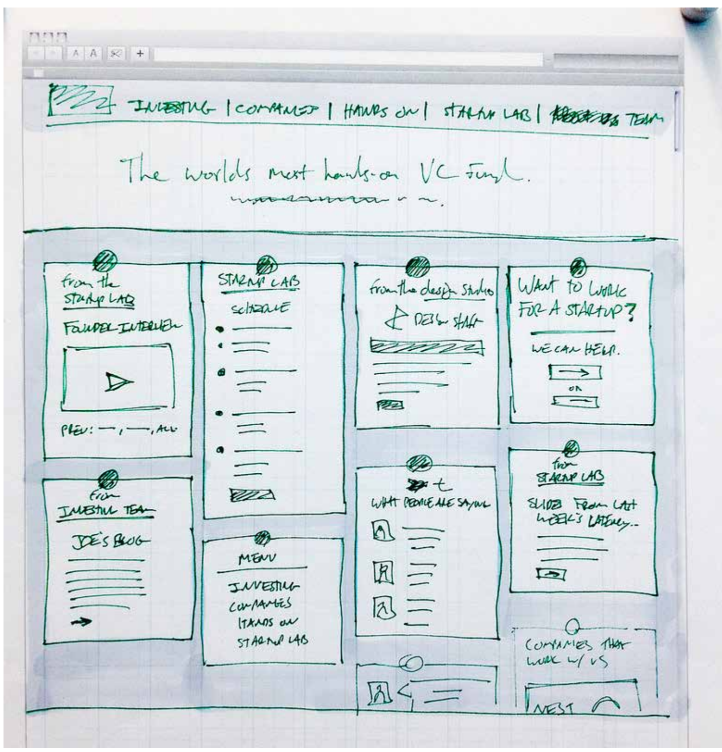


Figure 13 Wireframe Creative Commons 2.0 Braden Kowitz

Storyboarding

Storyboarding a website is quite similar to storyboarding a TV show, comic, or other forms of media. Using our wireframes, we can begin to add color, font, and rough images to our documents. Keep in mind at this point we are probably still in a graphics editor or document style program like PowerPoint, Photoshop, etc. Real code is coming soon, but we can do more mock-ups faster without taking the time to make it function.

As you storyboard, you will create separate pages, or panels, for the screen a user would see as they complete the most important processes on your site. If you are selling something, for example, your storyboard may include examples of product pages, adding items to their cart, logging in, and completing their purchase.

By paging through these panels, you can see how the user experience will progress and identify potential problems like a confusing check out process, or you may discover that your shopping cart block from wireframing may be better off in a different, more predominant location. This process may be repeated several times until a final version is accepted by everyone in the decision process.

Figure 14 Storyboarding Creative Commons 2.0 Rob Enslin

Color Schemes

The process of determining the color(s) involved in your site could fill a book. In fact, it does. Regardless of the varying opinions of what emotions colors instill, or represent, the quickest way to alienate a user is to give them a visual experience that is unappealing. The layout, appearance, and cohesiveness of your site are something that is immediately judged when a user first visits. These elements influence everything from their impression of what the site represents, its reputability, and even its trustworthiness as an ecommerce option. If your site appears to be disorganized, dated and out of style, or seems too "busy" or complicated, you can lose users in less than ten seconds.

You can address this issue (even without an artistic eye) by following the techniques we discussed earlier to plan out a simple, intuitive interface, and by using tools to help you select and compare color schemes like http://colorschemedesigner.com/ or http://www.colorsontheweb.com/colorwizard.asp.

Whether you choose to study other books on the subject or not, a great way to keep current is to get ideas from what others are doing by following sites that list or rate sites by appearances such as the annual Times review and http://www.thebestdesigns.com/.

Development

File Organization

As we begin to create more and more files to complete our website, keeping everything in one folder will quickly grow cluttered. To address this, we can create folders just like we do when sorting files in My Documents. Traditionally you can find folders for images, scripts, pages or files, or for different sections of content or tasks, like an admin folder or ecommerce store. How and why these folders are created varies to personal taste or group determinations, and in some cases is done to maintain a particular method of writing code such as model-view-controller.

Pseudo-Code First!

Whiteboards, notepads, and napkins are your friends. Writing out how you plan to tackle a particular problem will help you identify logic problems before you are halfway through coding them, and will help you keep track of what you need to work on as you progress. Creating pseudo-code is the process of writing out in loosely structured sentences what needs to be done. For example, if my task is to look at each element in an array and tell the user if it is true or false, we might draft the following:

```
foreach(thing in array){
    if(thing / 2 is 0) then show Even
    else show Odd
}
```

Imagine that while writing this example out we realize that we want to store the responses for use again later, not just show them to the user. So, let us update our pseudo-code to take that use case into mind:

```
foreach(thing in array){
    if(thing / 2 is 0) then add to even array
    else add to odd array
    send arrays back in an array
}
```

Reviewing what we have now, not much looks different, but so far we have not had to rewrite any code either. After some thought, it might occur to us that creating two additional arrays could be more memory intensive and less efficient than editing the one we already have. So, to simplify things and possibly improve performance, we might try this:

```
foreach(thing in array){
    if(thing / 2 is 0) then add it to even array, delete from this array
    send arrays back in an array
}
```

Finally, since we are now editing our existing array, we need to make sure we reference it (ensuring our changes are reflected after the for each completes), which also means we only have to pass back our even array:

```
foreach(reference! Thing in array){
    if(thing / 2 is 0) then add it to even array, delete from this array
    send even array back
}
```

While none of the above examples would work (they are just pseudo-code), we were able to edit a conceptual version of our program four times. Alternatively, we would have spent time creating and revising actual lines of code three times, only to keep finding an issue, backing up, and make lots of changes.

Comments

To quote Eagleson’s Law, “Any code you have not looked at for six or more months might as well have been written by someone else.” This is not to say that your style or approach will change drastically over time, but that your immediate memory of what variables mean, why certain exceptions were made, or what the code, ultimately, was meant to address may not be as apparent as when you last worked on the file. It is natural for us to feel that we will remember these details as they are so obvious when we are creating them. The need for good commenting becomes immediately apparent when reviewing someone else's work, and you are wallowing in frustration trying to figure out what that person was trying to do. This is not meant to endorse comments that are obvious to the reader from the line, like:

```
$int = 2 + 2 //we added 2 and 2 together
```

Rather, comments are best suited to explaining why agreed upon methods were not used or what difficult to understand code might be doing, like:

```
// This block of code checks each text file in the folder for the given date and dele
// This is legacy code from when Frank M. worked here. If we change anything, payroll
```

Spacing

Just as we use spacing in documents to convey that a topic change is occurring, we can break up a longer string of commands by putting spaces around lines that are grouped together to complete a particular task, signifying that the next set of lines is for another task.

Brackets

Some languages require the programmer to use a combination of parenthesis and brackets to identify what pieces of code belong together. This allows the engine or compiler to delineate between the code that should be tested as a logic statement, code that gets executed if that statement is true, and code that belongs to functions or classes.

As we write our code and reach instances where we need these elements, it is good practice to immediately enter both the starting and ending marker first, then create space between them to enter your code. This will help ensure that you do not forget to close brackets later or close them in the wrong places when nesting code. Depending on your text editor, it may assist you by automatically adding closing brackets and helping you identify which opening and closing brackets go together.

Indentation

To make your code easier to read, you can use indentation to give the reader an idea of what lines of code belong to different sections. This is typically done inside functions, classes, and control structures. When we nest code, extra indentations are added for each layer within, moving those blocks of text further right to visually distinguish. As we finish the contents of a loop or function, our closing bracket is lined up with our function definition or logic statement to show that the section of code belonging to it is complete.

While our program will run just fine without indentation, it makes it easier to see where you are in your program and where the line you are looking at is intended to be in the logic flow.

Figure 18 Code Formatting Examples

Meaningful Variable Names

When you create variables and functions, try to create names that will have meaning not only to you but to others who may read your code. While it can be tempting to use a lot of short variable names while writing your code, it will be more difficult later to follow what the variable is supposed to represent. You might decide to use short names like `queryResult` or `query_result` or something longer like `numberOfResumesReceived` . While the latter takes longer to type while coding, the name is very clear on what it represents. As spaces are generally prohibited in variable names, these examples show us a few ways to approach longer names. The method you use is up to you but should be used consistently throughout your code to reduce confusion. Differences in how and where you use capitalization or underscores can be used to represent different types of variables like classes or groups of variables.

Short variable names like a simple x or generic name like temp may have their places in your code but are best reserved for when they identify a small variable or one which will have a very short shelf life in your code.

Versioning

This is the process of creating multiple versions of your software, instead of continuously overwriting your sole edition of the code. By creating different copies of your program as you create new features, you can preserve working copies or even create different versions of your program. This allows you to "roll back" or restore previous versions if unforeseen errors are created in new code, or to allow different features to be tried and discarded over time. Naming conventions for different versions of your code might involve numbers, letters, release stages (i.e. alpha, beta, release candidate, and release) a combination of all of these, or just "development" and "live."

Templates

Similar to the idea behind frameworks, templates are sets of files that dictate the basic structure that provides a layout to your site. Templates typically create a grid format you can select from, like two or three columns, fixed or relative width and height, etc. If you are starting a site fresh and putting it into an empty template, there may be some placeholder content and styling as well. Templates are useful for getting the look and feel of a site up and running fast and there is little concern about the particulars of appearance, or whenever the template meets your needs well. When inserting your content dynamically, multiple templates can be used for one site to change the look and feel quickly based on which one is applied. This might be determined by what type of device your guest is using, or what type of authentication they are using.
Templates can be both freestanding or can be an extension of a content management system or framework.

Integrated Development Environments

This list is by no means comprehensive. These editors are sufficient to get you started. If you wish to continue in web programming, you may elect to invest in a development platform like Adobe Dreamweaver or another professional product that supports the more advanced design or try any number of other IDEs available that focus on a variety of different languages.

You might consider the following programs to help you write your code (listed in no particular order). Each of these has features particular for web development and should be sufficiently capable to get you through the examples in this text.

Jedit

A free editor based around Java. Works on multiple platforms (Windows, Mac, and Linux) and includes syntax highlighting.

Notepad++

Notepad++ is a source code text editor with syntax highlighting, multiple document handling using tabs, auto-completion of keywords (customizable), regular expressions in the search and replace function, macro recording and playback, brace and indent highlighting, collapsing and expanding of sections of code, and more.

Bluefish

Supports many programming and markup languages. An open source development project, multi-platform, and runs on Linux, FreeBSD, MacOS-X, Windows, OpenBSD, and Solaris.

TextWrangler

This editor is related to BBEdit. It does not include as many tools but retains syntax highlighting and the ability to use FTP within the editor.

HTML-Kit

This editor is intended for use by web developers and comes with support for writing HTML, XML, and scripts. Among its features is the internal preview of your web, integration with HTML Tidy, auto-completion of keywords, etc. Look for the "Previous" version for their free copy.

9.2: HTML 101

What is HTML

Hypertext Markup Language (HTML) is a standard markup language for tagging text files to achieve font, color, graphic, and hyperlink effects to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript.

Web browsers obtain HTML documents from either a web server or local storage and produce the documents into multimedia web pages. HTML elements are the foundation of HTML pages. Images and other objects such as interactive forms may be embedded into the rendered page using HTML constructs. HTML provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes, and other items. Tags are written using angle brackets to delineate HTML elements. Browsers do not display the HTML tags but use them to interpret the content of the page.

Creating HTML Files

File Format

Before we can create our first web page, we need to create a file that our service will recognize as a web page. To do this, we can open our chosen text editor (see a shortlist of potentials in the section on Integrated Development Environments), and create a new document if one was not created automatically. We will immediately select “Save As” from your editor’s File menu, and give your new page its name. If this is going to be the front page or first page you want a user to see for your site, you should name it index. The index is the default file name most web servers look for in any folder of your website; it saves your users from having to know and type the page name as part of the URL.

In many text editors, underneath or near where you enter the file name is another drop-down selector that allows you to pick a file type. This is the extension (what comes after the period in the file name), or file type, that identifies what kind of data the file represents. This tells our operating system, applications, and browsers what conventions were used to create the content so it can be reassembled into a usable form. Since we are creating a basic web page, we will use the .htm extension (.html is also acceptable, just be consistent to make your life easier). If your editor does not have .htm or .html in its list, then select “All” and make index.htm your file name.

Additional Notes

If you ever come across an unfamiliar extension and want to know more about it, sites like filext.com can help you determine what programs can open it and what it is for. Once we have saved our file as index.htm, we are ready to begin. Saving as soon as we create a file is useful as the text editor will then know what syntax is expected. This will enable features like color coding and highlighting that your editor supports.

Document Type

Every HTML page we create should declare its document type (doctype) in the first line. This will identify which spec of HTML is included so the browser knows how to interpret the tags within. Earlier version of the HTML specifications used two definitions for HTML: HTML 4.01 and XHTML. Both of these contained two additional properties of strict and transitional.

With HTML5, much of this has been eliminated, leaving one general doctype declaration of <!DOCTYPE html>. This should be the first line of code in any HTML page you create. We will not cover the older doctype formats as all of our examples will focus on HTML5. Keep in mind, though, that code examples you find online with anything other than the tag above may be outdated approaches to what is shown.

Learn more

Keywords, search terms: xml, html, css, dom, document markup,

W3C Documentation for XML: http://www.w3.org/TR/2004/REC-xml11-20040204/

W3C Documentation for CSS: http://www.w3.org/Style/CSS/

W3C Specifications for HTML:http://www.w3.org/community/webed/wiki/HTML/Specifications

Navigation

A feature found on almost every website is a navigation system for moving between pages. These are typically menus, where groups of common pages are created to give the site a hierarchical organization. While the approach to visual styling and interaction with menus comes in great variety, most follow a basic principle of using unordered lists of links, and the application of CSS to those lists in order to turn them into colorful, interactive elements we are accustomed to. While there are drawbacks that we will discuss in Visually Impaired Considerations, alternative approaches can still utilize linked lists to some extent.

Since we created our menu earlier, we already know the contents and structure of our navigation. Our group label, or top-level labels, and the nested s represent the contents of the list for that menu item.

Some popular approaches to providing a means of navigation are menu bars with dropdowns, bread crumbs, and event driven responses. Menu bars are the most frequently utilized element, where hovering or clicking over an item in the menu brings up additional choices related to the main item. Typically referred to as drop-down menus, they can be styled to move in any direction. Nesting lists within lists can give us a multi-tier menu that allows us the ability to select from a large number of pages with little effort.

Breadcrumbs are typically depicted as a horizontal delimited list of pages, similar to:

1. Home >> Sports >> Football >> Buffalo Bills >> Patriots >> Golf

The breadcrumb does not follow a hierarchical notation but acts more like a brief history of where you have been on the site, allowing you to skip back several steps at once without using your browser's back button. These can be helpful in sites with large amounts of content where the user's experience may not be particularly linear, as they move between topics or sections, like news or reference sites.

Event-driven navigation is useful in narrowing the user experience to a fixed set of paths. This method will only make certain links available under certain conditions, restricting the options a user has on a particular page to what they are allowed to do, which may be based on a variety of rules such as if they are logged in, previous links, or decisions they have made, or if something in the system needs their attention.

These approaches can be used by themselves, or in combination to provide your user experience.

Linking

Links in HTML can take two forms, both of which are created with the anchor tag (<a>). They can either point to a resource in another location, or to a location within the document. The former is used far more frequently than the latter, however, internal links are coming back into popularity with the rise of infinite scrolling.

Absolute, Base, and Relative Path

The href attribute of an anchor tag defines the actual location the link will represent. Absolute and relative paths are two reference methods for connecting sites and pages. While both methods can be used when creating links that point to content in our own site, only absolute can be used when pointing to content that is outside of your domain.

Absolute paths are the entire length of the link required to identify one resource, whether it is a page, image, script, or media file. The URL http://www.msn.com/news/index.htm tells us we want to go to the index page in the news folder of the msn.com website. If this was our site, and we wanted to go to the index.htm file in the sports folder, we could write it as http://www.msn.com/sports/index.htm (absolute) or ../sports/index.htm (relative). The initial .. instructs the browser that our intention is to go back one layer of depth (i.e. "up" one level in folders) and then into the sports folder, which in this example sits in the same parent folder as our news page.

Using just an initial / without .. tells the server that we want to start at the root folder of the server and navigate from there, meaning we start with the base path.

A base path is everything needed to get us to the index page of the root folder of the site. This is typically http://www.yoursitename.com, and is the part you find missing in the relative path above. The combination of the base path, and relative path, equals your absolute path.

Target

While the anchor tag supports several attributes, one of the most important of these is "target." This attribute describes where links will be loaded, like a new tab or the same tab/browser window we are already using. The attribute can take any of the following values to define that location.

Table 10.2.2.110.2.2.1: *Anchor Targets*

Value	Description
_blank	Opens the linked document in a new window or tab
_self	Opens the linked document in the same frame as it was clicked (this is default)
_parent	Opens the linked document in the parent frame
_top	Opens the linked document in the full body of the window
framename	Opens the linked document in a named frame

From php.net manual, creative commons 3.0 Attribution

Within the Page

We can add links to a page that move the user around the page itself, which is useful on pages with long content. To do this, we use an anchor tag to define where we want our destination to be. When we create our link, we simply reference the name of our anchor, preceded by a pound sign in place of a traditional URL.

- Some text here.
- Click here to go further down.
- Some more text.
- Even more text!
- <a name="ourDestination">
- This is where we want to "jump" to.

Some text here.
Click here to go further down.
Some more text.
Even more text!
This is where we want to “jump” to.

Graphics

Images are the greatest contributors to the visual appeal of your site, and typically account for the majority of bandwidth used in loading your pages. By using a combination of image types, and newer techniques found in HTML5 like canvas, and reproducing images using CSS, we can balance quality against size to reduce our bandwidth needs and allow our site to be more flexible.

Formats

Images are files, just like any other document in your computer, but they can be coded and formatted differently to reproduce the image you want to see. We find these referred to as raster and vector graphics. These formats represent two very different methods of creating an image.

Raster

The image files most of us are already familiar with using are typically raster format. Examples of these are JPEG, GIF and BMP. When we interact with pictures we took on digital cameras for example, we are dealing with JPEG or JPG files. Raster files recreate an image by recording the color value of pixels, which represent the smallest single point on a screen that can be assigned a color by the display. The higher the number of pixels (or density, measured as pixels per inch) translates to how sharp the image is, and how large it can be rendered without losing quality.

The number of colors available in the image file is based on the length of the value available to each point. If we only allowed a single binary character for each pixel point, we would be able to keep our file size as small as possible. This however would mean we could only represent our image in black and white (binary only allows us two options, 0 or 1, so we can only represent two colors.). When we allow longer values to represent a single point, we can assign values a larger range of colors. Once we scale these up, however, we trade away our smaller image sizes in order to have more colorful pictures. Large images can slow down the user experience, and if loading takes too long, users will leave.

Traditionally, we have faced this trade off by using different image formats in different areas of our site. While reserving JPG for our larger images or photos, we can use GIF for smaller icons and indicators. GIFs limit us to 256 colors, but since most icons use few colors, we are able to capitalize on the benefits of this format here. It is important to note that raster images will quickly lose quality when rendered at sizes larger than the original image's width or height.

Vector

Vector images store information about colors and locations as definitions of angles, lines, and curves in mathematical format. The benefit of a vector formatted image is that it can be scaled both up and down in size without distortion or degradation in quality. This is due to the fact that the image is "drawn" by the browser each time it is loaded, and the processor performs the steps necessary to recreate the image. Since the image can be scaled, the same image file can be drawn very large, or very small, without changing the file size. We will get some hands-on experience in how vector images are drawn when we look at the new Canvas features in HTML.

Table 10.2.3.110.2.3.1: *Image Formats*

Format	Compression	Platforms	Colors	Notes
JPEG (Joint Photographic Experts Group)	Lossy	Unix, Win, Mac	24-bit per pixel; 16.7 million colors.	JPEG is a compression algorithm; the format is actually JFIF (JPEG File Interchange Format)
GIF (Graphic Interchange Format)	Lossless	Unix, Win, Mac	8-bit; 256 colors (216 web palette). Allows transparency.	LZW compression algorithm developed by CompuServe; patent now held by Unisys, which charges for use of the code in graphics programs. Once Unisys began enforcing its patent (in 1995), programs began moving to PNG.
BMP (Bitmap graphics)	Uncompressed	Win	24-bit; 16.7 million colors.	Like all uncompressed formats, these files are very large.
PICT	Lossless	Mac		Very little compression; large files
TIFF (Tag Interchange File Format)	Lossless or uncompressed	Unix, Mac, Win		TIFF-LZW uses the proprietary LZW compressions (see GIF).
PNG (Portable Network Graphics)	Lossless	Unix, Mac, Win	48-bit; "true color" plus transparency	Will likely replace GIF. Supported in IE, NN 4 and above. A WC3specification.

http://mason.gmu.edu/~dtaciuch

You may notice the compression column. This is the act of removing or modifying the data that represents a file in a manner that makes its overall file size smaller. By doing this, we can transmit files faster, and they will take up less space in memory. When we discuss compression in terms of graphics we need to consider whether it will result in a lossy or lossless result. A lossless result means the compression techniques used do not remove data from the original copy, so we can restore the image to its exact original size and appearance. A lossy compressions structure can result in greater compression, but achieves the extra advantage by removing information from the file.

As an example, imagine a picture of you and your friends on the beach with a clear blue sky behind you. The data in the image file will measure the "blueness" of the sky in varying colors of blue, at a level greater than the eye can distinguish. By averaging the blueness and making more of the sky pixels the same color, we have eliminated information. Certain levels of lossy compression will still be indistinguishable from the original, but at any level, the lossy-compressed version of the file will not be restorable to the original because the extra values were eliminated.

Which is better? As usual, it depends on your intent. If the image can be lossy compressed and is still acceptable to you and your users, and having the smallest possible file sizes (good, of course, for mobile devices) is a priority, then go for it. Quality optimized scenarios will likely call for a lossless compression, like in sites that use large images as their background.

Font and Text Decoration

Color for Heading Tags

When we began our testing site, we started using HTML tags wherever we could to provide structure to our content with heading tags. We did this with the understanding that later we would redefine those tags so our headings looked how we wanted them to. That time has come. To start with some basics, we can use what we have already learned above by changing the color of the text and background for our heading tags:

1. <style>
2. h1{
3. color:red;
4. background-color:yellow;
5. }
6. </style>
7. <h1>This is an H1 heading</h1>

We can also adjust our font family and size. You will notice that none of these changes affect anything outside of our headings. While you may see examples using the key terms for size, ranging from "extra extra small" (xx-small) to "extra extra large" (xx-large), it is a good idea to always be as specific as possible, as key terms can be treated differently between browsers. Instead, our examples will use percentage based and fixed font sizes. To do so, we will make our h1 content italicized and bring the size down to 20px tall:

1. <style>
2. h1{
3. color:red;
4. background-color:yellow;
5. **font-style:italic;**
6. **font-size:20px;**
7. }
8. </style>
9. <h1>This is an H1 heading</h1>

These are just a few examples of the full power of font through CSS. Some "fancier" methods include effects like capitalizing, while simultaneously shrinking, your text (small-caps):

1. <style>
2. h1{
3. color:red;
4. background-color:yellow;
5. **font-variant:small-caps;**
6. }
7. </style>
8. <h1>This is an H1 heading</h1>

Additional notes

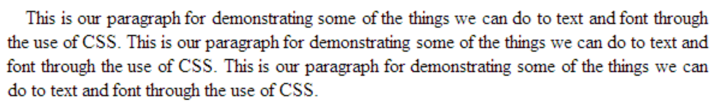
Order is Important! Active style rules must come after hover rules, and hover must come after link and visited! Since a link being hovered over can already have been visited, and the active link can be the one with hover on it, this ensures the correct order of application of style.

Text Styles

While our next example seems like it applies more to font than text, a good way to remember what noun you want to use in your rule is whether the affect changes the way the letters appear or not. If they do, you probably want font. If not, then you probably want text as in these next examples.

First, we might want to add the lead spaces back into our paragraph’s definition to make them appear more like a written document. We can also move our text around in our containing element by setting it to left (default), right, center, or stretch to fit with justify:

```
{
text-indent:15px;
text-align:justify;
}
</style>
<p>
```

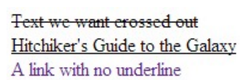


This is our paragraph for demonstrating some of the things we can do to text and font through the use of CSS. This is our paragraph for demonstrating some of the things we can do to text and font through the use of CSS. This is our paragraph for demonstrating some of the things we can do to text and font through the use of CSS.

This is our paragraph for demonstrating some of the things we can do to text and font through the use of CSS. This is our paragraph for demonstrating some of the things we can do to text and font through the use of CSS. This is our paragraph for demonstrating some of the things we can do to text and font through the use of CSS.

In addition to adjusting the font itself, we can decorate it with more affects like crossing it out, underlining it, or specifying that it should *not* be decorated, which is especially useful in eliminating the default lines under links:

```
<style>
.strikeOut{text-decoration:line-through;}
.titles{text-decoration:underline;}
a {text-decoration:none;}
</style>
<span class="strikeOut">Text we want crossed out</span><br/>
<span class="titles">Hitchiker's Guide to the Galaxy</span><br/>
<span><a href="">A link with no underline</span>
```



~~Text we want crossed out~~
Hitchiker's Guide to the Galaxy
A link with no underline

Anchors

Following up on our ability to remove the underline from a link, there are some other special features we can control for our page anchors using CSS. By specifying link, visited, or hover in our link selector, we can control what happens before, during and after a link has been clicked. We can think of these like applying attributes in our HTML tags, except in CSS the special features are called pseudo-classes. Since we can specify any valid CSS rule, we can have our links change colors, alter the backgrounds, change text and font properties, and everything else we will look at. To see some of the basics in action, we will change our text colors for each action, and also our background color when we are hovering. Since you will need to interact with the links to see these in action, we will forgo an image here and you can test the code yourself:

```
<style>
a:link {color:#FF0000; background-color:yellow;} /* unvisited link */
a:visited {color:#00FF00; background-color:orange;} /* visited link */
a:hover {color:#FF00FF; background-color:green;} /* mouse over link */
a:active {color:#0000FF; background-color:white;} /* selected link */
</style>
```

7. Here is our fake link!

Title

The title tag defines the title of the document. The title must be text-only, and it is shown in the browser's title bar or in the page's tab. The title tag is required in HTML documents. The contents of a page title is very important for search engine optimization (SEO)! The page title is used by search engine algorithms to decide the order when listing pages in search results. The <title> element defines a title in the browser toolbar, provides a title for the page when it is added to favorites and displays a title for the page in search-engine results.

Example

Define a title for your HTML document:

```
<!DOCTYPE html>
<html>
<head>
  <title>HTML Elements Reference</title>
</head>
<body>

<h1>This is a heading</h1>
<p>This is a paragraph.</p>

</body>
</html>
```

Paragraph

HTML paragraph <p> is a tag to defines a paragraph on the webpage. The text between the starting tag and the end tag represents a paragraph in HTML. Since the browsers automatically add a single blank line before and after each paragraph element, there is no need to add a line of separation before and after a paragraph.

Example

```
<p>This is a paragraph.</p>
<p>This is another paragraph.</p>
```

Body

The HTML body tag <body> defines the main content of an HTML document displays on the browser. It can contain all the contents of an HTML document, such as headings, paragraphs, images, hyperlinks, tables, lists, etc. The <body> must be the second element after the <head> tag or it should be placed between </head> and </html> tags. This tag is required for every HTML document and should only use once in the whole HTML document.

Example

A simple HTML document:

```
<html>
<head>
  <title>Title of the document</title>
</head>

<body>
  <h1>This is a heading</h1>
  <p>This is a paragraph.</p>
</body>

</html>
```

Horizontal Rule

A horizontal Rule <hr> defines a thematic break to separate two lines or two paragraphs between any two objects. To apply the horizontal rule, the tag is used in an HTML code. The horizontal rule element displays a horizontal rule that is used to separate content (or define a change) in an HTML page.

Example

Use the <hr> tag to define thematic changes in the content:

```
<h1>The Main Languages of the Web</h1>

<p>HTML is the standard markup language for creating Web pages. HTML describes the structure of a Web page, and consists of a series of elements. HTML elements tell the browser how to display the content.</p>

<hr>

<p>CSS is a language that describes how HTML elements are to be displayed on screen, paper, or in other media. CSS saves a lot of work, because it can control the layout of multiple web pages all at once.</p>

<hr>

<p>JavaScript is the programming language of HTML and the Web. JavaScript can change HTML content and attribute values. JavaScript can change CSS. JavaScript can hide and show HTML elements, and more.</p>
```

9.3: Summary

Summary

The internet came about because the government and university researchers were looking for a better means to communicate and share research.

The government and University decided to collaborate.

Most machines were owned by universities, large corporations, or government organizations due to the staffing demands, size, and cost to acquire and maintain them.

Data entry was done by using punched cards of paper, or the newest method of the time, magnetic tapes.

The basic elements of a web server are hardware, operating system, and an HTTP server.

Web development is best achieved as a linear process but is usually completed asynchronously.

Before we can create our first web page, we need to create a file that our service will recognize as a web page.

Maintenance is the **o**ngoing tasks, modifications to support changing requirements or new tasks created by the discovery of bugs.

When creating web pages, the include, graphics, navigation, title, HTML paragraph <p>,The HTML body tag <body>, and horizontal Rule <hr>.

CHAPTER OVERVIEW

10: Internet Privacy, Internet Security, and Netiquette

The growing internet intrusions have devastating implications for our right to privacy. But more than just privacy is threatened when everything we say, everywhere we go, and everyone we associate with is fair game. There are certain rules required to protect your data and ensure proper etiquette when using the internet.

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- Define computer security
- Understand what industries are affected by computer security
- Use tools for communicating online

10.1: Internet Privacy

Internet Privacy

Internet privacy involves the right or mandate of personal privacy concerning the storing, re-purposing, provision to third parties, and displaying of information pertaining to oneself via the Internet. Internet privacy is a subset of data privacy. Privacy concerns have been articulated from the beginnings of large-scale computer sharing.

Privacy can entail either Personally-Identifying Information (PII) or non-PII information such as a site visitor's behavior on a website. PII refers to any information that can be used to identify an individual. For example, age and physical address alone could identify who an individual is without explicitly disclosing their name, as these two factors are unique enough to typically identify a specific person.

Some experts such as Steve Rambam, a private investigator specializing in Internet privacy cases, believe that privacy no longer exists; saying, "Privacy is dead – get over it". In fact, it has been suggested that the "appeal of online services is to broadcast personal information on purpose. On the other hand, in his essay The Value of Privacy, security expert Bruce Schneier says, "Privacy protects us from abuses by those in power, even if we're doing nothing wrong at the time of surveillance."

Levels of Privacy

Internet and digital privacy are viewed differently from traditional expectations of privacy. Internet privacy is primarily concerned with protecting user information. Law Professor Jerry Kang explains that the term privacy expresses space, decision, and information. In terms of space, individuals have an expectation that their physical spaces (i.e. homes, cars) not be intruded on. Privacy within the realm of decision is best illustrated by the landmark case Roe v. Wade. Lastly, information privacy is in regards to the collection of user information from a variety of sources, which produces great discussion.

The 1997 Information Infrastructure Task Force (IITF) created under President Clinton defined information privacy as "an individual's claim to control the terms under which personal information–information identifiable to the individual–is acquired, disclosed, and used." At the end of the 1990s, with the rise of the internet, it became clear that the internet and companies would need to abide by new rules to protect individuals' privacy. With the rise of the internet and mobile networks, the salience of internet privacy is a daily concern for users.

People with only a casual concern for Internet privacy need not achieve total anonymity. Internet users may protect their privacy through controlled disclosure of personal information. The revelation of IP addresses, non-personally-identifiable profiling, and similar information might become acceptable trade-offs for the convenience that users could otherwise lose using the workarounds needed to suppress such details rigorously. On the other hand, some people desire much stronger privacy. In that case, they may try to achieve Internet anonymity to ensure privacy — use of the Internet without giving any third parties the ability to link the Internet activities to personally-identifiable information of the Internet user. In order to keep their information private, people need to be careful with what they submit and look at online. When filling out forms and buying merchandise, that becomes tracked and because the information was not private, companies are now sending Internet users spam and advertising on similar products.

There are also several governmental organizations that protect individuals' privacy and anonymity on the Internet, to a point. In an article presented by the FTC, in October 2011, a number of pointers were brought to attention that helps an individual internet user avoid possible identity theft and other cyber-attacks. Preventing or limiting the usage of Social Security numbers online, being wary and respectful of emails including spam messages, being mindful of personal financial details, creating and managing strong passwords, and intelligent web-browsing behaviors are recommended, among others.

Posting things on the Internet can be harmful or in danger of malicious attack. Some information posted on the Internet is permanent, depending on the terms of service, and privacy policies of particular services offered online. This can include comments written on blogs, pictures, and Internet sites, such as Facebook and Twitter. It is absorbed into cyberspace and once it is posted, anyone can potentially find it and access it. Some employers may research a potential employee by searching online for the details of their online behave ours, possibly affecting the outcome of the success of the candidate.

Risks to Internet Privacy

Companies are hired to watch what internet sites people visit, and then use the information, for instance by sending advertising based on one's browsing history. There are many ways in which people can divulge their personal information, for instance by use of "social media" and by sending bank and credit card information to various websites. Moreover, directly observed behavior, such as browsing logs, search queries, or contents of the Facebook profile can be automatically processed to infer potentially more intrusive details about an individual, such as sexual orientation, political and religious views, race, substance use, intelligence, and personality.

Those concerned about Internet privacy often cite a number of privacy risks — events that can compromise privacy — which may be encountered through Internet use. These range from the gathering of statistics on users to more malicious acts such as the spreading of spyware and the exploitation of various forms of bugs (software faults).

Several social networking sites try to protect the personal information of their subscribers. On Facebook, for example, privacy settings are available to all registered users: they can block certain individuals from seeing their profile, they can choose their "friends", and they can limit who has access to one's pictures and videos. Privacy settings are also available on other social networking sites such as Google Plus and Twitter. The user can apply such settings when providing personal information on the internet.

In late 2007 Facebook launched the Beacon program where user rental records were released to the public for friends to see. Many people were enraged by this breach of privacy, and the Lane v. Facebook, Inc. case ensued.

Children and adolescents often use the Internet (including social media) in ways that risk their privacy: a cause for growing concern among parents. Young people also may not realize that all their information and browsing can and may be tracked while visiting a particular site and that it is up to them to protect their own privacy. They must be informed about all these risks. For example, on Twitter, threats include shortened links that lead one to potentially harmful places. In their e-mail inbox, threats include email scams and attachments that get them to install malware and disclose personal information. On Torrent sites, threats include malware hiding in video, music, and software downloads. Even when using a smartphone, threats include geolocation, meaning that one's phone can detect where they are and post it online for all to see. Users can protect themselves by updating virus protection, using security settings, downloading patches, installing a firewall, screening e-mail, shutting down spyware, controlling cookies, using encryption, fending off browser hijackers, and blocking pop-ups.

However, most people have little idea how to go about doing many of these things. How can the average user with no training be expected to know how to run their own network security (especially as things are getting more complicated all the time)? Many businesses hire professionals to take care of these issues, but most individuals can only do their best to learn about all this.

In 1998, the Federal Trade Commission in the USA considered the lack of privacy for children on the Internet and created the Children Online Privacy Protection Act (COPPA). COPPA limits the options which gather information from children and created warning labels if potential harmful information or content was presented. In 2000, Children's Internet Protection Act (CIPA) was developed to implement safe Internet policies such as rules and filter software. These laws, awareness campaigns, parental and adult supervision strategies, and Internet filters can all help to make the Internet safer for children around the world.

HTTP Cookies

An HTTP cookie is data stored on a user's computer that assists in automated access to websites or web features, or other state information required in complex websites. It may also be used for user-tracking by storing special usage history data in a cookie, and such cookies—for example, those used by Google Analytics—are called tracking cookies. Cookies are a common concern in the field of Internet privacy. Although website developers most commonly use cookies for legitimate technical purposes, cases of abuse occur. In 2009, two researchers noted that social networking profiles could be connected to cookies, allowing the social networking profile to be connected to browsing habits.

In the past, websites have not generally made the user explicitly aware of the storing of cookies, however, tracking cookies and especially third-party tracking cookies are commonly used as ways to compile long-term records of individuals' browsing histories — a privacy concern that prompted European and US lawmakers to take action in 2011. Cookies can also have implications for computer forensics. In past years, most computer users were not completely aware of cookies, but recently, users have become conscious of possible detrimental effects of Internet cookies: a recent study done has shown that 58% of users have at least once, deleted cookies from their computer and that 39% of users delete cookies from their computer every month. Since cookies are advertisers' main way of targeting potential customers, and some customers are deleting cookies, some advertisers started to use persistent Flash cookies and zombie cookies, but modern browsers and anti-malware software can now block or detect and remove such cookies.

The original developers of cookies intended that only the website that originally distributed cookies to users could retrieve them, therefore returning only data already possessed by the website. However, in practice programmers can circumvent this restriction. Possible consequences include:

- the placing of a personally-identifiable tag in a browser to facilitate web profiling, or,
- use of cross-site scripting or other techniques to steal information from a user's cookies.

Cookies do have benefits that many people may not know. One benefit is that for websites that one frequently visits that require a password, cookies make it so they do not have to sign in every time. A cookie can also track one's preferences to show them websites that might interest them. Cookies make more websites free to use without any type of payment. Some of these benefits are also seen as negative. For example, one of the most common ways of theft is hackers taking one's user name and password that a cookie saves. While a lot of sites are free, they have to make a profit somehow so they sell their space to advertisers. These ads, which are personalized to one's likes, can often freeze one's computer or cause annoyance. Cookies are mostly harmless except for third-party cookies. These cookies are not made by the website itself, but by web banner advertising companies. These third-party cookies are so dangerous because they take the same information that regular cookies do, such as browsing habits and frequently visited websites, but then they give out this information to other companies.

Cookies are often associated with pop-up windows because these windows are often, but not always, tailored to a person's preferences. These windows are an irritation because they are often hard to close out of because the close button is strategically hidden in an unlikely part of the screen. In the worst cases, these pop-up ads can take over the screen and while trying to exit out of it, can take one to another unwanted website.

Cookies are seen so negatively because they are not understood and go unnoticed while someone is simply surfing the Internet. The idea that every move one makes while on the Internet is being watched, would frighten most users.

Some users choose to disable cookies in their web browsers. Such an action can reduce some privacy risks, but may severely limit or prevent the functionality of many websites. All significant web browsers have this disabling ability built-in, with no external program required. As an alternative, users may frequently delete any stored cookies. Some browsers (such as Mozilla Firefox and Opera) offer the option to clear cookies automatically whenever the user closes the browser. A third option involves allowing cookies in general, but preventing their abuse. There are also a host of wrapper applications that will redirect cookies and cache data to some other location. Concerns exist that the privacy benefits of deleting cookies have been over-stated.

The process of profiling (also known as "tracking") assembles and analyzes several events, each attributable to a single originating entity, in order to gain information (especially patterns of activity) relating to the originating entity. Some organizations engage in the profiling of people's web browsing, collecting the URLs of sites visited. The resulting profiles can potentially link with information that personally identifies the individual who did the browsing.

Some web-oriented marketing-research organizations may use this practice legitimately, for example: in order to construct profiles of 'typical Internet users'. Such profiles, which describe average trends of large groups of Internet users rather than of actual individuals, can then prove useful for market analysis. Although the aggregate data does not constitute a privacy violation, some people believe that the initial profiling does.

Profiling becomes a more contentious privacy issue when data-matching associates the profile of an individual with personally-identifiable information of the individual.

Governments and organizations may set up honeypot websites – featuring controversial topics – with the purpose of attracting and tracking unwary people. This constitutes a potential danger for individuals.

Flash Cookies

When some users choose to disable HTTP cookie to reduce privacy risks as noted, new types of cookies were invented: since cookies are advertisers' main way of targeting potential customers, and some customers were deleting cookies, some advertisers started to use persistent Flash cookies and zombie cookies. In a 2009 study, Flash cookies were found to be a popular mechanism for storing data on the top 100 most visited sites. Another 2011 study of social media found that, "Of the top 100 web sites, 31 had at least one overlap between HTTP and Flash cookies." However, modern browsers and anti-malware software can now block or detect and remove such cookies.

Flash cookies, also known as Local Shared Objects, work the same ways as normal cookies and are used by the Adobe Flash Player to store information at the user's computer. They exhibit a similar privacy risk as normal cookies, but are not as easily blocked, meaning that the option in most browsers to not accept cookies does not affect Flash cookies. One way to view and control them is with browser extensions or add-ons. Flash cookies are unlike HTTP cookies in a sense that they are not transferred from the client back to the server. Web browsers read and write these cookies and can track any data by web usage.

Although browsers such as Internet Explorer 8 and Firefox 3 have added a ‘Privacy Browsing’ setting, they still allow Flash cookies to track the user and operate fully. However, the Flash player browser plugin can be disabled or uninstalled, and Flash cookies can be disabled on a per-site or global basis. Adobe’s Flash and (PDF) Reader are not the only browser plugins whose past security defects have allowed spyware or malware to be installed: there have also been problems with Oracle’s Java.

Evercookies

Evercookies, created by Samy Kamkar, are JavaScript-based applications that produce cookies in a web browser that actively “resist” deletion by redundantly copying themselves in different forms on the user’s machine (e.g., Flash Local Shared Objects, various HTML5 storage mechanisms, window.name caching, etc.), and resurrecting copies that are missing or expired. Evercookie accomplishes this by storing the cookie data in several types of storage mechanisms that are available on the local browser. It has the ability to store cookies in over ten types of storage mechanisms so that once they are on one’s computer they will never be gone. Additionally, if evercookie has found the user has removed any of the types of cookies in question, it recreates them using each mechanism available. Evercookies are one type of zombie cookie. However, modern browsers and anti-malware software can now block or detect and remove such cookies.

Photographs on the internet

Today many people have digital cameras and post their photographs online, for example, street photography practitioners do so for artistic purposes and social documentary photography practitioners do so to document the common people in everyday life. The people depicted in these photos might not want to have them appear on the Internet. Police arrest photos, considered public records in many jurisdictions, are often posted on the internet by numerous online mug-shot publishing sites.

Some organizations attempt to respond to this privacy-related concern. For example, the 2005 Wikimania conference required that photographers have the prior permission of the people in their pictures, albeit this made it impossible for photographers to practice candid photography, and doing the same in a public place would violate the photographers’ free speech rights. Some people wore a ‘no photos’ tag to indicate they would prefer not to have their photo taken

The Harvard Law Review published a short piece called “In The Face of Danger: Facial Recognition and Privacy Law”, much of it explaining how “privacy law, in its current form, is of no help to those unwillingly tagged.”Any individual can be unwillingly tagged in a photo and displayed in a manner that might violate them personally in some way, and by the time Facebook gets to taking down the photo, many people will have already had the chance to view, share, or distribute it. Furthermore, traditional tort law does not protect people who are captured by a photograph in public because this is not counted as an invasion of privacy. The extensive Facebook privacy policy covers these concerns and much more. For example, the policy states that they reserve the right to disclose member information or share photos with companies, lawyers, courts, government entities, etc. if they feel it absolutely necessary. The policy also informs users that profile pictures are mainly to help friends connect to each other. However, these, as well as other pictures, can allow other people to invade a person’s privacy by finding out information that can be used to track and locate a certain individual In an article featured in ABC News, it was stated that two teams of scientists found out that Hollywood stars could be giving up information about their private whereabouts very easily through pictures uploaded to the Internet. Moreover, it was found that pictures taken by some phones and tablets including iPhones automatically attach the latitude and longitude of the picture taken through metadata unless this function is manually disabled.

Face recognition technology can be used to gain access to a person's private data, according to a new study. Researchers at Carnegie Mellon University combined image scanning, cloud computing, and public profiles from social network sites to identify individuals in the offline world. Data captured even included a user's social security number. Experts have warned of the privacy risks faced by the increased merging of our online and offline identities. The researchers have also developed an 'augmented reality mobile app that can display personal data over a person's image captured on a smartphone screen. Since these technologies are widely available, our future identities may become exposed to anyone with a smartphone and an Internet connection. Researchers believe this could force us to reconsider our future attitudes to privacy.

10.2: Internet Security

Computer Security

Computer security, also known as cyber security or IT security, is the protection of information systems from theft or damage to the hardware, the software, and to the information on them, as well as from disruption or misdirection of the services they provide.

It includes controlling physical access to the hardware, as well as protecting against harm that may come via network access, data and code injection, and due to malpractice by operators, whether intentional, accidental, or due to them being tricked into deviating from secure procedures.

The field is of growing importance due to the increasing reliance on computer systems and the Internet in most societies, wireless networks such as Bluetooth and Wi-Fi – and the growth of “smart” devices, including smartphones, televisions and tiny devices as part of the Internet of Things.

Systems at Risk

Computer security is critical in almost any industry which uses computers. Currently, most electronic devices such as computers, laptops and cellphones come with built in firewall security software, but despite this, computers are not 100 percent accurate and dependable to protect our data (Smith, Grabosky & Urbas, 2004.) There are many different ways of hacking into computers. It can be done through a network system, clicking into unknown links, connecting to unfamiliar Wi-Fi, downloading software and files from unsafe sites, power consumption, electromagnetic radiation waves, and many more. However, computers can be protected through well built software and hardware. By having strong internal interactions of properties, software complexity can prevent software crash and security failure.

Financial Systems

Web sites and apps that accept or store credit card numbers, brokerage accounts, and bank account information are prominent hacking targets, because of the potential for immediate financial gain from transferring money, making purchases, or selling the information on the black market. In-store payment systems and ATMs have also been tampered with in order to gather customer account data and PINs.

Utilities and Industrial Equipment

Computers control functions at many utilities, including coordination of telecommunications, the power grid, nuclear power plants, and valve opening and closing in water and gas networks. The Internet is a potential attack vector for such machines if connected, but the Stuxnet worm demonstrated that even equipment controlled by computers not connected to the Internet can be vulnerable to physical damage caused by malicious commands sent to industrial equipment (in that case uranium enrichment centrifuges) which are infected via removable media. In 2014, the Computer Emergency Readiness Team, a division of the Department of Homeland Security, investigated 79 hacking incidents at energy companies. Vulnerabilities in smart meters (many of which use local radio or cellular communications) can cause problems with billing fraud.

Aviation

The aviation industry is very reliant on a series of complex system which could be attacked. A simple power outage at one airport can cause repercussions worldwide,much of the system relies on radio transmissions which could be disrupted, and controlling aircraft over oceans is especially dangerous because radar surveillance only extends 175 to 225 miles offshore. There is also potential for attack from within an aircraft.

The consequences of a successful attack range from loss of confidentiality to loss of system integrity, which may lead to more serious concerns such as exfiltration of data, network and air traffic control outages, which in turn can lead to airport closures, loss of aircraft, loss of passenger life, damages on the ground and to transportation infrastructure. A successful attack on a military aviation system that controls munitions could have even more serious consequences.

Europe has started to move towards centralised aviation network PENS (Pan-European Network Service) which do provide a common IP-based network service across the European region covering voice and data communication and later on even more with NewPENS, similar USA has in NextGen program.

Consumer Devices

Desktop computers and laptops are commonly infected with malware either to gather passwords or financial account information, or to construct a botnet to attack another target. Smart phones, tablet computers, smart watches, and other mobile devices such as Quantified Self devices like activity trackers have also become targets and many of these have sensors such as cameras, microphones, GPS receivers, compasses, and accelerometers which could be exploited, and may collect personal information, including sensitive health information. Wifi, Bluetooth, and cell phone networks on any of these devices could be used as attack vectors, and sensors might be remotely activated after a successful breach.

Home automation devices such as the Nest thermostat are also potential targets.

Large Corporations

Large corporations are common targets. In many cases this is aimed at financial gain through identity theft and involves data breaches such as the loss of millions of clients' credit card details by Home Depot, Staples, and Target Corporation. Medical records have been targeted for use in general identify theft, health insurance fraud, and impersonating patients to obtain prescription drugs for recreational purposes or resale.

Not all attacks are financially motivated however; for example security firm HBGary Federal suffered a serious series of attacks in 2011 from hacktivist group Anonymous in retaliation for the firm's CEO claiming to have infiltrated their group, and Sony Pictures was attacked in 2014 where the motive appears to have been to embarrass with data leaks, and cripple the company by wiping workstations and servers.

Automobiles

If access is gained to a car's internal controller area network, it is possible to disable the brakes and turn the steering wheel. Computerized engine timing, cruise control,anti-lock brakes, seat belt tensioners, door locks, airbags and advanced driver assistance systems make these disruptions possible, and self-driving cars go even further.Connected cars may use wifi and bluetooth to communicate with onboard consumer devices, and the cell phone network to contact concierge and emergency assistance services or get navigational or entertainment information; each of these networks is a potential entry point for malware or an attacker. Researchers in 2011 were even able to use a malicious compact disc in a car's stereo system as a successful attack vector, and cars with built-in voice recognition or remote assistance features have onboard microphones which could be used for eavesdropping.

A 2015 report by U.S. Senator Edward Markey criticized manufacturers' security measures as inadequate, and also highlighted privacy concerns about driving, location, and diagnostic data being collected, which is vulnerable to abuse by both manufacturers and hackers.

Government

Government and military computer systems are commonly attacked by activists and foreign powers. Local and regional government infrastructure such as traffic light controls, police and intelligence agency communications, personnel records, student records, and financial systems are also potential targets as they are now all largely computerized. Passports and government ID cards that control access to facilities which use RFID can be vulnerable to cloning.

Internet of Things and physical vulnerabilities

The Internet of Things (IoT) is the network of physical objects such as devices, vehicles, and buildings that are embedded with electronics, software, sensors, and network connectivity that enables them to collect and exchange data – and concerns have been raised that this is being developed without appropriate consideration of the security challenges involved.

While the IoT creates opportunities for more direct integration of the physical world into computer-based systems, it also provides opportunities for misuse. In particular, as the Internet of Things spreads widely, cyber attacks are likely to become an increasingly physical (rather than simply virtual) threat. If a front door's lock is connected to the Internet, and can be locked/unlocked from a phone, then a criminal could enter the home at the press of a button from a stolen or hacked phone. People could stand to lose much more than their credit card numbers in a world controlled by IoT-enabled devices. Thieves have also used electronic means to circumvent non-Internet-connected hotel door locks.

Medical devices have either been successfully attacked or had potentially deadly vulnerabilities demonstrated, including both in-hospital diagnostic equipment and implanted devices including pacemakers and insulin pumps.

Impact of security breaches

Serious financial damage has been caused by security breaches, but because there is no standard model for estimating the cost of an incident, the only data available is that which is made public by the organizations involved. “Several computer security consulting firms produce estimates of total worldwide losses attributable to virus and wormattacks and to hostile digital acts in general. The 2003 loss estimates by these firms range from $13 billion (worms and viruses only) to $226 billion (for all forms of covert attacks). The reliability of these estimates is often challenged; the underlying methodology is basically anecdotal.”

However, reasonable estimates of the financial cost of security breaches can actually help organizations make rational investment decisions. According to the classic Gordon-Loeb Model analyzing the optimal investment level in information security, one can conclude that the amount a firm spends to protect information should generally be only a small fraction of the expected loss (i.e., the expected value of the loss resulting from a cyber/information security breach).

Attacker motivation

As with physical security, the motivations for breaches of computer security vary between attackers. Some are thrill-seekers or vandals, others are activists or criminals looking for financial gain. State-sponsored attackers are now common and well resourced, but started with amateurs such as Markus Hess who hacked for the KGB, as recounted by Clifford Stoll, in The Cuckoo’s Egg.

A standard part of threat modelling for any particular system is to identify what might motivate an attack on that system, and who might be motivated to breach it. The level and detail of precautions will vary depending on the system to be secured. A home personal computer, bank, and classified military network face very different threats, even when the underlying technologies in use are similar.

10.3: Netiquette

Network etiquette (Netiquette) is the set of rules that determines how to properly communicate and browse the web.

One important part of netiquette concerns your online safety. By following these basic rules and tips you will avoid most online threats such as phishing attempts, malware infections and more. Why is this important for us? Given the rise of cybercriminal activity in recent years, the need to stay safe on the Internet has never been more pronounced. Most people believe all they need to be secure online is to have an antivirus program and do frequent back-ups, but the truth is, cybercriminals know lot of ways around these. What really keeps you safe is the ability to recognize online threats and how to avoid them, and this is why netiquette is so important. A study from University of Maryland University College analyzed users' knowledge on cyber-security and came up with a few interesting conclusions:

- the more you know about cybersecurity, the likelier you are to adopt online safety measures;
- cybersecurity training and education will make you more open to new online safety methods.

Netiquette is important, so here are the 10 best internet safety tips and netiquette rules you need to be aware of when online.

Keep your software/apps updated and delete the ones you don't use.

Developers and cybercriminals are almost always caught in a cat and mouse game, where cybercriminals search for exploits and developers rush to patch them before too much damage is done. Most of the times, vulnerabilities are quickly updated, but the real issue is that most users will still use an outdated version and will be vulnerable to the exploit. Keeping your software up-to-date will go a long way into keeping you safe. But an equally important step is to remove software and apps you no longer use. Many of these programs still communicate in the background with various servers, and in case of a breach, this data might fall into the wrong hands.

Be careful when dealing with emails from unknown sources

Have you received an email from an unknown source? Do you frequently receive -mails from people you don't know? Just as in real life, an important rule to stay safe on the internet is to be suspicious of strangers. First, don't trust emails from people you never met, especially those that ask you to click a link, open an attachment or send a file to the sender. It's pretty easy to spot phishing emails. What gives most of the away is the urgency with which they ask you to do something, either because your account may be compromised or your online purchase may have encountered some issues you need to sort. To fool the potential victims, the latest trend in e-crime is to deploy spear phishing attacks, where emails appear to come from well-known individuals or banking authority. But in order for a spear phishing attack to be successful, it needs to appear as genuine and believable. So they launch an elaborate identity theft operation to target and steal your sensitive information. So when you do get a phishing email, ignore it completely and don't reply to the e-mail, don't click the (malicious) attachment, and don't click the dangerous links in the e-mail that could download malware on the system

Don't click that link or online ad

You found a pretty cool link on the Internet and it keeps tempting you to click it. But the source of the link (website or email) seems fishy. You ask yourself: What can the bad guys do to me if I click this link? The answer is simple: a lot of things could go wrong. Just by clicking a link in an email or a pop-up window, you could turn your PC into a botnet slave, allow cybercriminals to inject malware into your device, or expose your personal information. You may think that you are safe from all these dangers because you have a good antivirus product, but nowadays traditional antivirus protection isn't enough anymore and you need additional weapons in the fight against online dangers.

Just because it's free, it doesn't mean it's safe

As a rule of thumb, paid software is almost always secure and safe. It wouldn't make sense for a cybercriminal to limit the number of victims by imposing a paywall. Double check free software using Google, especially if it's not a well-known one, such as WinRAR. This filtering process can expose bad reviews, or reveal that the free software is actually a vehicle for malware. While the free software itself might be safe and legit, the website itself may try to compromise your security through drive-by downloads. These are secret downloads carried in the background, that target software vulnerabilities and system exploits. Make sure you use a trusted, free program that automatically updates your vulnerable software applications to close security holes in your system.

Do not reveal sensitive information online

Social media is the first place criminals check to gather information on you. They will use any piece of information they can find, such as your name, birthdate, address, city, your spouse/partners name, what sort of pet you have, name of the pet. Literally, anything that you can think of. Improving your social media security settings is a good first step in preventing an identity theft or doxxing. So, be a bit skeptical about people you meet online and about their intentions. Many people exaggerate their Facebook lives, and depending on how you use it, social media may come back to haunt you.

Keep your account information for yourself

Our credentials for online accounts, user names and passwords, hold the keys to a lot of important information, and for that reason they are the most sought after targets by cybercriminals. Using the same password for more than one site is a risky move, since if a malicious hacker breaks into just one account, he can then take over the others. One of the better online safety tips is to use a different password for every website you register. The downside however is the difficulty of memorizing them. One workaround is to use password managers, these will remember the login details of every site you use and can even generate some strong password for you use.

Report illegal activities or offending content

If you notice offending language attacks, like cyber-bullying, hate speech or any form of harassment, do not hesitate to report it. Using the "Flag as offensive" or "Flag as spam" buttons is proper internet etiquette and should be sufficient to fight back against minor online threats, but other types of content, such child pornography or arms trafficking should be immediately reported to the police. Reporting cyberbullying is takes on a different importance when children are the targets. A report on cyber bullying on 2014 gives the following troublesome conclusions:

- 25% of teenagers report that they have experienced repeated bullying
- 52% off young people report being cyber bullied

- 95% of teens who witnessed bullying on social media report they have ignored the behavior
- cyber bullying affects everyone
- the most common types of cyber bullying tactics reported are mean, hurtful comments.

What you post online stays online forever

- We post photos, remarks, location updates and similar content, which we think is fine because we use an anonymous username on a small niche forum. But niche forums are also indexed by search results, and if you reuse the username, other people can start to connect the dots. This is called doxxing, where people hunt for information about an internet user until they manage to figure out who they are in real life. And you may think that your posts and comments are usually ignored or don't receive much attention, but they still remain there and you never know when they come back at you. Not to mention the fact that search engines save and classify your content on so many online servers. To keep it short, when it comes to posting personal content on forums and similar places, follow these guidelines:

- is this information too personal?
- delete/edit past posts which reveal too much about you
- could your content affect your personal or professional life in the future?

Use anti malware protection before you go online

Don't go online until you have the best anti-malware protection that money can buy. You may think that avoiding adult websites and that sort of thing will keep you safe, but did you know that hackers now hide malicious code even in legitimate websites? And sometimes not even that is enough. Some ransomware programs are so sophisticated, it is technologically impossible to decrypt them, no matter how much you try. And in this case, you really need the best tools out there.

Create back-up copies for your important stuff

Though you may have all the security protection in the world, disaster may still hit your system and your valuable files. It may be a system crash, a hard disk failure, a ransomware attack that encrypts your entire operating system or it may be a human mistake. There are so many reasons something can go wrong for you and your sensitive information, even if you followed all the netiquette rules in the book. To emphasize the importance of this point, we can tell you that a while ago we interviewed a series of security experts in the IT industry, and one recurring theme was the importance of backing up your information.

Conclusion

Common sense is vital on the internet. Most of the threats you will find online rely on human error to break into your system. Technology hasn't made us safer because safety is not a function of technology. Real security comes from people. Technology is just a security tool. In the end, knowledge and netiquette become the two sides of the same coin: Internet protection.

10.4: Summary

Summary

Internet privacy involves the right or mandate of personal privacy concerning the storing, re-purposing, provision to third parties, and displaying of information pertaining to oneself via the Internet. Internet privacy is a subset of data privacy.

Risks to Internet privacy may be encountered through Internet use.

Flash cookies, also known as Local Shared Objects, work the same ways as normal cookies and are used by the Adobe Flash Player to store information at the user's computer.

Evercookies are JavaScript-based applications that produce cookies in a web browser that actively "resist" deletion by redundantly copying themselves in different forms on the user's machine (e.g., Flash Local Shared Objects, various HTML5 storage mechanisms, window.name caching, etc.), and resurrecting copies that are missing or expired.

People depicted in these photos might not want to have them appear on the Internet. Face recognition technology can be used to gain access to a person's private data, according to a new study.

Computer security, or IT security, is the protection of information systems from theft or damage to the hardware, the software, and to the information on them, as well as from disruption or misdirection of the services they provide.

Network etiquette (Netiquette) is the set of rules that determines how to properly communicate and browse the web.

11.1: Internet Browsers

(unable to fetch text document from uri [status: 403 (Forbidden)])

11.2: Basic Browser Tools

(unable to fetch text document from uri [status: 403 (Forbidden)])

11.3: Data and Customer Rights

(unable to fetch text document from uri [status: 403 (Forbidden)])

11.4: Web Searching

(unable to fetch text document from uri [status: 403 (Forbidden)])

11.5: Identifying Search Results

(unable to fetch text document from uri [status: 403 (Forbidden)])

SECTION OVERVIEW

11.6: Web Browsers and the Internet

A web browser is a type of software that allows you to find and view websites on the Internet. There are many different web browsers, but some of the most common ones include Google Chrome, Internet Explorer, Safari, Microsoft Edge, and Mozilla Firefox.

11.6.1: Web Browser Interface

Web Browser Interface

Every web browser contains a different interface, however, there are some common attributes of all of them.

URLs and the address bar

Each website has a unique address, called a URL (short for Uniform Resource Locator). URL's are street addresses for computers, letting your computer know where to connect to on the internet. URLs are entered on the browser's address bar. The address bar is usually located at the top of the website window. Sometimes, address bars also have built-in search features, to search for the address if you do not know the exact website address. Website addresses usually contain **HTTP://** or **https://** which stands for *Hypertext Transfer Protocol or Hypertext Transfer Protocol Secure*. https:// addresses are more secure. The next letters are usually **www** which stands for *World Wide Web*. After *www*, the website address is displayed. Usually, website addresses end in **.com** (commercial), **.org** (organization), or **.edu** (educational institution) depending on their classification. In the address bar, please type or www.clcmn.edu to take you to the CLC website.

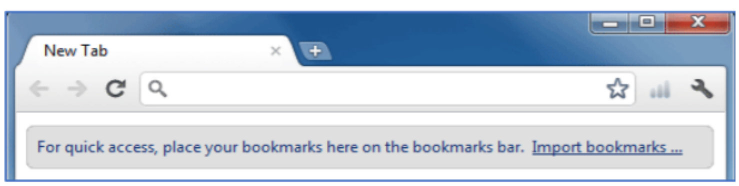
Hyperlinks

The website www.clcmn.edu displayed above that is underlined in blue is an example of a hyperlink, which provides the user with a direct address to a location on the web. Usually, the mouse pointer icon will switch to a  indicating the text is a hyperlink.

Navigation buttons

The Back and Forward arrow buttons allow you to move backward and forward through websites you've recently viewed. You can also click and hold either button to see your recent history. The Refresh button  will load the current website you are viewing a second time. Often a user will click on the refresh button when a website stops working to correct the error or to view new information because the website was updated while the user was viewing the information.

Browser Tabs



Many browsers allow you to open links in a new tab. Each tab is a new instance of the browser to help you manage your applications more efficiently. The more browser tabs that you open, the more system resources you consume. To open a link in a new tab, right-click the link and select Open link in a new tab (the exact wording may vary from browser to browser). You can close a tab by clicking the "X" on the particular tab, and open a new tab by clicking the "+" button on the particular tab.

11.6.2: Bookmarks and Internet History

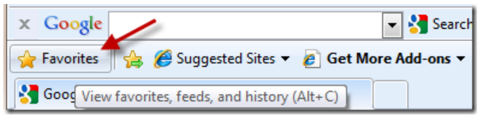


Figure 11.6.2.1 This Photo by Unknown Author is licensed under CC BY-NC-ND

If you find a website you want to access later, it can be difficult to remember all of the websites you want to visit. Bookmarks or favorites allow your computer to remember your favorite website URLs so you do not have to. To create a bookmark, select the Star icon to bookmark the current website. In addition to bookmarks, your browser will also keep track of every website you visit through your internet history. To view your history, open your browser settings—usually by clicking the icon in the upper-right corner—and select History.

11.6.3: Checking for Understanding

1. Demonstrate how to download and retrieve a file.
2. Save a picture from the web to your computer.
3. Please demonstrate how to clear browser history and internet cookies.
4. Navigate to the website www.clcmn.edu. Create a bookmark.
5. List methods to prevent yourself from online.
6. Identify Cyberbullying and share strategies to get help if you are being cyberbullied.

11.7: Web Literacy for Student Fact-Checker Topics

11.7.1: Four Strategies

Four Strategies

What people need most when confronted with a claim which may not be 100% true is *things they can do to get closer to the truth*. They need something we have decided to call strategies.

Strategies represent intermediate goals in the fact-checking process. They are associated with specific tactics. Here are our strategies:

- **Check for previous work:** Look around to see if someone else has already fact-checked the claim or provided a synthesis of research.
- **Go upstream to the source:** Go "upstream" to the source of the claim. Most web content is not original. Get to the original source to understand the trustworthiness of the information.
- **Read laterally:** Read laterally.[1] Once you get to the source of a claim, read what other people say about the source (publication, author, etc.). The truth is in the network.
- **Circle back:** If you get lost, or hit dead ends, or find yourself going down an increasingly confusing rabbit hole, back up and start over knowing what you know now. You're likely to take a more informed path with different search terms and better decisions.

In general, you can try these strategies in sequence, and at each stage if you find success your work might be done.

When you first see a claim you want to check, your first move might be to look to see if sites like Politifact, or Snopes, or even Wikipedia have researched the claim. (Check for previous work).

If you can't find previous work on the claim, the real work begins. It starts by trying to trace the claim to the source. If the claim is about research, can you find the journal it appeared in? If the claim is about an event, can you find the news publication in which it was originally reported? (Go upstream).

Maybe you get lucky, and the source is something known to be reputable — some recognizable source such as the journal Science, or the newspaper *The New York Times*. Again, if so, you can stop there. If not, you're going to need to *read laterally*, finding out more about this source you've ended up at. Is it trustworthy? (Read laterally).

And if at any point you fail — if the source you find is not trustworthy, complex questions emerge, or the claim turns out to have multiple sub-claims — then you circle back, and start a new process. Rewrite the claim. Try a new search of fact-checking sites, or find an alternate source. (Circle back).

1. We are indebted to researcher Sam Wineburg for this language. ↵

Contributors and Attributions

- Template:ContribWebLiteracyStudentFactcheckersCaulfield

11.7.2: Fact-checking Sites

Some Reputable Fact-Checking Organizations

The following organizations are generally regarded as reputable fact-checking organizations focused on U.S. national news:

- Politifact
- Factcheck.org
- Washington Post Fact Checker
- Snopes
- Truth be Told
- NPR Fact-Check
- Lie Detector (Univision, Spanish language)
- Hoax Slayer

Respected specialty sites cover niche areas such as climate or celebrities. Here's a few examples:

- Climate Feedback
- SciCheck
- Quote Investigator

There are many fact-checking sites outside the U.S. Here is a small sample.

- FactsCan (Canada)
- TrudeauMetre (Canada)
- El Polígrafo (Mexico)
- The Hound (Mexico)
- Guardian Reality Check (UK)
- BBC Reality Check (UK)
- Channel 4 Fact Check
- Full Fact (UK)

Contributors and Attributions

- Template:ContribWebLiteracyStudentFactcheckersCaulfield

11.7.3: Wikipedia

Wikipedia is broadly misunderstood by faculty and students alike. While Wikipedia must be approached with caution, especially with articles that are covering contentious subjects or evolving events, it is often the best source to get a consensus viewpoint on a subject. Because the Wikipedia community has strict rules about sourcing facts to reliable sources, and because authors must adopt a neutral point of view, articles are often the best available introduction to a subject on the web.

The focus on sourcing all claims has another beneficial effect. If you can find a claim expressed in a Wikipedia article, you can almost always follow the footnote on the claim to a reliable source. Scholars, reporters, and students all can benefit from using Wikipedia to quickly find authoritative sources for claims.

Each footnote leads to a reliable source. The article as a whole contains over 160 footnotes. If you are researching a complex question, starting with the resources and summaries provided by Wikipedia can give you a substantial running start on an issue.

11.7.4: Activity- Verify a Twitter Account

Activity: Verify a Twitter Account

Kellogg's Rant

General Kellogg was promoted by President Trump as acting head of the National Security Council on February 13, 2017. Is this Twitter account his?

https://twitter.com/GenKeithKellogg/...25494009638912

Explain your reasons.

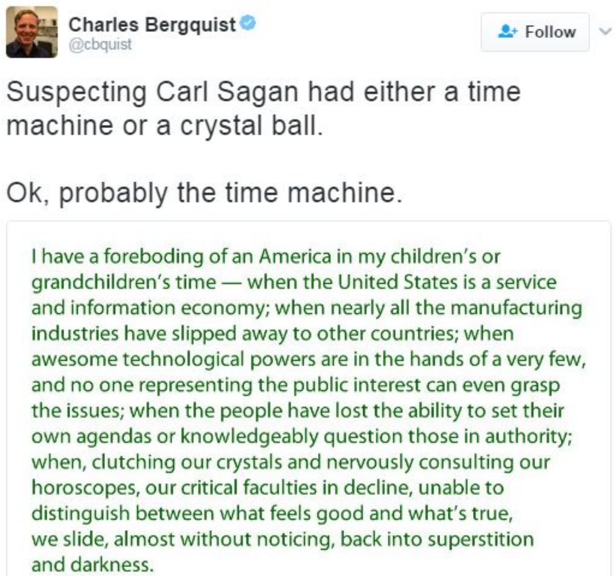
Contributors and Attributions

- Template:ContribWebLiteracyStudentFactcheckersCaulfield

11.7.5: Using Google Books to Track Down Quotes

Using Google Books to Track Down Quotes

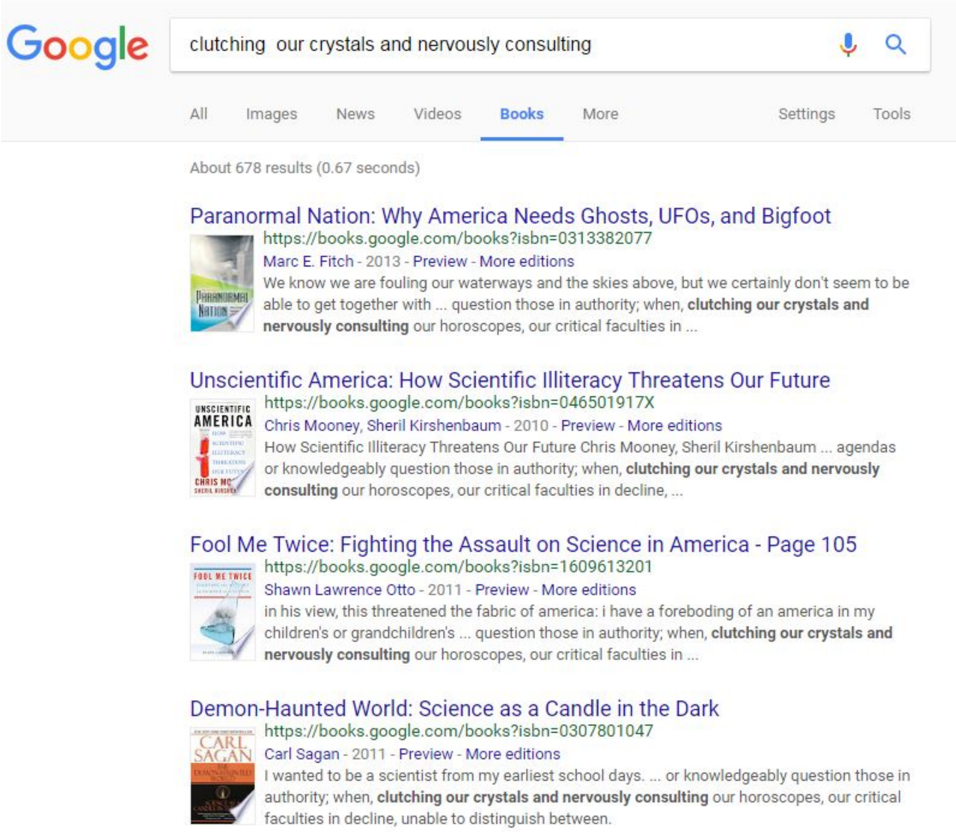
Did Carl Sagan say this?



Quotes are the internet are some of the most commonly faked content. People misattribute quotes to give them significance, or fabricate controversial quotes to create controversy. (For some examples of fact-checking historical quotes, check out Quote Investigator).

In our case, if we know that Carl Sagan is an author of many books, rather than start in Google or Duck Duck Go's general search we might start in Google Books, which will likely get us to the source of the quote faster. Additionally, even if we cannot find the source, we might find a someone quoting this in a book from a major publisher, which is likely to have a more developed fact-checking process than some guy on Twitter.

So we go to Google Books and we pick out just a short snippet of unique phrasing. I'm going to choose "clutching our crystals and nervously consulting":



Down there at the bottom, the fourth result, is a book by Carl Sagan. It says its from 2011, but don't be fooled by this date: this is just the date of the edition here indexed. Let's click through to the book to check the quote and then sort out the date later.

Clicking through the book we find the quote is accurate. More importantly we find the surrounding context and find that this quote is not being taken out of context. Sagan was truly worried about this, and his prediction was very much that a media obsessed with sound bites combined with a sort of celebration of ignorance would drag us backwards. Understanding the world was becoming more difficult at the same time the ways that understanding was communicated were becoming more shallow.

United States is a service and information economy; when nearly all the key manufacturing industries have slipped away to other countries; when awesome technological powers are in the hands of a very few, and no one representing the public interest can even grasp the issues; when the people have lost the ability to set their own agendas or knowledgeably question those in authority; when, clutching our crystals and nervously consulting our

Copy

horoscopes, our critical faculties in decline, unable to distinguish between what feels good and what's true, we slide, almost without noticing, back into superstition and darkness.

The dumbing down of America is most evident in the slow decay of substantive content in the enormously influential media, the 30-second sound bites (now down to 10 seconds or less), lowest common denominator programming, credulous presentations on pseudoscience and superstition, but especially a kind of celebration of ignorance. As I write,

You can find out the original publication date of this work a number of ways — there's a "more versions" option on the Google Books interface. You could go look for the book's article on Wikipedia, as they will usually give you the publication date. But the easiest way is usually to turn to the front pages of the book and find the date, just as you would with a physical book.

A Ballantine Book
Published by The Random House Publishing Group

Copyright © 1996 by Carl Sagan

All rights reserved.

Published in the United States by Ballantine Books, an imprint of The Random House Publishing Group, a division of Random House, Inc., New

Contributors and Attributions

- Template:ContribWebLiteracyStudentFactcheckersCaulfield

11.8: Getting Started in Chrome - Bookmarking in Chrome and more

https://edu.gcfglobal.org/en/chrome/bookmarking-in-chrome/1/

11.9: Introduction to Internet Safety

https://edu.gcfglobal.org/en/internetsafety/introduction-to-internet-safety/1/

11.10: Using Windows Defender Scan

To scan your PC for viruses and malware using Windows Defender Antivirus, complete either a quick scan or full scan. To perform a Quick Scan, complete the following:

1. Open Windows Defender Security Center.
2. Click on Virus & threat protection.
3. Click the Quick scan button.

To perform a Full Scan or Custom Scan, complete the following.

1. Open Windows Defender Security Center.
2. Click on Virus & threat protection.
3. Click the link of the Advanced scan.
4. Select the Full Scan or Custom Scan option and select the appropriate folders to scan.

Sometimes if you're dealing with a tough virus or another type of malware, the antivirus may not be able to remove it while Windows 10 is running. If this is the case, save your work, close any running application, and run an offline scan using the following steps:

1. Open Windows Defender Security Center.
2. Click on Virus & threat protection.
3. Click the link of the Advanced scan.
4. Select the Windows Defender Offline scan option.
5. Click the Scan now button.
6. Click the Scan button to continue.

When you start an offline scan, your device will restart and boot using the Windows Defender Antivirus environment. The latest definition update will be installed (if an internet connection is available), a scan will be performed, and if malicious software is detected, it'll be removed or quarantined automatically.

11.11: Internet and Computer Safety Tips

The following article was retrieved from usa.kapersky.com (Kapersky.com, n.d.)

A 19-year-old running for public office in New Hampshire found out about the importance of following Internet safety rules the hard way. As Seacoast Online reports, his opponents found images in his social media posts that were sexually suggestive and referenced past drug use. Just like that, his political career crashed and burned upon takeoff. But, unfortunately, he isn't the only one, as careless Internet habits have left others exposed to scams, identity theft and physical harm at the hands of people they met online. With more users accessing the Internet through mobile devices, these risks are changing and growing quickly. Even though apps loom larger in most people's daily online interactions than traditional websites do, that does not mean that the basic Internet safety rules have changed. Hackers are still on the lookout for personal information they can use to access your credit card and bank information. Unsafe surfing can also lead to other threats—from embarrassing personal comments or images that, once online, are nearly impossible to erase, to getting mixed up with people you'd rather have had nothing to do with.

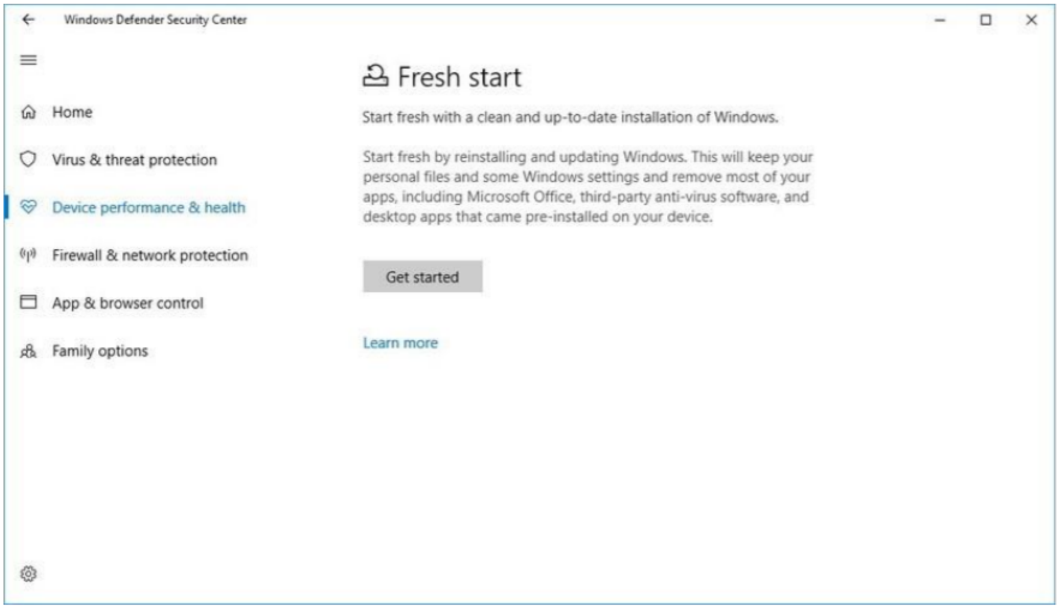
Here are the Top 10 Internet safety rules to help you avoid getting into trouble online (and offline).

1. Keep Personal Information Professional and Limited: Potential employers or customers don't need to know your personal relationship status or your home address. They do need to know about your expertise and professional background, and how to get in touch with you. You wouldn't hand purely personal information out to strangers individually—don't hand it out to millions of people online.
2. Keep Your Privacy Settings On: Marketers love to know all about you, and so do hackers. Both can learn a lot from your browsing and social media usage. But you can take charge of your information. As noted by Lifehacker, both web browsers and mobile operating systems have settings available to protect your privacy online. Major websites like Facebook also have privacyenhancing settings available. These settings are sometimes (deliberately) hard to find because companies want your personal information for its marketing value. Make sure you have enabled these privacy safeguards and keep them enabled.
3. Practice Safe Browsing: You wouldn't choose to walk through a dangerous neighborhood—don't visit dangerous neighborhoods online. Cybercriminals use lurid content as bait. They know people are sometimes tempted by dubious content and may let their guard down when searching for it. The Internet's demimonde is filled with hard-to-see pitfalls, where one careless click could expose personal data or infect your device with malware. By resisting the urge, you don't even give the hackers a chance.
4. Make Sure Your Internet Connection is Secure. Make sure your device is secure, and when in doubt, wait for a better time (i.e., until you're able to connect to a secure Wi-Fi network) before providing information such as your bank account number.
5. Be Careful What You Download: A top goal of cybercriminals is to trick you into downloading malware—programs or apps that carry malware or try to steal information. This malware can be disguised as an app: anything from a popular game to something that checks traffic or the weather. Don't download apps that look suspicious or come from a site you don't trust.
6. Choose Strong Passwords: Passwords are one of the biggest weak spots in the whole Internet security structure, but there's currently no way around them. And the problem with passwords is that people tend to choose easy ones to remember (such as "password" and "123456"), which are also easy for cyber thieves to guess. Select strong passwords that are harder for cybercriminals to demystify. Password manager software can help you to manage multiple passwords so that you don't forget them. A strong password is one that is unique and complex—at least 15 characters long, mixing letters, numbers and special characters.
7. Make Online Purchases from Secure Sites: Any time you make a purchase online, you need to provide credit card or bank account information—just what cybercriminals are most eager to get their hands on. Only supply this information to sites that provide secure, encrypted connections. As Boston University notes, you can identify secure sites by looking for an address that starts with https: (the S stands for secure) rather than simply Http: They may also be marked by a padlock icon next to the address bar.
8. Be Careful What You Post: The Internet does not have a delete key, as that young candidate in New Hampshire found out. Any comment or image you post online may stay online forever because removing the original (say, from Twitter) does not remove any copies that other people made. There is no way for you to "take back" a remark you wish you hadn't made or get rid of that embarrassing selfie you took at a party. Don't put anything online that you wouldn't want your mom or a prospective employer to see.
9. Be Careful Who You Meet Online: People you meet online are not always who they claim to be. Indeed, they may not even be real. Be as cautious and sensible in your online social life as you are in your in-person social life.
10. Keep Your Antivirus Program Up to Date: Internet security software cannot protect against every threat, but it will detect and remove most malware—though you should make sure it's to date. Be sure to stay current with your operating system's updates

and updates to the applications you use. They provide a vital layer of security.

Keep these 10 basic Internet safety rules in mind and you'll avoid many of the nasty surprises that lurk online for the careless

11.12: Fresh Start



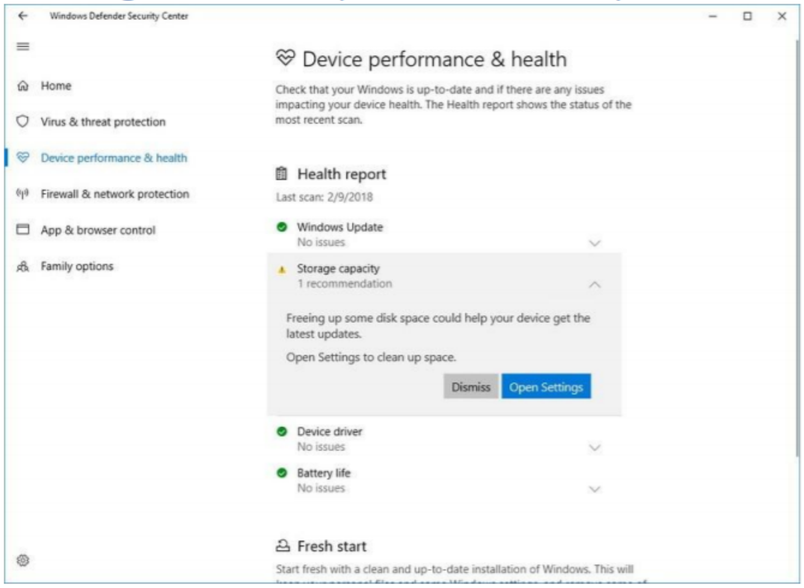
If your PC is having performance problems, such as issues with memory, shutting down, or everything is working very slow, you can click the Additional info link under "Fresh start" to re-install Windows 10 with the latest updates.

To perform a Fresh Start, perform the following:

1. Click the Get started button
2. Click the Next button.
3. Review the apps that will be removed.
4. Click the Next button.
5. Click the Start button to begin the process.

This process won't delete your files, it'll keep some Windows settings, but it'll remove everything else to help fix performance issues with your device (Windows Central, n.d.).

11.12.1: Viewing the Health/ Performance of your PC



Windows Defender Security Center also includes an area that surfaces information about the health and performance of your computer. To view your PC health and performance report, do the following:

1. Open Windows Defender Security Center.
2. Click on Device performance & health.

The report includes a status for Windows Update, storage, device driver, and battery. If action needs to be taken, you'll see an alert with a recommendation on how to remedy the issue (Windows Central, n.d.).

11.13: Automatic Updates

By default, Windows 10 is set up to automatically download the latest updates for the Windows Operating system as soon as they are available. It is important to allow those updates, as often crucial errors or "bugs" that may make your computer susceptible to a computer virus or malware are fixed or "patched" by updates. When updates are installing, it is important to not turn off your computer until they are completed.

Figure 11.13.1 This Photo by Unknown Author is licensed under under CC BY-SA-NC

11.14: Avast Academy – How to Open Incognito Mode in Chrome

https://www.avast.com/c-incognito-mode-chrome

11.15: Avast Academy – How to Clear Your Browser History

https://www.avast.com/c-how-to-clear-browser-history

11.16: Avast Academy – What are cookies and how to clear them

https://www.avast.com/c-delete-browser-cookies

11.17: Email Basics

https://edu.gcfglobal.org/en/email101/

11.18: INTERNET PRIVACY, COMPUTER SECURITY AND NETIQUETTE

11.19: INTERNET PRIVACY, COMPUTER SECURITY AND NETIQUETTE

CHAPTER OVERVIEW

Front Matter

Table of Contents

- 12.10.4: Programming Languages
    - 12.10.4.1: Language Generations
    - 12.10.4.2: Compiled vs Interpreted
    - 12.10.4.3: Procedural / Object Oriented
  - 12.10.5: IDE / CASE
  - 12.10.6: Home Grown or Purchased
  - 12.10.7: Web Service
  - 12.10.8: End User Dev
  - 12.10.9: Project Implementation
  - 12.10.10: Summary
- 12.11: Globalization and the Digital Divide
  Globalization is the term used to refer to the integration of goods, services, and culture among the nations of the world. Globalization is not necessarily a new phenomenon; in many ways, we have been experiencing globalization since the days of European colonization. Further advances in telecommunication and transportation technologies accelerated globalization. The advent of the the worldwide Internet has made all nations next-door neighbors.

  - 12.11.1: Globalization?
    - 12.11.1.1: The World...
  - 12.11.2: Going Global
  - 12.11.3: Digital Divide
  - 12.11.4: OLPC
  - 12.11.5: Changing the Divide
  - 12.11.6: Summary
- 12.12: The Ethical and Legal Implications of Information Systems
  New technologies create new situations that we have never dealt with before. How do we handle the new capabilities that these devices empower us with? What new laws are going to be needed to protect us from ourselves? This chapter will kick off with a discussion of the impact of information systems on how we behave (ethics). This will be followed with the new legal structures being put in place, with a focus on intellectual property and privacy.

  - 12.12.1: Ethics
    - 12.12.1.1: Code of Ethics
  - 12.12.2: Acceptable Use
  - 12.12.3: Intellectual Property
    - 12.12.3.1: Copyright
      - 12.12.3.1.1: Copyright details
      - 12.12.3.1.2: Digital Millennium Copyright
    - 12.12.3.2: Patent
      - 12.12.3.2.1: Obtain Protection
    - 12.12.3.3: Trademark
  - 12.12.4: Privacy
    - 12.12.4.1: Other Privacy Laws
    - 12.12.4.2: Non-Obvious Relationship
  - 12.12.5: Section 2 Summary
- 12.13: Future Trends in Information Systems
  Information systems have evolved at a rapid pace ever since their introduction in the 1950s. Today, devices that we can hold in one hand are more powerful than the computers used to land a man on the moon. The Internet has made the entire world accessible to us, allowing us to communicate and collaborate with each other like never before.

  - 12.13.1: Global
  - 12.13.2: Social and Personal

SECTION OVERVIEW

12.1: Introduction to Information Systems?

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- define what an information system is by identifying its major components;
- describe the basic history of information systems; and
- describe the basic argument behind the article "Does IT Matter?" by Nicholas Carr.

Introduction

If you are reading this, you are most likely taking a course in information systems, but do you even know what the course is going to cover? When you tell your friends or your family that you are taking a course in information systems, can you explain what it is about? For the past several years, I have taught an Introduction to Information Systems course. The first day of class I ask my students to tell me what they think an information system is. I generally get answers such as "computers," "databases," or "Excel." These are good answers, but definitely incomplete ones. The study of information systems goes far beyond understanding some technologies. Let's begin our study by defining information systems.

Defining Information Systems

Almost all programs in business require students to take a course in something called *information systems*. But what exactly does that term mean? Let's take a look at some of the more popular definitions, first from Wikipedia and then from a couple of textbooks:

- "Information systems (IS) is the study of complementary networks of hardware and software that people and organizations use to collect, filter, process, create, and distribute data."[1]
- "Information systems are combinations of hardware, software, and telecommunications networks that people build and use to collect, create, and distribute useful data, typically in organizational settings."[2]
- "Information systems are interrelated components working together to collect, process, store, and disseminate information to support decision making, coordination, control, analysis, and viualization in an organization."[3]

As you can see, these definitions focus on two different ways of describing information systems: the *components* that make up an information system and the *role* that those components play in an organization. Let's take a look at each of these.

12.1.1: Components of an Information System

The Components of Information Systems

As I stated earlier, I spend the first day of my information systems class discussing exactly what the term means. Many students understand that an information system has something to do with databases or spreadsheets. Others mention computers and e-commerce. And they are all right, at least in part: information systems are made up of different components that work together to provide value to an organization.

The first way I describe information systems to students is to tell them that they are made up of five components: hardware, software, data, people, and process. The first three, fitting under the category *technology* , are generally what most students think of when asked to define information systems. But the last two, people and process, are really what separate the idea of information systems from more technical fields, such as computer science. In order to fully understand information systems, students must understand how all of these components work together to bring value to an organization.

Technology

Technology can be thought of as the application of scientific knowledge for practical purposes. From the invention of the wheel to the harnessing of electricity for artificial lighting, technology is a part of our lives in so many ways that we tend to take it for granted. As discussed before, the first three components of information systems – hardware, software, and data – all fall under the category of technology. Each of these will get its own chapter and a much lengthier discussion, but we will take a moment here to introduce them so we can get a full understanding of what an information system is.

Hardware

Information systems hardware is the part of an information system you can touch – the physical components of the technology. Computers, keyboards, disk drives, iPads, and flash drives are all examples of information systems hardware. We will spend some time going over these components and how they all work together in chapter 2.

Software

Software is a set of instructions that tells the hardware what to do. Software is not tangible – it cannot be touched. When programmers create software programs, what they are really doing is simply typing out lists of instructions that tell the hardware what to do. There are several categories of software, with the two main categories being operating-system software, which makes the hardware usable, and application software, which does something useful. Examples of operating systems include Microsoft Windows on a personal computer and Google's Android on a mobile phone. Examples of application software are Microsoft Excel and Angry Birds. Software will be explored more thoroughly in chapter 3.

Data

The third component is data. You can think of data as a collection of facts. For example, your street address, the city you live in, and your phone number are all pieces of data. Like software, data is also intangible. By themselves, pieces of data are not really very useful. But aggregated, indexed, and organized together into a database, data can become a powerful tool for businesses. In fact, all of the definitions presented at the beginning of this chapter focused on how information systems manage data. Organizations collect all kinds of data and use it to make decisions. These decisions can then be analyzed as to their effectiveness and the organization can be improved. Chapter 4 will focus on data and databases, and their uses in organizations.

Networking Communication: A Fourth Technology Piece?

Besides the components of hardware, software, and data, which have long been considered the core technology of information systems, it has been suggested that one other component should be added: communication. An information system can exist without the ability to communicate – the first personal computers were stand-alone machines that did not access the Internet. However, in today's hyper-connected world, it is an extremely rare computer that does not connect to another device or to a network. Technically, the networking communication component is made up of hardware and software, but it is such a core feature of today's information systems that it has become its own category. We will be covering networking in chapter 5.

People

When thinking about information systems, it is easy to get focused on the technology components and forget that we must look beyond these tools to fully understand how they integrate into an organization. A focus on the people involved in information systems is the next step. From the front-line help-desk workers, to systems analysts, to programmers, all the way up to the chief information officer (CIO), the people involved with information systems are an essential element that must not be overlooked. The people component will be covered in chapter 9.

Process

The last component of information systems is process. A process is a series of steps undertaken to achieve a desired outcome or goal. Information systems are becoming more and more integrated with organizational processes, bringing more productivity and better control to those processes. But simply automating activities using technology is not enough – businesses looking to effectively utilize information systems do more. Using technology to manage and improve processes, both within a company and externally with suppliers and customers, is the ultimate goal. Technology buzzwords such as "business process reengineering," "business process management," and "enterprise resource planning" all have to do with the continued improvement of these business procedures and the integration of technology with them. Businesses hoping to gain an advantage over their competitors are highly focused on this component of information systems. We will discuss processes in chapter 8.

12.1.2: The Role of Information Systems

The Role of Information Systems

Now that we have explored the different components of information systems, we need to turn our attention to the role that information systems play in an organization. So far we have looked at what the components of an information system are, but what do these components actually do for an organization? From our definitions above, we see that these components collect, store, organize, and distribute data throughout the organization. In fact, we might say that one of the roles of information systems is to take data and turn it into information, and then transform that into organizational knowledge. As technology has developed, this role has evolved into the backbone of the organization. To get a full appreciation of the role information systems play, we will review how they have changed over the years.

IBM 704 Mainframe (Copyright: Lawrence Livermore National Laboratory)

The Mainframe Era

From the late 1950s through the 1960s, computers were seen as a way to more efficiently do calculations. These first business computers were room-sized monsters, with several refrigerator-sized machines linked together. The primary work of these devices was to organize and store large volumes of information that were tedious to manage by hand. Only large businesses, universities, and government agencies could afford them, and they took a crew of specialized personnel and specialized facilities to maintain. These devices served dozens to hundreds of users at a time through a process called time-sharing. Typical functions included scientific calculations and accounting, under the broader umbrella of “data processing.”

Registered trademark of International Business Machines

In the late 1960s, the Manufacturing Resources Planning (MRP) systems were introduced. This software, running on a mainframe computer, gave companies the ability to manage the manufacturing process, making it more efficient. From tracking inventory to creating bills of materials to scheduling production, the MRP systems (and later the MRP II systems) gave more businesses a reason to want to integrate computing into their processes. IBM became the dominant mainframe company. Nicknamed “Big Blue,” the company became synonymous with business computing. Continued improvement in software and the availability of cheaper hardware eventually brought mainframe computers (and their little sibling, the minicomputer) into most large businesses.

The PC Revolution

In 1975, the first microcomputer was announced on the cover of *Popular Mechanics*: the Altair 8800. Its immediate popularity sparked the imagination of entrepreneurs everywhere, and there were quickly dozens of companies making these “personal computers.” Though at first just a niche product for computer hobbyists, improvements in usability and the availability of practical software led to growing sales. The most prominent of these early personal computer makers was a little company known as Apple Computer, headed by Steve Jobs and Steve Wozniak, with the hugely successful “Apple II.” Not wanting to be left out of the revolution, in 1981 IBM (teaming with a little company called Microsoft for their operating-system software) hurriedly released their own version of the personal computer, simply called the “PC.” Businesses, who had used IBM mainframes for years to run their businesses, finally had the permission they needed to bring personal computers into their companies, and the IBM PC took off. The IBM PC was named *Time* magazine's "Man of the Year" for 1982.

Because of the IBM PC's open architecture, it was easy for other companies to copy, or "clone" it. During the 1980s, many new computer companies sprang up, offering less expensive versions of the PC. This drove prices down and spurred innovation. Microsoft developed itsWindows operating system and made the PC even easier to use. Common uses for the PC during this period included word processing, spreadsheets, and databases. These early PCs were not connected to any sort of network; for the most part they stood alone as islands of innovation within the larger organization.

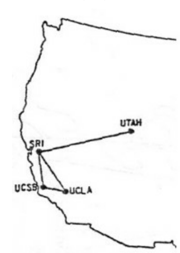
Client-Server

In the mid-1980s, businesses began to see the need to connect their computers together as a way to collaborate and share resources. This networking architecture was referred to as "client-server" because users would log in to the local area network (LAN) from their PC (the "client") by connecting to a powerful computer called a "server," which would then grant them rights to different resources on the network (such as shared file areas and a printer). Software companies began developing applications that allowed multiple users to access the same data at the same time. This evolved into software applications for communicating, with the first real popular use of electronic mail appearing at this time.

This networking and data sharing all stayed within the confines of each business, for the most part. While there was sharing of electronic data between companies, this was a very specialized function. Computers were now seen as tools to collaborate internally, within an organization. In fact, these networks of computers were becoming so powerful that they were replacing many of the functions previously performed by the larger mainframe computers at a fraction of the cost. It was during this era that the first Enterprise Resource Planning (ERP) systems were developed and run on the client-server architecture. An ERP system is a software application with a centralized database that can be used to run a company's entire business. With separate modules for accounting, finance, inventory, human resources, and many, many more, ERP systems, with Germany's SAP leading the way, represented the state of the art in information systems integration. We will discuss ERP systems as part of the chapter on process (chapter 9).

The World Wide Web and E-Commerce

First invented in 1969, the Internet was confined to use by universities, government agencies, and researchers for many years. Its ratherarcane commands and user applications made it unsuitable for mainstream use in business. One exception to this was the ability to expand electronic mail outside the confines of a single organization. While the first e-mail messages on the Internet were sent in the early 1970s, companies who wanted to expand their LAN-based e-mail started hooking up to the Internet in the 1980s. Companies began connecting their internal networks to the Internet in order to allow communication between their employees and employees at other companies. It was with these early Internet connections that the computer truly began to evolve from a computational device to a communications device.



In 1989, Tim Berners-Lee developed a simpler way for researchers to share information over the network at CERN laboratories, a concept he called the World Wide Web.[4] This invention became the launching point of the growth of the Internet as a way for businesses to share information about themselves. As web browsers and Internet connections became the norm, companies rushed to grab domain names and create websites.

Registered trademark of Amazon Technologies, Inc.

In 1991, the National Science Foundation, which governed how the Internet was used, lifted restrictions on its commercial use. The year 1994 saw the establishment of both eBay and Amazon.com, two true pioneers in the use of the new digital marketplace. A mad rush of investment in Internet-based businesses led to the dot-com boom through the late 1990s, and then the dot-com bust in 2000. While much can be learned from the speculation and crazy economic theories espoused during that bubble, one important outcome for businesses was that thousands of miles of Internet connections were laid around the world during that time. The world became truly “wired” heading into the new millenium, ushering in the era of globalization, which we will discuss in chapter 11.

As it became more expected for companies to be connected to the Internet, the digital world also became a more dangerous place. Computer viruses and worms, once slowly propagated through the sharing of computer disks, could now grow with tremendous speed via the Internet. Software written for a disconnected world found it very difficult to defend against these sorts of threats. A whole new industry of computer and Internet security arose. We will study information security in chapter 6.

Web 2.0

As the world recovered from the dot-com bust, the use of technology in business continued to evolve at a frantic pace. Websites became interactive; instead of just visiting a site to find out about a business and purchase its products, customers wanted to be able to customize their experience and interact with the business. This new type of interactive website, where you did not have to know how to create a web page or do any programming in order to put information online, became known as web 2.0. Web 2.0 is exemplified by blogging, social networking, and interactive comments being available on many websites. This new web-2.0 world, in which online interaction became expected, had a big impact on many businesses and even whole industries. Some industries, such as bookstores, found themselves relegated to a niche status. Others, such as video rental chains and travel agencies, simply began going out of business as they were replaced by online technologies. This process of technology replacing a middleman in a transaction is called disintermediation.

As the world became more connected, new questions arose. Should access to the Internet be considered a right? Can I copy a song that I downloaded from the Internet? How can I keep information that I have put on a website private? What information is acceptable to collect from children? Technology moved so fast that policymakers did not have enough time to enact appropriate laws, making for a Wild West–type atmosphere. Ethical issues surrounding information systems will be covered in chapter 12.

The Post-PC World

After thirty years as the primary computing device used in most businesses, sales of the PC are now beginning to decline as sales of tablets and smartphones are taking off. Just as the mainframe before it, the PC will continue to play a key role in business, but will no longer be the primary way that people interact and do business. The limited storage and processing power of these devices is being offset by a move to “cloud” computing, which allows for storage, sharing, and backup of information on a massive scale. This will require new rounds of thinking and innovation on the part of businesses as technology continues to advance.

The Eras of Business Computing

Era	Hardware	Operating System	Applications
Mainframe (1970s)	Terminals connected to mainframe computer.	Time-sharing (TSO) on MVS	Custom-written MRP software
PC (mid-1980s)	IBM PC or compatible. Sometimes connected to mainframe computer via expansion card.	MS-DOS	WordPerfect, Lotus 1-2-3
Client-Server (late 80s to early 90s)	IBM PC “clone” on a Novell Network.	Windows for Workgroups	Microsoft Word, Microsoft Excel
World Wide Web (mid-90s to early 2000s)	IBM PC “clone” connected to company intranet.	Windows XP	Microsoft Office, Internet Explorer
Web 2.0 (mid-2000s to present)	Laptop connected to company Wi-Fi.	Windows 7	Microsoft Office, Firefox

Era	Hardware	Operating System	Applications
Post-PC (today and beyond)	Apple iPad	iOS	Mobile-friendly websites, mobile apps

12.1.3: Competitive Advantage

Can Information Systems Bring Competitive Advantage?

It has always been the assumption that the implementation of information systems will, in and of itself, bring a business competitive advantage. After all, if installing one computer to manage inventory can make a company more efficient, won't installing several computers to handle even more of the business continue to improve it?

In 2003, Nicholas Carr wrote an article in the *Harvard Business Review* that questioned this assumption. The article, entitled "IT Doesn't Matter," raised the idea that information technology has become just a commodity. Instead of viewing technology as an investment that will make a company stand out, it should be seen as something like electricity: It should be managed to reduce costs, ensure that it is always running, and be as risk-free as possible.

As you might imagine, this article was both hailed and scorned. Can IT bring a competitive advantage? It sure did for Walmart (see sidebar). We will discuss this topic further in chapter 7.

Sidebar: Walmart Uses Information Systems to Become the World's Leading Retailer

File:Walmart_exteriorcropped.jpg)

Registered trademark of Wal-Mart Stores, Inc.

Walmart is the world's largest retailer, earning $15.2 billion on sales of $443.9 billion in the fiscal year that ended on January 31, 2012. Walmart currently serves over 200 million customers every week, worldwide.[5] Walmart's rise to prominence is due in no small part to their use of information systems.

One of the keys to this success was the implementation of Retail Link, a supply-chain management system. This system, unique when initially implemented in the mid-1980s, allowed Walmart's suppliers to directly access the inventory levels and sales information of their products at any of Walmart's more than ten thousand stores. Using Retail Link, suppliers can analyze how well their products are selling at one or more Walmart stores, with a range of reporting options. Further, Walmart requires the suppliers to use Retail Link to manage their own inventory levels. If a supplier feels that their products are selling out too quickly, they can use Retail Link to petition Walmart to raise the levels of inventory for their products. This has essentially allowed Walmart to "hire" thousands of product managers, all of whom have a vested interest in the products they are managing. This revolutionary approach to managing inventory has allowed Walmart to continue to drive prices down and respond to market forces quickly.

Today, Walmart continues to innovate with information technology. Using its tremendous market presence, any technology that Walmart requires its suppliers to implement immediately becomes a business standard.

12.1.4: Section #1 Summary

Summary

In this chapter, you have been introduced to the concept of information systems. We have reviewed several definitions, with a focus on the components of information systems: technology, people, and process. We have reviewed how the business use of information systems has evolved over the years, from the use of large mainframe computers for number crunching, through the introduction of the PC and networks, all the way to the era of mobile computing. During each of these phases, new innovations in software and technology allowed businesses to integrate technology more deeply.

We are now to a point where every company is using information systems and asking the question: Does it bring a competitive advantage? In the end, that is really what this book is about. Every businessperson should understand what an information system is and how it can be used to bring a competitive advantage. And that is the task we have before us.

Study Questions

1. What are the five components that make up an information system?
2. What are three examples of information system hardware?
3. Microsoft Windows is an example of which component of information systems?
4. What is application software?
5. What roles do people play in information systems?
6. What is the definition of a process?
7. What was invented first, the personal computer or the Internet (ARPANET)?
8. In what year were restrictions on commercial use of the Internet first lifted? When were eBay and Amazon founded?
9. What does it mean to say we are in a “post-PC world”?
10. What is Carr’s main argument about information technology?

Exercises

1. Suppose that you had to explain to a member of your family or one of your closest friends the concept of an information system. How would you define it? Write a one-paragraph description *in your own words* that you feel would best describe an information system to your friends or family.
2. Of the five primary components of an information system (hardware, software, data, people, process), which do you think is the most important to the success of a business organization? Write a one-paragraph answer to this question that includes an example from your personal experience to support your answer.
3. We all interact with various information systems every day: at the grocery store, at work, at school, even in our cars (at least some of us). Make a list of the different information systems you interact with every day. See if you can identify the technologies, people, and processes involved in making these systems work.
4. Do you agree that we are in a post-PC stage in the evolution of information systems? Some people argue that we will always need the personal computer, but that it will not be the primary device used for manipulating information. Others think that a whole new era of mobile and biological computing is coming. Do some original research and make your prediction about what business computing will look like in the next generation.
5. The Walmart case study introduced you to how that company used information systems to become the world’s leading retailer. Walmart has continued to innovate and is still looked to as a leader in the use of technology. Do some original research and write a one-page report detailing a new technology that Walmart has recently implemented or is pioneering.

---

1. Wikipedia entry on "Information Systems," as displayed on August 19, 2012. *Wikipedia: The Free Encyclopedia*. San Francisco: Wikimedia Foundation. http://en.wikipedia.org/wiki/Informa...s_(discipline).↵
2. Excerpted from *Information Systems Today - Managing in the Digital World*, fourth edition. Prentice-Hall, 2010.↵
3. Excerpted from *Management Information Systems*, twelfth edition, Prentice-Hall, 2012.↵
4. CERN's "The Birth of the Web." http://public.web.cern.ch/public/en/about/web-en.html↵
5. Walmart 2012 Annual Report.↵

SECTION OVERVIEW

12.2: Hardware

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe information systems hardware;
- identify the primary components of a computer and the functions they perform; and
- explain the effect of the commoditization of the personal computer.

Introduction

As we learned in the first chapter, an information system is made up of five components: hardware, software, data, people, and process. The physical parts of computing devices – those that you can actually touch – are referred to as hardware. In this chapter, we will take a look at this component of information systems, learn a little bit about how it works, and discuss some of the current trends surrounding it.

As stated above, computer hardware encompasses digital devices that you can physically touch. This includes devices such as the following:

- desktop computers
- laptop computers
- mobile phones
- tablet computers
- e-readers
- storage devices, such as flash drives
- input devices, such as keyboards, mice, and scanners
- output devices such as printers and speakers.

Besides these more traditional computer hardware devices, many items that were once not considered digital devices are now becoming computerized themselves. Digital technologies are now being integrated into many everyday objects, so the days of a device being labeled categorically as computer hardware may be ending. Examples of these types of digital devices include automobiles, refrigerators, and even soft-drink dispensers. In this chapter, we will also explore digital devices, beginning with defining what we mean by the term itself.

Digital Devices

A digital device processes electronic signals that represent either a one ("on") or a zero ("off"). The *on* state is represented by the presence of an electronic signal; the *off* state is represented by the absence of an electronic signal. Each one or zero is referred to as a *bit* (a contraction of *binary digit*); a group of eight bits is a *byte*. The first personal computers could process 8 bits of data at once; modern PCs can now process 64 bits of data at a time, which is where the term *64-bit processor* comes from.

Sidebar: Understanding Binary

As you know, the system of numbering we are most familiar with is base-ten numbering. In base-ten numbering, each column in the number represents a power of ten, with the far-right column representing 10^0 (ones), the next column from the right representing 10^1 (tens), then 10^2 (hundreds), then 10^3 (thousands), etc. For example, the number 1010 in decimal represents: (1 x 1000) + (0 x 100) + (1 x 10) + (0 x 1).

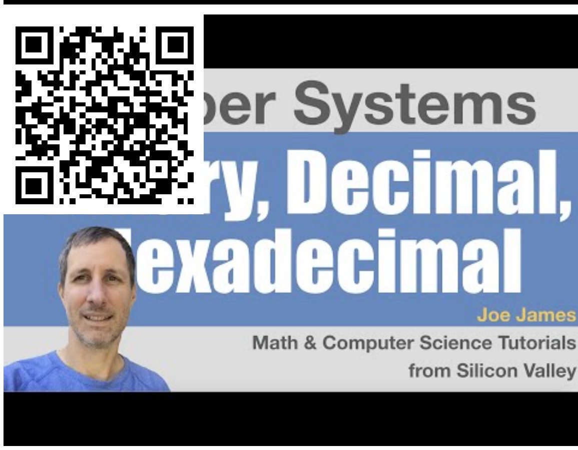
Computers use the base-two numbering system, also known as binary. In this system, each column in the number represents a power of two, with the far-right column representing 2^0 (ones), the next column from the right representing 2^1 (tens), then 2^2 (fours), then 2^3 (eights), etc. For example, the number 1010 in binary represents (1 x 8) + (0 x 4) + (1 x 2) + (0 x 1). In base ten, this evaluates to 10.

As the capacities of digital devices grew, new terms were developed to identify the capacities of processors, memory, and disk storage space. Prefixes were applied to the word *byte* to represent different orders of magnitude. Since these are digital specifications, the prefixes were originally meant to represent multiples of 1024 (which is 2^{10}), but have more recently been rounded to mean multiples of 1000.

A Listing of Binary Prefixes

Prefix	Represents	Example
kilo	one thousand	kilobyte=one thousand bytes
mega	one million	megabyte=one million bytes
giga	one billion	gigabyte=one billion bytes
tera	one trillion	terabyte=one trillion bytes

Here is a short video that explains numbering systems - decimal, binary and hexidecimal: https://www.youtube.com/watch?v=aW3qCcH6Dao



12.2.1: Hardware Components

12.2.2: More Computing Devices

12.2.3: Computers are part of life

12.2.4: Summary

12.2.1: Hardware Components

Tour of a PC

All personal computers consist of the same basic components: a CPU, memory, circuit board, storage, and input/output devices. It also turns out that almost every digital device uses the same set of components, so examining the personal computer will give us insight into the structure of a variety of digital devices. So let's take a "tour" of a personal computer and see what makes them function.

Processing Data: The CPU

As stated above, most computing devices have a similar architecture. The core of this architecture is the central processing unit, or CPU. The CPU can be thought of as the "brains" of the device. The CPU carries out the commands sent to it by the software and returns results to be acted upon.

The earliest CPUs were large circuit boards with limited functionality. Today, a CPU is generally on one chip and can perform a large variety of functions. There are two primary manufacturers of CPUs for personal computers: Intel and Advanced Micro Devices (AMD).

The speed ("clock time") of a CPU is measured in hertz. A hertz is defined as one cycle per second. Using the binary prefixes mentioned above, we can see that a kilohertz (abbreviated kHz) is one thousand cycles per second, a megahertz (mHz) is one million cycles per second, and a gigahertz (gHz) is one billion cycles per second. The CPU's processing power is increasing at an amazing rate (see the sidebar about Moore's Law). Besides a faster clock time, many CPU chips now contain multiple processors per chip. These chips, known as dual-core (two processors) or quad-core (four processors), increase the processing power of a computer by providing the capability of multiple CPUs.

Sidebar: Moore's Law

We all know that computers get faster every year. Many times, we are not sure if we want to buy today's model of smartphone, tablet, or PC because next week it won't be the most advanced any more. Gordon Moore, one of the founders of Intel, recognized this phenomenon in 1965, noting that microprocessor transistor counts had been doubling every year.[1] His insight eventually evolved into Moore's Law, which states that the number of transistors on a chip will double every two years. This has been generalized into the concept that computing power will double every two years for the same price point. Another way of looking at this is to think that the price for the same computing power will be cut in half every two years. Though many have predicted its demise, Moore's Law has held true for over forty years (see figure below).

A graphical representation of Moore's Law (CC-BY-SA: Wgsimon)

There will be a point, someday, where we reach the limits of Moore's Law, where we cannot continue to shrink circuits any further. But engineers will continue to seek ways to increase performance.

Motherboard

Motherboard (click image to enlarge)

The motherboard is the main circuit board on the computer. The CPU, memory, and storage components, among other things, all connect into the motherboard. Motherboards come in different shapes and sizes, depending upon how compact or expandable the computer is designed to be. Most modern motherboards have many integrated components, such as video and sound processing, which used to require separate components.

The motherboard provides much of the bus of the computer (the term *bus* refers to the electrical connection between different computer components). The bus is an important determiner of the computer's speed: the combination of how fast the bus can transfer data and the number of data bits that can be moved at one time determine the speed.

Random-Access Memory

When a computer starts up, it begins to load information from the hard disk into its working memory. This working memory, called random-access memory (RAM), can transfer data much faster than the hard disk. Any program that you are running on the computer is loaded into RAM for processing. In order for a computer to work effectively, some minimal amount of RAM must be installed. In most cases, adding more RAM will allow the computer to run faster. Another characteristic of RAM is that it is "volatile." This means that it can store data as long as it is receiving power; when the computer is turned off, any data stored in RAM is lost.

Memory DIMM (click image to enlarge)

RAM is generally installed in a personal computer through the use of a dual-inline memory module (DIMM). The type of DIMM accepted into a computer is dependent upon the motherboard. As described by Moore's Law, the amount of memory and speeds of DIMMs have increased dramatically over the years.

Hard Disk

Hard disk enclosure (click image to enlarge)

While the RAM is used as working memory, the computer also needs a place to store data for the longer term. Most of today's personal computers use a hard disk for long-term data storage. A hard disk is where data is stored when the computer is turned off and where it is retrieved from when the computer is turned on. Why is it called a hard disk? A hard disk consists of a stack of disks inside a hard metal case. A floppy disk (discussed below) was a removable disk that, in some cases at least, was flexible, or "floppy."

Solid-State Drives

A relatively new component becoming more common in some personal computers is the solid-state drive (SSD). The SSD performs the same function as a hard disk: long-term storage. Instead of spinning disks, the SSD uses flash memory, which is much faster.

Solid-state drives are currently quite a bit more expensive than hard disks. However, the use of flash memory instead of disks makes them much lighter and faster than hard disks. SSDs are primarily utilized in portable computers, making them lighter and more efficient. Some computers combine the two storage technologies, using the SSD for the most accessed data (such as the operating system) while using the hard disk for data that is accessed less frequently. As with any technology, Moore's Law is driving up capacity and speed and lowering prices of solid-state drives, which will allow them to proliferate in the years to come.

Removable Media

Besides fixed storage components, removable storage media are also used in most personal computers. Removable media allows you to take your data with you. And just as with all other digital technologies, these media have gotten smaller and more powerful as the years have gone by. Early computers used floppy disks, which could be inserted into a disk drive in the computer. Data was stored on a magnetic disk inside an enclosure. These disks ranged from 8″ in the earliest days down to 3 1/2″.

Floppy-disk evolution (8″ to 5 1/4″ to 3 1/2″) (Public Domain)

Around the turn of the century, a new portable storage technology was being developed: the USB flash drive (more about the USB port later in the chapter). This device attaches to the universal serial bus (USB) connector, which became standard on all personal computers beginning in the late 1990s. As with all other storage media, flash drive storage capacity has skyrocketed over the years, from initial capacities of eight megabytes to current capacities of 64 gigabytes and still growing.

Network Connection

When personal computers were first developed, they were stand-alone units, which meant that data was brought into the computer or removed from the computer via removable media, such as the floppy disk. Beginning in the mid-1980s, however, organizations began to see the value in connecting computers together via a digital network. Because of this, personal computers needed the ability to connect to these networks. Initially, this was done by adding an expansion card to the computer that enabled the network connection, but by the mid-1990s, a network port was standard on most personal computers. As wireless technologies began to dominate in the early 2000s, many personal computers also began including wireless networking capabilities. Digital communication technologies will be discussed further in chapter 5.

Input and Output

USB connector (click image to enlarge)

In order for a personal computer to be useful, it must have channels for receiving input from the user and channels for delivering output to the user. These input and output devices connect to the computer via various connection ports, which generally are part of the motherboard and are accessible outside the computer case. In early personal computers, specific ports were designed for each type of output device. The configuration of these ports has evolved over the years, becoming more and more standardized over time. Today, almost all devices plug into a computer through the use of a USB port. This port type, first introduced in 1996, has increased in its capabilities, both in its data transfer rate and power supplied.

Bluetooth

Besides USB, some input and output devices connect to the computer via a wireless-technology standard called Bluetooth. Bluetooth was first invented in the 1990s and exchanges data over short distances using radio waves. Bluetooth generally has a range of 100 to 150 feet. For devices to communicate via Bluetooth, both the personal computer and the connecting device must have a Bluetooth communication chip installed.

Input Devices

All personal computers need components that allow the user to input data. Early computers used simply a keyboard to allow the user to enter data or select an item from a menu to run a program. With the advent of the graphical user interface, the mouse became a standard component of a computer. These two components are still the primary input devices to a personal computer, though variations of each have been introduced with varying levels of success over the years. For example, many new devices now use a touch screen as the primary way of entering data.

Besides the keyboard and mouse, additional input devices are becoming more common. Scanners allow users to input documents into a computer, either as images or as text. Microphones can be used to record audio or give voice commands. Webcams and other types of video cameras can be used to record video or participate in a video chat session.

Output Devices

Output devices are essential as well. The most obvious output device is a display, visually representing the state of the computer. In some cases, a personal computer can support multiple displays or be connected to larger-format displays such as a projector or large-screen television. Besides displays, other output devices include speakers for audio output and printers for printed output.

Sidebar: What Hardware Components Contribute to the Speed of My Computer?

The speed of a computer is determined by many elements, some related to hardware and some related to software. In hardware, speed is improved by giving the electrons shorter distances to traverse to complete a circuit. Since the first CPU was created in the early 1970s, engineers have constantly worked to figure out how to shrink these circuits and put more and more circuits onto the same chip. And this work has paid off – the speed of computing devices has been continuously improving ever since.

The hardware components that contribute to the speed of a personal computer are the CPU, the motherboard, RAM, and the hard disk. In most cases, these items can be replaced with newer, faster components. In the case of RAM, simply adding more RAM can also speed up the computer. The table below shows how each of these contributes to the speed of a computer. Besides upgrading hardware, there are many changes that can be made to the software of a computer to make it faster.

How Hardware Components Contribute to Computer Speed

Component	Speed measured by	Units	Description
CPU	Clock speed	gHz	The time it takes to complete a circuit.
Motherboard	Bus speed	mHz	How much data can move across the bus simultaneously.
RAM	Data transfer rate	MB/s	The time it takes for data to be transferred from memory to system.
Hard Disk	Access time	ms	The time it takes before the disk can transfer data.
	Data transfer rate	MBit/s	The time it takes for data to be transferred from disk to system.

12.2.2: More Computing Devices

Other Computing Devices

A personal computer is designed to be a general-purpose device. That is, it can be used to solve many different types of problems. As the technologies of the personal computer have become more commonplace, many of the components have been integrated into other devices that previously were purely mechanical. We have also seen an evolution in what defines a computer. Ever since the invention of the personal computer, users have clamored for a way to carry them around. Here we will examine several types of devices that represent the latest trends in personal computing.

Portable Computers

A modern laptop

In 1983, Compaq Computer Corporation developed the first commercially successful portable personal computer. By today's standards, the Compaq PC was not very portable: weighing in at 28 pounds, this computer was portable only in the most literal sense – it could be carried around. But this was no laptop; the computer was designed like a suitcase, to be lugged around and laid on its side to be used. Besides portability, the Compaq was successful because it was fully compatible with the software being run by the IBM PC, which was the standard for business.

In the years that followed, portable computing continued to improve, giving us laptop and notebook computers. The "luggable" computer has given way to a much lighter clamshell computer that weighs from 4 to 6 pounds and runs on batteries. In fact, the most recent advances in technology give us a new class of laptop that is quickly becoming the standard: these laptops are extremely light and portable and use less power than their larger counterparts. The MacBook Air is a good example of this: it weighs less than three pounds and is only 0.68 inches thick!

Finally, as more and more organizations and individuals are moving much of their computing to the Internet, laptops are being developed that use "the cloud" for all of their data and application storage. These laptops are also extremely light because they have no need of a hard disk at all! A good example of this type of laptop (sometimes called a netbook) is Samsung's Chromebook.

Smartphones

The first modern-day mobile phone was invented in 1973. Resembling a brick and weighing in at two pounds, it was priced out of reach for most consumers at nearly four thousand dollars. Since then, mobile phones have become smaller and less expensive; today mobile phones are a modern convenience available to all levels of society. As mobile phones evolved, they became more like small computers. These smartphones have many of the same characteristics as a personal computer, such as an operating system and memory. The first smartphone was the IBM Simon, introduced in 1994.

In January of 2007, Apple introduced the iPhone. Its ease of use and intuitive interface made it an immediate success and solidified the future of smartphones. Running on an operating system called iOS, the iPhone was really a small computer with a touch-screen interface. In 2008, the first Android phone was released, with similar functionality.

Tablet Computers

A tablet computer is one that uses a touch screen as its primary input and is small enough and light enough to be carried around easily. They generally have no keyboard and are self-contained inside a rectangular case. The first tablet computers appeared in the early 2000s and used an attached pen as a writing device for input. These tablets ranged in size from small personal digital assistants (PDAs), which were handheld, to full-sized, 14-inch devices. Most early tablets used a version of an existing computer operating system, such as Windows or Linux.

These early tablet devices were, for the most part, commercial failures. In January, 2010, Apple introduced the iPad, which ushered in a new era of tablet computing. Instead of a pen, the iPad used the finger as the primary input device. Instead of using the operating system of their desktop and laptop computers, Apple chose to use iOS, the operating system of the iPhone. Because the iPad had a user interface that was the same as the iPhone, consumers felt comfortable and sales took off. The iPad has set the standard for tablet computing. After the success of the iPad, computer manufacturers began to develop new tablets that utilized operating systems that were designed for mobile devices, such as Android.

The Rise of Mobile Computing

Mobile computing is having a huge impact on the business world today. The use of smartphones and tablet computers is rising at double-digit rates each year. The Gartner Group, in a report issued in April, 2013, estimates that over 1.7 million mobile phones will ship in the US in 2013 as compared to just over 340,000 personal computers. Over half of these mobile phones are smartphones.[2] Almost 200,000 tablet computers are predicted to ship in 2013. According to the report, PC shipments will continue to decline as phone and tablet shipments continue to increase. [3]

Integrated Computing

Along with advances in computers themselves, computing technology is being integrated into many everyday products. From automobiles to refrigerators to airplanes, computing technology is enhancing what these devices can do and is adding capabilities that would have been considered science fiction just a few years ago. Here are two of the latest ways that computing technologies are being integrated into everyday products:

- The Smart House
- The Self-Driving Car

12.2.3: Computers are part of life

The Commoditization of the Personal Computer

Over the past thirty years, as the personal computer has gone from technical marvel to part of our everyday lives, it has also become a commodity. The PC has become a commodity in the sense that there is very little differentiation between computers, and the primary factor that controls their sale is their price. Hundreds of manufacturers all over the world now create parts for personal computers. Dozens of companies buy these parts and assemble the computers. As commodities, there are essentially no differences between computers made by these different companies. Profit margins for personal computers are razor-thin, leading hardware developers to find the lowest-cost manufacturing.

There is one brand of computer for which this is not the case – Apple. Because Apple does not make computers that run on the same open standards as other manufacturers, they can make a unique product that no one can easily copy. By creating what many consider to be a superior product, Apple can charge more for their computers than other manufacturers. Just as with the iPad and iPhone, Apple has chosen a strategy of differentiation, which, at least at this time, seems to be paying off.

The Problem of Electronic Waste

Electronic waste (Public Domain)

Personal computers have been around for over thirty-five years. Millions of them have been used and discarded. Mobile phones are now available in even the remotest parts of the world and, after a few years of use, they are discarded. Where does this electronic debris end up?

Often, it gets routed to any country that will accept it. Many times, it ends up in dumps in developing nations. These dumps are beginning to be seen as health hazards for those living near them. Though many manufacturers have made strides in using materials that can be recycled, electronic waste is a problem with which we must all deal.

12.2.4: Summary

Summary

Information systems hardware consists of the components of digital technology that you can touch. In this chapter, we reviewed the components that make up a personal computer, with the understanding that the configuration of a personal computer is very similar to that of any type of digital computing device. A personal computer is made up of many components, most importantly the CPU, motherboard, RAM, hard disk, removable media, and input/output devices. We also reviewed some variations on the personal computer, such as the tablet computer and the smartphone. In accordance with Moore's Law, these technologies have improved quickly over the years, making today's computing devices much more powerful than devices just a few years ago. Finally, we discussed two of the consequences of this evolution: the commoditization of the personal computer and the problem of electronic waste.

Study Questions

1. Write your own description of what the term *information systems hardware* means.
2. What is the impact of Moore's Law on the various hardware components described in this chapter?
3. Write a summary of one of the items linked to in the "Integrated Computing" section.
4. Explain why the personal computer is now considered a commodity.
5. The CPU can also be thought of as the _____________ of the computer.
6. List the following in increasing order (slowest to fastest): megahertz, kilohertz, gigahertz.
7. What is the bus of a computer?
8. Name two differences between RAM and a hard disk.
9. What are the advantages of solid-state drives over hard disks?
10. How heavy was the first commercially successful portable computer?

Exercises

1. Review the sidebar on the binary number system. How would you represent the number 16 in binary? How about the number 100? Besides decimal and binary, other number bases are used in computing and programming. One of the most used bases is *hexadecimal*, which is base-16. In base-16, the numerals 0 through 9 are supplemented with the letters A (10) through F (15). How would you represent the decimal number 100 in hexadecimal?
2. Review the timeline of computers at the Old Computers website. Pick one computer from the listing and write a brief summary. Include the specifications for CPU, memory, and screen size. Now find the specifications of a computer being offered for sale today and compare. Did Moore's Law hold true?
3. The Homebrew Computer Club was one of the original clubs for enthusiasts of the first personal computer, the Altair 8800. Read some of their newsletters and then discuss some of the issues surrounding this early personal computer.
4. If you could build your own personal computer, what components would you purchase? Put together a list of the components you would use to create it, including a computer case, motherboard, CPU, hard disk, RAM, and DVD drive. How can you be sure they are all compatible with each other? How much would it cost? How does this compare to a similar computer purchased from a vendor such as Dell or HP?
5. Review the Wikipedia entry on electronic waste. Now find at least two more scholarly articles on this topic. Prepare a slideshow that summarizes the issue and then recommend a possible solution based on your research.
6. As with any technology text, there have been advances in technologies since publication. What technology that has been developed recently would you add to this chapter?
7. What is the current state of solid-state drives vs. hard disks? Do original research online where you can compare price on solid-state drives and hard disks. Be sure you note the differences in price, capacity, and speed.

1. Moore, Gordon E. (1965). "Cramming more components onto integrated circuits" (PDF). Electronics Magazine. p. 4. Retrieved 2012-10-18.↵
2. Smartphone shipments to surpass feature phones this year. CNet, June 4, 2013. http://news.cnet.com/8301-1035_3-575...nes-this-year/↵
3. Gartner Press Release. April 4, 2013. http://www.gartner.com/newsroom/id/2408515↵

SECTION OVERVIEW

12.3: Software

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- define the term *software*;
- describe the two primary categories of software;
- describe the role ERP software plays in an organization;
- describe cloud computing and its advantages and disadvantages for use in an organization; and
- define the term *open-source* and identify its primary characteristics.

Introduction

The second component of an information system is software. Simply put: Software is the set of instructions that tell the hardware what to do. Software is created through the process of programming (we will cover the creation of software in more detail in chapter 10). Without software, the hardware would not be functional.

Types of Software

Software can be broadly divided into two categories: operating systems and application software. *Operating systems* manage the hardware and create the interface between the hardware and the user. *Application software* is the category of programs that do something useful for the user.

12.3.1: Software Creation and Open Source Software

Software Creation

How is software created? If software is the set of instructions that tells the hardware what to do, how are these instructions written? If a computer reads everything as ones and zeroes, do we have to learn how to write software that way?

Modern software applications are written using a programming language. A programming language consists of a set of commands and syntax that can be organized logically to execute specific functions. This language generally consists of a set of readable words combined with symbols. Using this language, a programmer writes a program (called the source code) that can then be compiled into machine-readable form, the ones and zeroes necessary to be executed by the CPU. Examples of well-known programming languages today include Java, PHP, and various flavors of C (Visual C, C++, C#). Languages such as HTML and Javascript are used to develop web pages. Most of the time, programming is done inside a programming environment; when you purchase a copy of Visual Studio from Microsoft, it provides you with an editor, compiler, and help for many of Microsoft's programming languages.

Software programming was originally an individual process, with each programmer working on an entire program, or several programmers each working on a portion of a larger program. However, newer methods of software development include a more collaborative approach, with teams of programmers working on code together. We will cover information-systems development more fully in chapter 10.

Open-Source Software

When the personal computer was first released, it did not serve any practical need. Early computers were difficult to program and required great attention to detail. However, many personal-computer enthusiasts immediately banded together to build applications and solve problems. These computer enthusiasts were happy to share any programs they built and solutions to problems they found; this collaboration enabled them to more quickly innovate and fix problems.

As software began to become a business, however, this idea of sharing everything fell out of favor, at least with some. When a software program takes hundreds of man-hours to develop, it is understandable that the programmers do not want to just give it away. This led to a new business model of restrictive software licensing, which required payment for software, a model that is still dominant today. This model is sometimes referred to as *closed source*, as the source code is not made available to others.

There are many, however, who feel that software should not be restricted. Just as with those early hobbyists in the 1970s, they feel that innovation and progress can be made much more rapidly if we share what we learn. In the 1990s, with Internet access connecting more and more people together, the open-source movement gained steam.

Open-source software is software that makes the source code available for anyone to copy and use. For most of us, having access to the source code of a program does us little good, as we are not programmers and won't be able to do much with it. The good news is that open-source software is also available in a compiled format that we can simply download and install. The open-source movement has led to the development of some of the most-used software in the world, including the Firefox browser, the Linux operating system, and the Apache web server. Many also think open-source software is superior to closed-source software. Because the source code is freely available, many programmers have contributed to open-source software projects, adding features and fixing bugs.

Many businesses are wary of open-source software precisely because the code is available for anyone to see. They feel that this increases the risk of an attack. Others counter that this openness actually decreases the risk because the code is exposed to thousands of programmers who can incorporate code changes to quickly patch vulnerabilities.

There are many arguments on both sides of the aisle for the benefits of the two models. Some benefits of the open-source model are:

- The software is available for free.
- The software source-code is available; it can be examined and reviewed before it is installed.
- The large community of programmers who work on open-source projects leads to quick bug-fixing and feature additions.

Some benefits of the closed-source model are:

- By providing financial incentive for software development, some of the brightest minds have chosen software development as a career.
- Technical support from the company that developed the software.

Today there are thousands of open-source software applications available for download. For example, as we discussed previously in this chapter, you can get the productivity suite from Open Office. One good place to search for open-source software is sourceforge.net, where thousands of software applications are available for free download.

12.3.2: Operating Systems

Operating Systems

The operating system provides several essential functions, including:

1. managing the hardware resources of the computer;
2. providing the user-interface components;
3. providing a platform for software developers to write applications.

All computing devices run an operating system. For personal computers, the most popular operating systems are Microsoft's Windows, Apple's OS X, and different versions of Linux. Smartphones and tablets run operating systems as well, such as Apple's iOS, Google's Android, Microsoft's Windows Mobile, and Blackberry.

Early personal-computer operating systems were simple by today's standards; they did not provide multitasking and required the user to type commands to initiate an action. The amount of memory that early operating systems could handle was limited as well, making large programs impractical to run. The most popular of the early operating systems was IBM's Disk Operating System, or DOS, which was actually developed for them by Microsoft.

In 1984, Apple introduced the Macintosh computer, featuring an operating system with a graphical user interface. Though not the first graphical operating system, it was the first one to find commercial success. In 1985, Microsoft released the first version of Windows. This version of Windows was not an operating system, but instead was an application that ran on top of the DOS operating system, providing a graphical environment. It was quite limited and had little commercial success. It was not until the 1990 release of Windows 3.0 that Microsoft found success with a graphical user interface. Because of the hold of IBM and IBM-compatible personal computers on business, it was not until Windows 3.0 was released that business users began using a graphical user interface, ushering us into the graphical-computing era. Since 1990, both Apple and Microsoft have released many new versions of their operating systems, with each release adding the ability to process more data at once and access more memory. Features such as multitasking, virtual memory, and voice input have become standard features of both operating systems.

Linux logo (Copyright: Larry Ewing)

A third personal-computer operating system family that is gaining in popularity is Linux (pronounced "linn-ex"). Linux is an open source operating system. Linux is used on a wide variety of computing devices, including smartphones, tables, personal computers, Raspberry Pi, and even supercomputers. In 1991 a Finnish computer science student, Linus Torvalds, decided to create a new free operating system. Linux was the result. Linux has many variations and now powers a large percentage of web servers in the world. It is also an example of *open-source software*, a topic we will cover later in this chapter.

Sidebar: Mac vs. Windows

Are you a Mac? Are you a PC? Ever since its introduction in 1984, users of the Apple Macintosh have been quite biased about their preference for the Macintosh operating system (now called OS X) over Microsoft's. When Microsoft introduced Windows, Apple sued Microsoft, claiming that they copied the "look and feel" of the Macintosh operating system. In the end, Microsoft successfully defended themselves.

Over the past few years, Microsoft and Apple have traded barbs with each other, each claiming to have a better operating system and software. While Microsoft has always had the larger market share (see sidebar), Apple has been the favorite of artists, musicians, and the technology elite. Apple also provides a lot of computers to elementary schools, thus gaining a following among the younger generation.

Sidebar: Why Is Microsoft Software So Dominant in the Business World?

If you've worked in the world of business, you may have noticed that almost all of the computers run a version of Microsoft's Windows operating system. Why is this? On almost all college campuses, you see a preponderance of Apple Macintosh laptops. In elementary schools, Apple reigns as well. Why has this not extended into the business world?

As we learned in chapter 1, almost all businesses used IBM mainframe computers back in the 1960s and 1970s. These same businesses shied away from personal computers until IBM released the PC in 1981. When executives had to make a decision about purchasing personal computers for their employees, they would choose the safe route and purchase IBM. The saying then was: “No one ever got fired for buying IBM.” So over the next decade, companies bought IBM personal computers (or those compatible with them), which ran an operating system called DOS. DOS was created by Microsoft, so when Microsoft released Windows as the next iteration of DOS, companies took the safe route and started purchasing Windows.

Microsoft soon found itself with the dominant personal-computer operating system for businesses. As the networked personal computer began to replace the mainframe computer as the primary way of computing inside businesses, it became essential for Microsoft to give businesses the ability to administer and secure their networks. Microsoft developed business-level server products to go along with their personal computer products, thereby providing a complete business solution. And so now, the saying goes: “No one ever got fired for buying Microsoft.”

12.3.3: Application Software

Application Software

The second major category of software is application software. Application software is, essentially, software that allows the user to accomplish some goal or purpose. For example, if you have to write a paper, you might use the application-software program Microsoft Word. If you want to listen to music, you might use iTunes. To surf the web, you might use Internet Explorer or Firefox. Even a computer game could be considered application software.

The “Killer” App

VisiCalc running on an Apple II. (Public Domain)

When a new type of digital device is invented, there are generally a small group of technology enthusiasts who will purchase it just for the joy of figuring out how it works. However, for most of us, until a device can actually do something useful we are not going to spend our hard-earned money on it. A “killer” application is one that becomes so essential that large numbers of people will buy a device just to run that application. For the personal computer, the killer application was the spreadsheet. In 1979, VisiCalc, the first personal-computer spreadsheet package, was introduced. It was an immediate hit and drove sales of the Apple II. It also solidified the value of the personal computer beyond the relatively small circle of technology geeks. When the IBM PC was released, another spreadsheet program, Lotus 1-2-3, was the killer app for business users.

Productivity Software

Along with the spreadsheet, several other software applications have become standard tools for the workplace. These applications, called productivity software, allow office employees to complete their daily work. Many times, these applications come packaged together, such as in Microsoft’s Office suite. Here is a list of these applications and their basic functions:

- Word processing: This class of software provides for the creation of written documents. Functions include the ability to type and edit text, format fonts and paragraphs, and add, move, and delete text throughout the document. Most modern word-processing programs also have the ability to add tables, images, and various layout and formatting features to the document. Word processors save their documents as electronic files in a variety of formats. By far, the most popular word-processing package is Microsoft Word, which saves its files in the DOCX format. This format can be read/written by many other word-processor packages.
- Spreadsheet: This class of software provides a way to do numeric calculations and analysis. The working area is divided into rows and columns, where users can enter numbers, text, or formulas. It is the formulas that make a spreadsheet powerful, allowing the user to develop complex calculations that can change based on the numbers entered. Most spreadsheets also include the ability to create charts based on the data entered. The most popular spreadsheet package is Microsoft Excel, which saves its files in the XLSX format. Just as with word processors, many other spreadsheet packages can read and write to this file format.
- Presentation: This class of software provides for the creation of slideshow presentations. Harkening back to the days of overhead projectors and transparencies, presentation software allows its users to create a set of slides that can be printed or projected on a screen. Users can add text, images, and other media elements to the slides. Microsoft’s PowerPoint is the most popular software right now, saving its files in PPTX format.
- Some office suites include other types of software. For example, Microsoft Office includes Outlook, its e-mail package, and OneNote, an information-gathering collaboration tool. The professional version of Office also includes Microsoft Access, a database package. (Databases are covered more in chapter 4.)

> Microsoft popularized the idea of the office-software productivity bundle with their release of Microsoft Office. This package continues to dominate the market and most businesses expect employees to know how to use this software. However, many competitors to Microsoft Office do exist and are compatible with the file formats used by Microsoft (see table below). Recently, Microsoft has begun to offer a web version of their Office suite. Similar to Google Drive, this suite allows users to edit and share documents online utilizing cloud-computing technology. Cloud computing will be discussed later in this chapter.

Comparison of office application software suites

Utility Software and Programming Software

Two subcategories of application software worth mentioning are utility software and programming software. Utility software includes software that allows you to fix or modify your computer in some way. Examples include antivirus software and disk defragmentation software. These types of software packages were invented to fill shortcomings in operating systems. Many times, a subsequent release of an operating system will include these utility functions as part of the operating system itself.

Programming software is software whose purpose is to make more software. Most of these programs provide programmers with an environment in which they can write the code, test it, and convert it into the format that can then be run on a computer.

Sidebar: "PowerPointed" to Death

As presentation software, specifically Microsoft PowerPoint, has gained acceptance as the primary method to formally present information in a business setting, the art of giving an engaging presentation is becoming rare. Many presenters now just read the bullet points in the presentation and immediately bore those in attendance, who can already read it for themselves.

The real problem is not with PowerPoint as much as it is with the person creating and presenting. Author and thinker Seth Godin put it this way: "PowerPoint could be the most powerful tool on your computer. But it's not. It's actually a dismal failure. Almost every PowerPoint presentation sucks rotten eggs."[1] The software used to help you communicate should not duplicate the presentation you want to give, but instead it should support it. I highly recommend the book *Presentation Zen* by Garr Reynolds to anyone who wants to improve their presentation skills.

Software developers are becoming aware of this problem as well. New digital presentation technologies are being developed, with the hopes of becoming "the next PowerPoint." One innovative new presentation application is Prezi. Prezi is a presentation tool that uses a single canvas for the presentation, allowing presenters to place text, images, and other media on the canvas, and then navigate between these objects as they present. Just as with PowerPoint, Prezi should be used to supplement the presentation. And we must always remember that sometimes the best presentations are made with no digital tools.

Sidebar: I Own This Software, Right? Well . . .

When you purchase software and install it on your computer, are you the owner of that software? Technically, you are not! When you install software, you are actually just being given a license to use it. When you first install a software package, you are asked to agree to the terms of service or the license agreement. In that agreement, you will find that your rights to use the software are limited. For example, in the terms of the Microsoft Office Excel 2010 software license, you will find the following statement: "This software is licensed, not sold. This agreement only gives you some rights to use the features included in the software edition you licensed."

For the most part, these restrictions are what you would expect: you cannot make illegal copies of the software and you may not use it to do anything illegal. However, there are other, more unexpected terms in these software agreements. For example, many software agreements ask you to agree to a limit on liability. Again, from Microsoft: "Limitation on and exclusion of damages. You can recover from Microsoft and its suppliers only direct damages up to the amount you paid for the software. You cannot recover any other damages, including consequential, lost profits, special, indirect or incidental damages." What this means is that if a problem with the software causes harm to your business, you cannot hold Microsoft or the supplier responsible for damages.

Applications for the Enterprise

As the personal computer proliferated inside organizations, control over the information generated by the organization began splintering. Say the customer service department creates a customer database to keep track of calls and problem reports, and the sales department also creates a database to keep track of customer information. Which one should be used as the master list of customers? As another example, someone in sales might create a spreadsheet to calculate sales revenue, while someone in finance creates a different one that meets the needs of their department. However, it is likely that the two spreadsheets will come up with different totals for revenue. Which one is correct? And who is managing all of this information?

Enterprise Resource Planning

In the 1990s, the need to bring the organization's information back under centralized control became more apparent. The enterprise resource planning (ERP) system (sometimes just called enterprise software) was developed to bring together an entire organization in one software application. Simply put, an ERP system is a software application utilizing a central database that is implemented throughout the entire organization. Let's take a closer look at this definition:

- "A software application": An ERP is a software application that is used by many of an organization's employees.
- "utilizing a central database": All users of the ERP edit and save their information from the data source. What this means practically is that there is only one customer database, there is only one calculation for revenue, etc.
- "that is implemented throughout the entire organization": ERP systems include functionality that covers all of the essential components of a business. Further, an organization can purchase modules for its ERP system that match specific needs, such as manufacturing or planning.

Registered trademark of SAP

ERP systems were originally marketed to large corporations. However, as more and more large companies began installing them, ERP vendors began targeting mid-sized and even smaller businesses. Some of the more well-known ERP systems include those from SAP, Oracle, and Microsoft.

In order to effectively implement an ERP system in an organization, the organization must be ready to make a full commitment. All aspects of the organization are affected as old systems are replaced by the ERP system. In general, implementing an ERP system can take two to three years and several million dollars. In most cases, the cost of the software is not the most expensive part of the implementation: it is the cost of the consultants!

So why implement an ERP system? If done properly, an ERP system can bring an organization a good return on their investment. By consolidating information systems across the enterprise and using the software to enforce best practices, most organizations see an overall improvement after implementing an ERP. Business processes as a form of competitive advantage will be covered in chapter 9.

Sidebar: Y2K and ERP

The initial wave of software-application development began in the 1960s, when applications were developed for mainframe computers. In those days, computing was expensive, so applications were designed to take as little space as possible. One shortcut that many programmers took was in the storage of dates, specifically the year. Instead of allocating four digits to hold the year, many programs allocated two digits, making the assumption that the first two digits were "19″. For example, to calculate how old someone was, the application would take the last two digits of the current year (for 1995, for example, that would be "95″) and then subtract the two digits stored for the birthday year ("65″ for 1965). 95 minus 65 gives an age of 30, which is correct.

However, as the year 2000 approached, many of these "legacy" applications were still being used, and businesses were very concerned that any software applications they were using that needed to calculate dates would fail. To update our age-calculation example, the application would take the last two digits of the current year (for 2012, that would be "12″) and then subtract the two digits stored for the birthday year ("65″ for 1965). 12 minus 65 gives an age of -53, which would cause an error. In order to solve this problem, applications would have to be updated to use four digits for years instead of two. Solving this would be a massive undertaking, as every line of code and every database would have to be examined.

This is where companies gained additional incentive to implement an ERP system. For many organizations that were considering upgrading to ERP systems in the late 1990s, this problem, known as Y2K (year 2000), gave them the extra push they needed to get their ERP installed before the year 2000. ERP vendors guaranteed that their systems had been designed to be Y2K compliant – which simply meant that they stored dates using four digits instead of two. This led to a massive increase in ERP installations in the years leading up to 2000, making the ERP a standard software application for businesses.

Customer Relationship Management

A customer relationship management (CRM) system is a software application designed to manage an organization's customers. In today's environment, it is important to develop relationships with your customers, and the use of a well-designed CRM can allow a business to personalize its relationship with each of its customers. Some ERP software systems include CRM modules. An example of a well-known CRM package is Salesforce.

Supply Chain Management

Many organizations must deal with the complex task of managing their supply chains. At its simplest, a supply chain is the linkage between an organization's suppliers, its manufacturing facilities, and the distributors of its products. Each link in the chain has a multiplying effect on the complexity of the process: if there are two suppliers, one manufacturing facility, and two distributors, for example, then there are 2 x 1 x 2 = 4 links to handle. However, if you add two more suppliers, another manufacturing facility, and two more distributors, then you have 4 x 2 x 4 = 32 links to manage.

A supply chain management (SCM) system manages the interconnection between these links, as well as the inventory of the products in their various stages of development. A full definition of a supply chain management system is provided by the Association for Operations Management: "The design, planning, execution, control, and monitoring of supply chain activities with the objective of creating net value, building a competitive infrastructure, leveraging worldwide logistics, synchronizing supply with demand, and measuring performance globally."[2] Most ERP systems include a supply chain management module.

12.3.4: Mobile Apps

Mobile Applications

Just as with the personal computer, mobile devices such as tablet computers and smartphones also have operating systems and application software. In fact, these mobile devices are in many ways just smaller versions of personal computers. A mobile app is a software application programmed to run specifically on a mobile device.

As we saw in chapter 2, smartphones and tablets are becoming a dominant form of computing, with many more smartphones being sold than personal computers. This means that organizations will have to get smart about developing software on mobile devices in order to stay relevant.

These days, most mobile devices run on one of two operating systems: Android or iOS. Android is an open-source operating system purchased and supported by Google; iOS is Apple's mobile operating system. In the fourth quarter of 2012, Android was installed on 70.1% of all mobile phones shipped, followed by 21.0% for iOS. Other mobile operating systems of note are Blackberry (3.2%) and Windows (2.6%). [3]

As organizations consider making their digital presence compatible with mobile devices, they will have to decide whether to build a mobile app. A mobile app is an expensive proposition, and it will only run on one type of mobile device at a time. For example, if an organization creates an iPhone app, those with Android phones cannot run the application. Each app takes several thousand dollars to create, so this is not a trivial decision for many companies.

One option many companies have is to create a website that is mobile-friendly. A mobile website works on all mobile devices and costs about the same as creating an app. We will discuss the question of whether to build a mobile app more thoroughly in Chapter 10.

12.3.5: Cloud Computing

Cloud Computing

Historically, for software to run on a computer, an individual copy of the software had to be installed on the computer, either from a disk or, more recently, after being downloaded from the Internet. The concept of "cloud" computing changes this, however.

To understand cloud computing, we first have to understand what the cloud is. "The cloud" refers to applications, services, and data storage on the Internet. These service providers rely on giant server farms and massive storage devices that are connected via Internet protocols. Cloud computing is the use of these services by individuals and organizations.

You probably already use cloud computing in some forms. For example, if you access your e-mail via your web browser, you are using a form of cloud computing. If you use Google Drive's applications, you are using cloud computing. While these are free versions of cloud computing, there is big business in providing applications and data storage over the web. Salesforce (see above) is a good example of cloud computing – their entire suite of CRM applications are offered via the cloud. Cloud computing is not limited to web applications: it can also be used for services such as phone or video streaming.

Advantages of Cloud Computing

- No software to install or upgrades to maintain.
- Available from any computer that has access to the Internet.
- Can scale to a large number of users easily.
- New applications can be up and running very quickly.
- Services can be leased for a limited time on an as-needed basis.
- Your information is not lost if your hard disk crashes or your laptop is stolen.
- You are not limited by the available memory or disk space on your computer.

Disadvantages of Cloud Computing

- Your information is stored on someone else's computer – how safe is it?
- You must have Internet access to use it. If you do not have access, you're out of luck.
- You are relying on a third-party to provide these services.

Cloud computing has the ability to really impact how organizations manage technology. For example, why is an IT department needed to purchase, configure, and manage personal computers and software when all that is really needed is an Internet connection?

Using a Private Cloud

Many organizations are understandably nervous about giving up control of their data and some of their applications by using cloud computing. But they also see the value in reducing the need for installing software and adding disk storage to local computers. A solution to this problem lies in the concept of a *private cloud*. While there are various models of a private cloud, the basic idea is for the cloud service provider to section off web server space for a specific organization. The organization has full control over that server space while still gaining some of the benefits of cloud computing.

Virtualization

One technology that is utilized extensively as part of cloud computing is "virtualization." Virtualization is the process of using software to simulate a computer or some other device. For example, using virtualization, a single computer can perform the functions of several computers. Companies such as EMC provide virtualization software that allows cloud service providers to provision web servers to their clients quickly and efficiently. Organizations are also implementing virtualization in order to reduce the number of servers needed to provide the necessary services. For more detail on how virtualization works, see this informational page from VMWare.

12.3.6: Summary

Summary

Software gives the instructions that tell the hardware what to do. There are two basic categories of software: operating systems and applications. Operating systems provide access to the computer hardware and make system resources available. Application software is designed to meet a specific goal. Productivity software is a subset of application software that provides basic business functionality to a personal computer: word processing, spreadsheets, and presentations. An ERP system is a software application with a centralized database that is implemented across the entire organization. Cloud computing is a method of software delivery that runs on any computer that has a web browser and access to the Internet. Software is developed through a process called programming, in which a programmer uses a programming language to put together the logic needed to create the program. While most software is developed using a closed-source model, the open-source movement is gaining more support today.

Study Questions

1. Come up with your own definition of software. Explain the key terms in your definition.
2. What are the functions of the operating system?
3. Which of the following are operating systems and which are applications: Microsoft Excel, Google Chrome, iTunes, Windows, Android, Angry Birds.
4. What is your favorite software application? What tasks does it help you accomplish?
5. What is a “killer” app? What was the killer app for the PC?
6. How would you categorize the software that runs on mobile devices? Break down these apps into at least three basic categories and give an example of each.
7. Explain what an ERP system does.
8. What is open-source software? How does it differ from closed-source software? Give an example of each.
9. What does a software license grant?
10. How did the Y2K (year 2000) problem affect the sales of ERP systems?

Exercises

1. Go online and find a case study about the implementation of an ERP system. Was it successful? How long did it take? Does the case study tell you how much money the organization spent?
2. What ERP system does your university or place of employment use? Find out which one they use and see how it compares to other ERP systems.
3. If you were running a small business with limited funds for information technology, would you consider using cloud computing? Find some web-based resources that support your decision.
4. Download and install Open Office. Use it to create a document or spreadsheet. How does it compare to Microsoft Office? Does the fact that you got it for free make it feel less valuable?
5. Go to sourceforge.net and review their most downloaded software applications. Report back on the variety of applications you find. Then pick one that interests you and report back on what it does, the kind of technical support offered, and the user reviews.
6. Review this article on the security risks of open-source software. Write a short analysis giving your opinion on the different risks discussed.
7. What are three examples of programming languages? What makes each of these languages useful to programmers?

---

1. From *Why are your PowerPoints so bad?* available for download at http://www.sethgodin.com/freeprize/reallybad-1.pdf.↵
2. http://www.apics.org/dictionary/dict...mation?ID=3984↵
3. Taken from IDC Worldwide Mobile Phone Tracker, February 14, 2013. Full report available at http://www.idc.com/getdoc.jsp?containerId=prUS23946013↵

SECTION OVERVIEW

12.4: Data and Databases

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe the differences between data, information, and knowledge;
- define the term *database* and identify the steps to creating one;
- describe the role of a database management system;
- describe the characteristics of a data warehouse; and
- define data mining and describe its role in an organization.

Introduction

You have already been introduced to the first two components of information systems: hardware and software. However, those two components by themselves do not make a computer useful. Imagine if you turned on a computer, started the word processor, but could not save a document. Imagine if you opened a music player but there was no music to play. Imagine opening a web browser but there were no web pages. Without data, hardware and software are not very useful! Data is the third component of an information system.

Data, Information, and Knowledge

Data are the raw bits and pieces of information with no context. If I told you, “15, 23, 14, 85,” you would not have learned anything. But I would have given you data.

Data can be quantitative or qualitative. Quantitative data is numeric, the result of a measurement, count, or some other mathematical calculation. Qualitative data is descriptive. “Ruby Red,” the color of a 2013 Ford Focus, is an example of qualitative data. A number can be qualitative too: if I tell you my favorite number is 5, that is qualitative data because it is descriptive, not the result of a measurement or mathematical calculation.

By itself, data is not that useful. To be useful, it needs to be given context. Returning to the example above, if I told you that “15, 23, 14, and 85″ are the numbers of students that had registered for upcoming classes, that would be *information*. By adding the context – that the numbers represent the count of students registering for specific classes – I have converted data into information.

Once we have put our data into context, aggregated and analyzed it, we can use it to make decisions for our organization. We can say that this consumption of information produces *knowledge*. This knowledge can be used to make decisions, set policies, and even spark innovation.

The final step up the information ladder is the step from knowledge (knowing a lot about a topic) to *wisdom*. We can say that someone has wisdom when they can combine their knowledge and experience to produce a deeper understanding of a topic. It often takes many years to develop wisdom on a particular topic, and requires patience.

Examples of Data

Almost all software programs require data to do anything useful. For example, if you are editing a document in a word processor such as Microsoft Word, the document you are working on is the data. The word-processing software can manipulate the data: create a new document, duplicate a document, or modify a document. Some other examples of data are: an MP3 music file, a video file, a spreadsheet, a web page, and an e-book. In some cases, such as with an e-book, you may only have the ability to read the data.

Databases

The goal of many information systems is to transform data into information in order to generate knowledge that can be used for decision making. In order to do this, the system must be able to take data, put the data into context, and provide tools for aggregation and analysis. A database is designed for just such a purpose.

A database is an organized collection of related information. It is an *organized* collection, because in a database, all data is described and associated with other data. All information in a database should be *related* as well; separate databases should be created to manage unrelated information. For example, a database that contains information about students should not also hold information about company stock prices. Databases are not always digital – a filing cabinet, for instance, might be considered a form of database. For the purposes of this text, we will only consider digital databases.

Topic hierarchy

12.4.1: Relational Database

Relational Databases

Databases can be organized in many different ways, and thus take many forms. The most popular form of database today is the relational database. Popular examples of relational databases are Microsoft Access, MySQL, and Oracle. A relational database is one in which data is organized into one or more tables. Each table has a set of fields, which define the nature of the data stored in the table. A record is one instance of a set of fields in a table. To visualize this, think of the records as the rows of the table and the fields as the columns of the table. In the example below, we have a table of student information, with each row representing a student and each column representing one piece of information about the student.

Rows and columns in a table

In a relational database, all the tables are related by one or more fields, so that it is possible to connect all the tables in the database through the field(s) they have in common. For each table, one of the fields is identified as a primary key. This key is the unique identifier for each record in the table. To help you understand these terms further, let's walk through the process of designing a database.

Other Types of Databases

The relational database model is the most used database model today. However, many other database models exist that provide different strengths than the relational model. The hierarchical database model, popular in the 1960s and 1970s, connected data together in a hierarchy, allowing for a parent/child relationship between data. The document-centric model allowed for a more unstructured data storage by placing data into "documents" that could then be manipulated.

Perhaps the most interesting new development is the concept of NoSQL (from the phrase "not only SQL"). NoSQL arose from the need to solve the problem of large-scale databases spread over several servers or even across the world. For a relational database to work properly, it is important that only one person be able to manipulate a piece of data at a time, a concept known as record-locking. But with today's large-scale databases (think Google and Amazon), this is just not possible. A NoSQL database can work with data in a looser way, allowing for a more unstructured environment, communicating changes to the data over time to all the servers that are part of the database.

12.4.1.1: Designing a Database

Designing a Database

Suppose a university wants to create an information system to track participation in student clubs. After interviewing several people, the design team learns that the goal of implementing the system is to give better insight into how the university funds clubs. This will be accomplished by tracking how many members each club has and how active the clubs are. From this, the team decides that the system must keep track of the clubs, their members, and their events. Using this information, the design team determines that the following tables need to be created:

- Clubs: this will track the club name, the club president, and a short description of the club.
- Students: student name, e-mail, and year of birth.
- Memberships: this table will correlate students with clubs, allowing us to have any given student join multiple clubs.
- Events: this table will track when the clubs meet and how many students showed up.

Now that the design team has determined which tables to create, they need to define the specific information that each table will hold. This requires identifying the fields that will be in each table. For example, Club Name would be one of the fields in the Clubs table. First Name and Last Name would be fields in the Students table. Finally, since this will be a relational database, every table should have a field in common with at least one other table (in other words: they should have a relationship with each other).

In order to properly create this relationship, a primary key must be selected for each table. This key is a unique identifier for each record in the table. For example, in the Students table, it might be possible to use students' last name as a way to uniquely identify them. However, it is more than likely that some students will share a last name (like Rodriguez, Smith, or Lee), so a different field should be selected. A student's e-mail address might be a good choice for a primary key, since e-mail addresses are unique. However, a primary key cannot change, so this would mean that if students changed their e-mail address we would have to remove them from the database and then re-insert them – not an attractive proposition. Our solution is to create a value for each student — a user ID — that will act as a primary key. We will also do this for each of the student clubs. This solution is quite common and is the reason you have so many user IDs!

You can see the final database design in the figure below:

Student Clubs database diagram

With this design, not only do we have a way to organize all of the information we need to meet the requirements, but we have also successfully related all the tables together. Here's what the database tables might look like with some sample data. Note that the Memberships table has the sole purpose of allowing us to relate multiple students to multiple clubs.

12.4.1.2: Normalization of a Database

Normalization

When designing a database, one important concept to understand is *normalization*. In simple terms, to normalize a database means to design it in a way that: 1) reduces duplication of data between tables and 2) gives the table as much flexibility as possible.

In the Student Clubs database design, the design team worked to achieve these objectives. For example, to track memberships, a simple solution might have been to create a Members field in the Clubs table and then just list the names of all of the members there. However, this design would mean that if a student joined two clubs, then his or her information would have to be entered a second time. Instead, the designers solved this problem by using two tables: Students and Memberships.

In this design, when a student joins their first club, we first must add the student to the Students table, where their first name, last name, e-mail address, and birth year are entered. This addition to the Students table will generate a student ID. Now we will add a new entry to denote that the student is a member of a specific club. This is accomplished by adding a record with the student ID and the club ID in the Memberships table. If this student joins a second club, we do not have to duplicate the entry of the student's name, e-mail, and birth year; instead, we only need to make another entry in the Memberships table of the second club's ID and the student's ID.

The design of the Student Clubs database also makes it simple to change the design without major modifications to the existing structure. For example, if the design team were asked to add functionality to the system to track faculty advisors to the clubs, we could easily accomplish this by adding a Faculty Advisors table (similar to the Students table) and then adding a new field to the Clubs table to hold the Faculty Advisor ID.

12.4.1.3: Data Types

Data Types

When defining the fields in a database table, we must give each field a data type. For example, the field Birth Year is a year, so it will be a number, while First Name will be text. Most modern databases allow for several different data types to be stored. Some of the more common data types are listed here:

- Text: for storing non-numeric data that is brief, generally under 256 characters. The database designer can identify the maximum length of the text.
- Number: for storing numbers. There are usually a few different number types that can be selected, depending on how large the largest number will be.
- Yes/No: a special form of the number data type that is (usually) one byte long, with a 0 for “No” or “False” and a 1 for “Yes” or “True”.
- Date/Time: a special form of the number data type that can be interpreted as a number or a time.
- Currency: a special form of the number data type that formats all values with a currency indicator and two decimal places.
- Paragraph Text: this data type allows for text longer than 256 characters.
- Object: this data type allows for the storage of data that cannot be entered via keyboard, such as an image or a music file.

There are two important reasons that we must properly define the data type of a field. First, a data type tells the database what functions can be performed with the data. For example, if we wish to perform mathematical functions with one of the fields, we must be sure to tell the database that the field is a number data type. So if we have, say, a field storing birth year, we can subtract the number stored in that field from the current year to get age.

The second important reason to define data type is so that the proper amount of storage space is allocated for our data. For example, if the First Name field is defined as a text(50) data type, this means fifty characters are allocated for each first name we want to store. However, even if the first name is only five characters long, fifty characters (bytes) will be allocated. While this may not seem like a big deal, if our table ends up holding 50,000 names, we are allocating 50 * 50,000 = 2,500,000 bytes for storage of these values. It may be prudent to reduce the size of the field so we do not waste storage space.

Sidebar: The Difference between a Database and a Spreadsheet

Many times, when introducing the concept of databases to students, they quickly decide that a database is pretty much the same as a spreadsheet. After all, a spreadsheet stores data in an organized fashion, using rows and columns, and looks very similar to a database table. This misunderstanding extends beyond the classroom: spreadsheets are used as a substitute for databases in all types of situations every day, all over the world.

To be fair, for simple uses, a spreadsheet can substitute for a database quite well. If a simple listing of rows and columns (a single table) is all that is needed, then creating a database is probably overkill. In our Student Clubs example, if we only needed to track a listing of clubs, the number of members, and the contact information for the president, we could get away with a single spreadsheet. However, the need to include a listing of events and the names of members would be problematic if tracked with a spreadsheet.

When several types of data must be mixed together, or when the relationships between these types of data are complex, then a spreadsheet is not the best solution. A database allows data from several entities (such as students, clubs, memberships, and events) to all be related together into one whole. While a spreadsheet does allow you to define what kinds of values can be entered into its cells, a database provides more intuitive and powerful ways to define the types of data that go into each field, reducing possible errors and allowing for easier analysis.

Though not good for replacing databases, spreadsheets can be ideal tools for analyzing the data stored in a database. A spreadsheet package can be connected to a specific table or query in a database and used to create charts or perform analysis on that data.

12.4.1.4: Structured Query Language (SQL)

Structured Query Language

Once you have a database designed and loaded with data, how will you do something useful with it? The primary way to work with a relational database is to use Structured Query Language, SQL (pronounced "sequel," or simply stated as S-Q-L). Almost all applications that work with databases (such as database management systems, discussed below) make use of SQL as a way to analyze and manipulate relational data. As its name implies, SQL is a language that can be used to work with a relational database. From a simple request for data to a complex update operation, SQL is a mainstay of programmers and database administrators. To give you a taste of what SQL might look like, here are a couple of examples using our Student Clubs database.

- The following query will retrieve a list of the first and last names of the club presidents:

```
SELECT "First Name", "Last Name" FROM "Students" WHERE "Students.ID" = "Clubs.Preside
```

- The following query will create a list of the number of students in each club, listing the club name and then the number of members:

```
SELECT "Clubs.Club Name", COUNT("Memberships.Student ID") FROM "Clubs" LEFT JOIN "Mem
```

An in-depth description of how SQL works is beyond the scope of this introductory text, but these examples should give you an idea of the power of using SQL to manipulate relational data. Many database packages, such as Microsoft Access, allow you to visually create the query you want to construct and then generate the SQL query for you.

12.4.2: Database Management Systems

Database Management Systems

Screen shot of the Open Office database management system

To the computer, a database looks like one or more files. In order for the data in the database to be read, changed, added, or removed, a software program must access it. Many software applications have this ability: iTunes can read its database to give you a listing of its songs (and play the songs); your mobile-phone software can interact with your list of contacts. But what about applications to create or manage a database? What software can you use to create a database, change a database's structure, or simply do analysis? That is the purpose of a category of software applications called database management systems (DBMS).

DBMS packages generally provide an interface to view and change the design of the database, create queries, and develop reports. Most of these packages are designed to work with a specific type of database, but generally are compatible with a wide range of databases.

For example, Apache OpenOffice.org Base (see screen shot) can be used to create, modify, and analyze databases in open-database (ODB) format. Microsoft's Access DBMS is used to work with databases in its own Microsoft Access Database format. Both Access and Base have the ability to read and write to other database formats as well.

Microsoft Access and Open Office Base are examples of personal database-management systems. These systems are primarily used to develop and analyze single-user databases. These databases are not meant to be shared across a network or the Internet, but are instead installed on a particular device and work with a single user at a time.

Enterprise Databases

A database that can only be used by a single user at a time is not going to meet the needs of most organizations. As computers have become networked and are now joined worldwide via the Internet, a class of database has emerged that can be accessed by two, ten, or even a million people. These databases are sometimes installed on a single computer to be accessed by a group of people at a single location. Other times, they are installed over several servers worldwide, meant to be accessed by millions. These relational enterprise database packages are built and supported by companies such as Oracle, Microsoft, and IBM. The open-source MySQL is also an enterprise database.

As stated earlier, the relational database model does not scale well. The term *scale* here refers to a database getting larger and larger, being distributed on a larger number of computers connected via a network. Some companies are looking to provide large-scale database solutions by moving away from the relational model to other, more flexible models. For example, Google now offers the App Engine Datastore, which is based on NoSQL. Developers can use the App Engine Datastore to develop applications that access data from anywhere in the world. Amazon.com offers several database services for enterprise use, including Amazon RDS, which is a relational database service, and Amazon DynamoDB, a NoSQL enterprise solution.

12.4.3: Big Data

Big Data

A new buzzword that has been capturing the attention of businesses lately is *big data*. The term refers to such massively large data sets that conventional database tools do not have the processing power to analyze them. For example, Walmart must process over one million customer transactions every hour. Storing and analyzing that much data is beyond the power of traditional database-management tools. Understanding the best tools and techniques to manage and analyze these large data sets is a problem that governments and businesses alike are trying to solve.

Sidebar: What Is Metadata?

The term *metadata* can be understood as "data about data." For example, when looking at one of the values of Year of Birth in the Students table, the data itself may be "1992″. The metadata about that value would be the field name Year of Birth, the time it was last updated, and the data type (integer). Another example of metadata could be for an MP3 music file, like the one shown in the image below; information such as the length of the song, the artist, the album, the file size, and even the album cover art, are classified as metadata. When a database is being designed, a "data dictionary" is created to hold the metadata, defining the fields and structure of the database.

12.4.4: Data Warehousing

Data Warehouse

As organizations have begun to utilize databases as the centerpiece of their operations, the need to fully understand and leverage the data they are collecting has become more and more apparent. However, directly analyzing the data that is needed for day-to-day operations is not a good idea; we do not want to tax the operations of the company more than we need to. Further, organizations also want to analyze data in a historical sense: How does the data we have today compare with the same set of data this time last month, or last year? From these needs arose the concept of the data warehouse.

The concept of the data warehouse is simple: extract data from one or more of the organization's databases and load it into the data warehouse (which is itself another database) for storage and analysis. However, the execution of this concept is not that simple. A data warehouse should be designed so that it meets the following criteria:

- It uses non-operational data. This means that the data warehouse is using a copy of data from the active databases that the company uses in its day-to-day operations, so the data warehouse must pull data from the existing databases on a regular, scheduled basis.
- The data is time-variant. This means that whenever data is loaded into the data warehouse, it receives a time stamp, which allows for comparisons between different time periods.
- The data is standardized. Because the data in a data warehouse usually comes from several different sources, it is possible that the data does not use the same definitions or units. For example, our Events table in our Student Clubs database lists the event dates using the mm/dd/yyyy format (e.g., 01/10/2013). A table in another database might use the format yy/mm/dd (e.g., 13/01/10) for dates. In order for the data warehouse to match up dates, a standard date format would have to be agreed upon and all data loaded into the data warehouse would have to be converted to use this standard format. This process is called extraction-transformation-load (ETL).

There are two primary schools of thought when designing a data warehouse: bottom-up and top-down. The bottom-up approach starts by creating small data warehouses, called data marts, to solve specific business problems. As these data marts are created, they can be combined into a larger data warehouse. The top-down approach suggests that we should start by creating an enterprise-wide data warehouse and then, as specific business needs are identified, create smaller data marts from the data warehouse.

Data warehouse process (top-down)

Benefits of Data Warehouses

Organizations find data warehouses quite beneficial for a number of reasons:

- The process of developing a data warehouse forces an organization to better understand the data that it is currently collecting and, equally important, what data is not being collected.
- A data warehouse provides a centralized view of all data being collected across the enterprise and provides a means for determining data that is inconsistent.
- Once all data is identified as consistent, an organization can generate one version of the truth. This is important when the company wants to report consistent statistics about itself, such as revenue or number of employees.
- By having a data warehouse, snapshots of data can be taken over time. This creates a historical record of data, which allows for an analysis of trends.
- A data warehouse provides tools to combine data, which can provide new information and analysis.

12.4.5: Data Mining and stuff

Data Mining

Data mining is the process of analyzing data to find previously unknown trends, patterns, and associations in order to make decisions. Generally, data mining is accomplished through automated means against extremely large data sets, such as a data warehouse. Some examples of data mining include:

- An analysis of sales from a large grocery chain might determine that milk is purchased more frequently the day after it rains in cities with a population of less than 50,000.
- A bank may find that loan applicants whose bank accounts show particular deposit and withdrawal patterns are not good credit risks.
- A baseball team may find that collegiate baseball players with specific statistics in hitting, pitching, and fielding make for more successful major league players.

In some cases, a data-mining project is begun with a hypothetical result in mind. For example, a grocery chain may already have some idea that buying patterns change after it rains and want to get a deeper understanding of exactly what is happening. In other cases, there are no presuppositions and a data-mining program is run against large data sets in order to find patterns and associations.

Privacy Concerns

The increasing power of data mining has caused concerns for many, especially in the area of privacy. In today's digital world, it is becoming easier than ever to take data from disparate sources and combine them to do new forms of analysis. In fact, a whole industry has sprung up around this technology: data brokers. These firms combine publicly accessible data with information obtained from the government and other sources to create vast warehouses of data about people and companies that they can then sell. This subject will be covered in much more detail in chapter 12 – the chapter on the ethical concerns of information systems.

Business Intelligence and Business Analytics

With tools such as data warehousing and data mining at their disposal, businesses are learning how to use information to their advantage. The term *business intelligence* is used to describe the process that organizations use to take data they are collecting and analyze it in the hopes of obtaining a competitive advantage. Besides using data from their internal databases, firms often purchase information from data brokers to get a big-picture understanding of their industries. *Business analytics* is the term used to describe the use of internal company data to improve business processes and practices.

Knowledge Management

We end the chapter with a discussion on the concept of knowledge management (KM). All companies accumulate knowledge over the course of their existence. Some of this knowledge is written down or saved, but not in an organized fashion. Much of this knowledge is not written down; instead, it is stored inside the heads of its employees. Knowledge management is the process of formalizing the capture, indexing, and storing of the company's knowledge in order to benefit from the experiences and insights that the company has captured during its existence.

12.4.6: Unit 4 Summary

Summary

In this chapter, we learned about the role that data and databases play in the context of information systems. Data is made up of small facts and information without context. If you give data context, then you have information. Knowledge is gained when information is consumed and used for decision making. A database is an organized collection of related information. Relational databases are the most widely used type of database, where data is structured into tables and all tables must be related to each other through unique identifiers. A database management system (DBMS) is a software application that is used to create and manage databases, and can take the form of a personal DBMS, used by one person, or an enterprise DBMS that can be used by multiple users. A data warehouse is a special form of database that takes data from other databases in an enterprise and organizes it for analysis. Data mining is the process of looking for patterns and relationships in large data sets. Many businesses use databases, data warehouses, and data-mining techniques in order to produce business intelligence and gain a competitive advantage.

Study Questions

1. What is the difference between data, information, and knowledge?
2. Explain in your own words how the data component relates to the hardware and software components of information systems.
3. What is the difference between quantitative data and qualitative data? In what situations could the number 42 be considered qualitative data?
4. What are the characteristics of a relational database?
5. When would using a personal DBMS make sense?
6. What is the difference between a spreadsheet and a database? List three differences between them.
7. Describe what the term *normalization* means.
8. Why is it important to define the data type of a field when designing a relational database?
9. Name a database you interact with frequently. What would some of the field names be?
10. What is metadata?
11. Name three advantages of using a data warehouse.
12. What is data mining?

Exercises

1. Review the design of the Student Clubs database earlier in this chapter. Reviewing the lists of data types given, what data types would you assign to each of the fields in each of the tables. What lengths would you assign to the text fields?
2. Download Apache OpenOffice.org and use the database tool to open the "Student Clubs.odb" file available here. Take some time to learn how to modify the database structure and then see if you can add the required items to support the tracking of faculty advisors, as described at the end of the Normalization section in the chapter. Here is a link to the Getting Started documentation.
3. Using Microsoft Access, download the database file of comprehensive baseball statistics from the website SeanLahman.com. (If you don't have Microsoft Access, you can download an abridged version of the file here that is compatible with Apache Open Office). Review the structure of the tables included in the database. Come up with three different data-mining experiments you would like to try, and explain which fields in which tables would have to be analyzed.
4. Do some original research and find two examples of data mining. Summarize each example and then write about what the two examples have in common.
5. Conduct some independent research on the process of business intelligence. Using at least two scholarly or practitioner sources, write a two-page paper giving examples of how business intelligence is being used.
6. Conduct some independent research on the latest technologies being used for knowledge management. Using at least two scholarly or practitioner sources, write a two-page paper giving examples of software applications or new technologies being used in this field.

SECTION OVERVIEW

12.5: Networking and Communication

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- understand the history and development of networking technologies;
- define the key terms associated with networking technologies;
- understand the importance of broadband technologies; and
- describe organizational networking.

Introduction

In the early days of computing, computers were seen as devices for making calculations, storing data, and automating business processes. However, as the devices evolved, it became apparent that many of the functions of telecommunications could be integrated into the computer. During the 1980s, many organizations began combining their once-separate telecommunications and information-systems departments into an information technology, or IT, department. This ability for computers to communicate with one another and, maybe more importantly, to facilitate communication between individuals and groups, has been an important factor in the growth of computing over the past several decades.

Computer networking really began in the 1960s with the birth of the Internet, as we'll see below. However, while the Internet and web were evolving, corporate networking was also taking shape in the form of local area networks and client-server computing. In the 1990s, when the Internet came of age, Internet technologies began to pervade all areas of the organization. Now, with the Internet a global phenomenon, it would be unthinkable to have a computer that did not include communications capabilities. This chapter will review the different technologies that have been put in place to enable this communications revolution.

Sidebar: An Internet Vocabulary Lesson

Networking communication is full of some very technical concepts based on some simple principles. Learn the terms below and you'll be able to hold your own in a conversation about the Internet.

- *Packet*: The fundamental unit of data transmitted over the Internet. When a device intends to send a message to another device (for example, your PC sends a request to YouTube to open a video), it breaks the message down into smaller pieces, called packets. Each packet has the sender's address, the destination address, a sequence number, and a piece of the overall message to be sent.
- *Hub*: A simple network device that connects other devices to the network and sends packets to all the devices connected to it.
- *Bridge*: A network device that connects two networks together and only allows packets through that are needed.
- *Switch*: A network device that connects multiple devices together and filters packets based on their destination within the connected devices.
- *Router*: A device that receives and analyzes packets and then routes them towards their destination. In some cases, a router will send a packet to another router; in other cases, it will send it directly to its destination.
- *IP Address*: Every device that communicates on the Internet, whether it be a personal computer, a tablet, a smartphone, or anything else, is assigned a unique identifying number called an IP (Internet Protocol) address. Historically, the IP-address standard used has been IPv4 (version 4), which has the format of four numbers between 0 and 255 separated by a period. For example, the domain Saylor.org has the IP address of 107.23.196.166. The IPv4 standard has a limit of 4,294,967,296 possible addresses. As the use of the Internet has proliferated, the number of IP addresses needed has grown to the point where the use of IPv4 addresses will be exhausted. This has led to the new IPv6 standard, which is currently being phased in. The IPv6 standard is formatted as eight groups of four hexadecimal digits, such as 2001:0db8:85a3:0042:1000:8a2e:0370:7334. The IPv6 standard has a limit of 3.4×10^{38} possible addresses. For more detail about the new IPv6 standard, see this Wikipedia article.
- *Domain name*: If you had to try to remember the IP address of every web server you wanted to access, the Internet would not be nearly as easy to use. A domain name is a human-friendly name for a device on the Internet. These names generally

consist of a descriptive text followed by the top-level domain (TLD). For example, Wikepedia's domain name is wikipedia.org; *wikipedia* describes the organization and *.org* is the top-level domain. In this case, the *.org* TLD is designed for nonprofit organizations. Other well-known TLDs include *.com*, *.net*, and *.gov*. For a complete list and description of domain names, see this Wikipedia article.

- *DNS*: DNS stands for "domain name system," which acts as the directory on the Internet. When a request to access a device with a domain name is given, a DNS server is queried. It returns the IP address of the device requested, allowing for proper routing.
- *Packet-switching*: When a packet is sent from one device out over the Internet, it does not follow a straight path to its destination. Instead, it is passed from one router to another across the Internet until it is reaches its destination. In fact, sometimes two packets from the same message will take different routes! Sometimes, packets will arrive at their destination out of order. When this happens, the receiving device restores them to their proper order. For more details on packet-switching, see this interactive web page.
- *Protocol*: In computer networking, a protocol is the set of rules that allow two (or more) devices to exchange information back and forth across the network.

12.5.1: History Lesson

In the Beginning: ARPANET

The story of the Internet, and networking in general, can be traced back to the late 1950s. The US was in the depths of the Cold War with the USSR, and each nation closely watched the other to determine which would gain a military or intelligence advantage. In 1957, the Soviets surprised the US with the launch of Sputnik, propelling us into the space age. In response to Sputnik, the US Government created the Advanced Research Projects Agency (ARPA), whose initial role was to ensure that the US was not surprised again. It was from ARPA, now called DARPA (Defense Advanced Research Projects Agency), that the Internet first sprang.

ARPA was the center of computing research in the 1960s, but there was just one problem: many of the computers could not talk to each other. In 1968, ARPA sent out a request for proposals for a communication technology that would allow different computers located around the country to be integrated together into one network. Twelve companies responded to the request, and a company named Bolt, Beranek, and Newman (BBN) won the contract. They began work right away and were able to complete the job just one year later: in September, 1969, the ARPANET was turned on. The first four nodes were at UCLA, Stanford, MIT, and the University of Utah.

12.5.2: Internet and the Web

The Internet and the World Wide Web

Over the next decade, the ARPANET grew and gained popularity. During this time, other networks also came into existence. Different organizations were connected to different networks. This led to a problem: the networks could not talk to each other. Each network used its own proprietary language, or protocol (see sidebar for the definition of *protocol*), to send information back and forth. This problem was solved by the invention of transmission control protocol/Internet protocol (TCP/IP). TCP/IP was designed to allow networks running on different protocols to have an intermediary protocol that would allow them to communicate. So as long as your network supported TCP/IP, you could communicate with all of the other networks running TCP/IP. TCP/IP quickly became the standard protocol and allowed networks to communicate with each other. It is from this breakthrough that we first got the term *Internet*, which simply means "an interconnected network of networks."

Worldwide Internet use over 24-hour period (click to go to site of origin). (Public Domain. Courtesy of the Internet Census 2012 project.)

As we moved into the 1980s, computers were added to the Internet at an increasing rate. These computers were primarily from government, academic, and research organizations. Much to the surprise of the engineers, the early popularity of the Internet was driven by the use of electronic mail (see sidebar below).

Using the Internet in these early days was not easy. In order to access information on another server, you had to know how to type in the commands necessary to access it, as well as know the name of that device. That all changed in 1990, when Tim Berners-Lee introduced his **World Wide Web** project, which provided an easy way to navigate the Internet through the use of linked text (hypertext). The World Wide Web gained even more steam with the release of the Mosaic browser in 1993, which allowed graphics and text to be combined together as a way to present information and navigate the Internet. The Mosaic browser took off in popularity and was soon superseded by Netscape Navigator, the first commercial web browser, in 1994. The Internet and the World Wide Web were now poised for growth. The chart below shows the growth in users from the early days until now.

Growth of internet usage, 1995–2012 (click to enlarge). Data taken from InternetWorldStats.com.

The Dot-Com Bubble

In the 1980s and early 1990s, the Internet was being managed by the National Science Foundation (NSF). The NSF had restricted commercial ventures on the Internet, which meant that no one could buy or sell anything online. In 1991, the NSF transferred its role to three other organizations, thus getting the US government out of direct control over the Internet and essentially opening up commerce online.

This new commercialization of the Internet led to what is now known as the dot-com bubble. A frenzy of investment in new dot-com companies took place in the late 1990s, running up the stock market to new highs on a daily basis. This investment bubble was driven by the fact that investors knew that online commerce would change everything. Unfortunately, many of these new companies had poor business models and ended up with little to show for all of the funds that were invested in them. In 2000 and 2001, the bubble burst and many of these new companies went out of business. Many companies also survived, including the still-thriving Amazon (started in 1994) and eBay (1995). After the dot-com bubble burst, a new reality became clear: in order to succeed online, e-business companies would need to develop real business models and show that they could survive financially using this new technology.

Web 2.0

In the first few years of the World Wide Web, creating and putting up a website required a specific set of knowledge: you had to know how to set up a server on the World Wide Web, how to get a domain name, how to write web pages in HTML, and how to troubleshoot various technical issues as they came up. Someone who did these jobs for a website became known as a webmaster.

As the web gained in popularity, it became more and more apparent that those who did not have the skills to be a webmaster still wanted to create online content and have their own piece of the web. This need was met with new technologies that provided a website framework for those who wanted to put content online. Blogger and Wikipedia are examples of these early Web 2.0

applications, which allowed anyone with something to say a place to go and say it, without the need for understanding HTML or web-server technology.

Starting in the early 2000s, Web 2.0 applications began a second bubble of optimism and investment. It seemed that everyone wanted their own blog or photo-sharing site. Here are some of the companies that came of age during this time: MySpace (2003), Photobucket (2003), Flickr (2004), Facebook (2004), WordPress (2005), Tumblr (2006), and Twitter (2006). The ultimate indication that Web 2.0 had taken hold was when *Time* magazine named "You" its "Person of the Year" in 2006.

Sidebar: E-mail Is the "Killer" App for the Internet

When the personal computer was created, it was a great little toy for technology hobbyists and armchair programmers. As soon as the spreadsheet was invented, however, businesses took notice, and the rest is history. The spreadsheet was the killer app for the personal computer: people bought PCs just so they could run spreadsheets.

The Internet was originally designed as a way for scientists and researchers to share information and computing power among themselves. However, as soon as electronic mail was invented, it began driving demand for the Internet. This wasn't what the developers had in mind, but it turned out that people connecting to people was the killer app for the Internet.

We are seeing this again today with social networks, specifically Facebook. Many who weren't convinced to have an online presence now feel left out without a Facebook account. The connections made between people using Web 2.0 applications like Facebook on their personal computer or smartphone is driving growth yet again.

Sidebar: The Internet and the World Wide Web Are Not the Same Thing

Many times, the terms "Internet" and "World Wide Web," or even just "the web," are used interchangeably. But really, they are *not* the same thing at all! The Internet is an interconnected network of networks. Many services run across the Internet: electronic mail, voice and video, file transfers, and, yes, the World Wide Web.

The World Wide Web is simply one piece of the Internet. It is made up of web servers that have HTML pages that are being viewed on devices with web browsers. It is really that simple.

12.5.3: Broadband

The Growth of Broadband

In the early days of the Internet, most access was done via a modem over an analog telephone line. A modem (short for "modulator-demodulator") was connected to the incoming phone line and a computer in order to connect you to a network. Speeds were measured in bits-per-second (bps), with speeds growing from 1200 bps to 56,000 bps over the years. Connection to the Internet via these modems is called *dial-up* access. Dial-up was very inconvenient because it tied up the phone line. As the web became more and more interactive, dial-up also hindered usage, as users wanted to transfer more and more data. As a point of reference, downloading a typical 3.5 mb song would take 24 minutes at 1200 bps and 2 minutes at 28,800 bps.

A *broadband* connection is defined as one that has speeds of at least 256,000 bps, though most connections today are much faster, measured in millions of bits per second (megabits or mbps) or even billions (gigabits). For the home user, a broadband connection is usually accomplished via the cable television lines or phone lines (DSL). Both cable and DSL have similar prices and speeds, though each individual may find that one is better than the other for their specific area. Speeds for cable and DSL can vary during different times of the day or week, depending upon how much data traffic is being used. In more remote areas, where cable and phone companies do not provide access, home Internet connections can be made via satellite. The average home broadband speed is anywhere between 3 mbps and 30 mbps. At 10 mbps, downloading a typical 3.5 mb song would take less than a second. For businesses who require more bandwidth and reliability, telecommunications companies can provide other options, such as T1 and T3 lines.

Growth of broadband use (Source: Pew Internet and American Life Project Surveys)

Broadband access is important because it impacts how the Internet is used. When a community has access to broadband, it allows them to interact more online and increases the usage of digital tools overall. Access to broadband is now considered a basic human right by the United Nations, as declared in their 2011 statement:

> "Broadband technologies are fundamentally transforming the way we live," the Broadband Commission for Digital Development, set up last year by the UN Educational Scientific and Cultural Organization (UNESCO) and the UN International Telecommunications Union (ITU), said in issuing "The Broadband Challenge" at a leadership summit in Geneva.
>
> "It is vital that no one be excluded from the new global knowledge societies we are building. We believe that communication is not just a human need – it is a right."[1]

12.5.4: Wireless

Wireless Networking

Today we are used to being able to access the Internet wherever we go. Our smartphones can access the Internet; Starbucks provides wireless “hotspots” for our laptops or iPads. These wireless technologies have made Internet access more convenient and have made devices such as tablets and laptops much more functional. Let’s examine a few of these wireless technologies.

Wi-Fi

Wi-Fi is a technology that takes an Internet signal and converts it into radio waves. These radio waves can be picked up within a radius of approximately 65 feet by devices with a wireless adapter. Several Wi-Fi specifications have been developed over the years, starting with 802.11b (1999), followed by the 802.11g specification in 2003 and 802.11n in 2009. Each new specification improved the speed and range of Wi-Fi, allowing for more uses. One of the primary places where Wi-Fi is being used is in the home. Home users are purchasing Wi-Fi routers, connecting them to their broadband connections, and then connecting multiple devices via Wi-Fi.

Mobile Network

As the cellphone has evolved into the smartphone, the desire for Internet access on these devices has led to data networks being included as part of the mobile phone network. While Internet connections were technically available earlier, it was really with the release of the 3G networks in 2001 (2002 in the US) that smartphones and other cellular devices could access data from the Internet. This new capability drove the market for new and more powerful smartphones, such as the iPhone, introduced in 2007. In 2011, wireless carriers began offering 4G data speeds, giving the cellular networks the same speeds that customers were used to getting via their home connection.

Sidebar: Why Doesn’t My Cellphone Work When I Travel Abroad?

As mobile phone technologies have evolved, providers in different countries have chosen different communication standards for their mobile phone networks. In the US, both of the two competing standards exist: GSM (used by AT&T and T-Mobile) and CDMA (used by the other major carriers). Each standard has its pros and cons, but the bottom line is that phones using one standard cannot easily switch to the other. In the US, this is not a big deal because mobile networks exist to support both standards. But when you travel to other countries, you will find that most of them use GSM networks, with the one big exception being Japan, which has standardized on CDMA. It is possible for a mobile phone using one type of network to switch to the other type of network by switching out the SIM card, which controls your access to the mobile network. However, this will not work in all cases. If you are traveling abroad, it is always best to consult with your mobile provider to determine the best way to access a mobile network.

Bluetooth

While Bluetooth is not generally used to connect a device to the Internet, it is an important wireless technology that has enabled many functionalities that are used every day. When created in 1994 by Ericsson, it was intended to replace wired connections between devices. Today, it is the standard method for connecting nearby devices wirelessly. Bluetooth has a range of approximately 300 feet and consumes very little power, making it an excellent choice for a variety of purposes. Some applications of Bluetooth include: connecting a printer to a personal computer, connecting a mobile phone and headset, connecting a wireless keyboard and mouse to a computer, and connecting a remote for a presentation made on a personal computer.

VoIP

A typical VoIP communication. Image courtesy of BroadVoice.

A growing class of data being transferred over the Internet is voice data. A protocol called voice over IP, or VoIP, enables sounds to be converted to a digital format for transmission over the Internet and then re-created at the other end. By using many existing technologies and software, voice communication over the Internet is now available to anyone with a browser (think Skype, Google Hangouts). Beyond this, many companies are now offering VoIP-based telephone service for business and home use.

12.5.5: LAN / WAN and Client Server

LAN and WAN

Scope of business networks

While the Internet was evolving and creating a way for organizations to connect to each other and the world, another revolution was taking place inside organizations. The proliferation of personal computers inside organizations led to the need to share resources such as printers, scanners, and data. Organizations solved this problem through the creation of local area networks (LANs), which allowed computers to connect to each other and to peripherals. These same networks also allowed personal computers to hook up to legacy mainframe computers.

An LAN is (by definition) a local network, usually operating in the same building or on the same campus. When an organization needed to provide a network over a wider area (with locations in different cities or states, for example), they would build a wide area network (WAN).

Client-Server

The personal computer originally was used as a stand-alone computing device. A program was installed on the computer and then used to do word processing or number crunching. However, with the advent of networking and local area networks, computers could work together to solve problems. Higher-end computers were installed as servers, and users on the local network could run applications and share information among departments and organizations. This is called client-server computing.

12.5.6: Intranet, Extranet, and Cloud

Intranet

Just as organizations set up web sites to provide global access to information about their business, they also set up internal web pages to provide information about the organization to the employees. This internal set of web pages is called an intranet. Web pages on the intranet are not accessible to those outside the company; in fact, those pages would come up as “not found” if an employee tried to access them from outside the company’s network.

Extranet

Sometimes an organization wants to be able to collaborate with its customers or suppliers while at the same time maintaining the security of being inside its own network. In cases like this a company may want to create an *extranet*, which is a part of the company’s network that can be made available securely to those outside of the company. Extranets can be used to allow customers to log in and check the status of their orders, or for suppliers to check their customers’ inventory levels.

Sometimes, an organization will need to allow someone who is not located physically within its internal network to gain access. This access can be provided by a virtual private network (VPN). VPNs will be discussed further in the chapter 6 (on information security).

Sidebar: Microsoft’s SharePoint Powers the Intranet

As organizations begin to see the power of collaboration between their employees, they often look for solutions that will allow them to leverage their intranet to enable more collaboration. Since most companies use Microsoft products for much of their computing, it is only natural that they have looked to Microsoft to provide a solution. This solution is Microsoft’s SharePoint.

SharePoint provides a communication and collaboration platform that integrates seamlessly with Microsoft’s Office suite of applications. Using SharePoint, employees can share a document and edit it together – no more e-mailing that Word document to everyone for review. Projects and documents can be managed collaboratively across the organization. Corporate documents are indexed and made available for search. No more asking around for that procedures document – now you just search for it in SharePoint. For organizations looking to add a social networking component to their intranet, Microsoft offers Yammer, which can be used by itself or integrated into SharePoint.

Cloud Computing

We covered cloud computing in chapter 3, but it should also be mentioned here. The universal availability of the Internet combined with increases in processing power and data-storage capacity have made cloud computing a viable option for many companies. Using cloud computing, companies or individuals can contract to store data on storage devices somewhere on the Internet. Applications can be “rented” as needed, giving a company the ability to quickly deploy new applications. You can read about cloud computing in more detail in chapter 3.

Sidebar: Metcalfe’s Law

Just as Moore’s Law describes how computing power is increasing over time, Metcalfe’s Law describes the power of networking. Specifically, Metcalfe’s Law states that the value of a telecommunications network is proportional to the square of the number of connected users of the system. Think about it this way: If none of your friends were on Facebook, would you spend much time there? If no one else at your school or place of work had e-mail, would it be very useful to you? Metcalfe’s Law tries to quantify this value.

12.5.7: Summary

Summary

The networking revolution has completely changed how the computer is used. Today, no one would imagine using a computer that was not connected to one or more networks. The development of the Internet and World Wide Web, combined with wireless access, has made information available at our fingertips. The Web 2.0 revolution has made us all authors of web content. As networking technology has matured, the use of Internet technologies has become a standard for every type of organization. The use of intranets and extranets has allowed organizations to deploy functionality to employees and business partners alike, increasing efficiencies and improving communications. Cloud computing has truly made information available everywhere and has serious implications for the role of the IT department.

Study Questions

1. What were the first four locations hooked up to the Internet (ARPANET)?
2. What does the term *packet* mean?
3. Which came first, the Internet or the World Wide Web?
4. What was revolutionary about Web 2.0?
5. What was the so-called killer app for the Internet?
6. What makes a connection a *broadband* connection?
7. What does the term VoIP mean?
8. What is an LAN?
9. What is the difference between an intranet and an extranet?
10. What is Metcalfe's Law?

Exercises

1. What is the IP address of your computer? How did you find out? What is the IP address of google.com? How did you find out? Did you get IPv4 or IPv6 addresses?
2. What is the difference between the Internet and the World Wide Web? Create at least three statements that identify the differences between the two.
3. Who are the broadband providers in your area? What are the prices and speeds offered?
4. Pretend you are planning a trip to three foreign countries in the next month. Consult your wireless carrier to determine if your mobile phone would work properly in those countries. What would the costs be? What alternatives do you have if it would not work?

---

1. "UN sets goal of bringing broadband to half developing world's people by 2015.", UN News Center website, http://www.un.org/apps/news/story.as...1#.Ut7JOmTTk1J↵

SECTION OVERVIEW

12.6: Information Systems Security

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- identify the information security triad;
- identify and understand the high-level concepts surrounding information security tools; and
- secure yourself digitally.

Introduction

As computers and other digital devices have become essential to business and commerce, they have also increasingly become a target for attacks. In order for a company or an individual to use a computing device with confidence, they must first be assured that the device is not compromised in any way and that all communications will be secure. In this chapter, we will review the fundamental concepts of information systems security and discuss some of the measures that can be taken to mitigate security threats. We will begin with an overview focusing on how organizations can stay secure. Several different measures that a company can take to improve security will be discussed. We will then follow up by reviewing security precautions that individuals can take in order to secure their personal computing environment.

SECTION OVERVIEW

12.6.1: The Ethical and Legal Implications of Information Systems

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe what the term *information systems ethics* means;
- explain what a code of ethics is and describe the advantages and disadvantages;
- define the term *intellectual property* and explain the protections provided by copyright, patent, and trademark; and
- describe the challenges that information technology brings to individual privacy.

Introduction

Information systems have had an impact far beyond the world of business. New technologies create new situations that we have never dealt with before. How do we handle the new capabilities that these devices empower us with? What new laws are going to be needed to protect us from ourselves? This chapter will kick off with a discussion of the impact of information systems on how we behave (ethics). This will be followed with the new legal structures being put in place, with a focus on intellectual property and privacy.

12.6.2: CIA

The security triad

The Information Security Triad: Confidentiality, Integrity, Availability (CIA)

Confidentiality

When protecting information, we want to be able to restrict access to those who are allowed to see it; everyone else should be disallowed from learning anything about its contents. This is the essence of confidentiality. For example, federal law requires that universities restrict access to private student information. The university must be sure that only those who are authorized have access to view the grade records.

Integrity

Integrity is the assurance that the information being accessed has not been altered and truly represents what is intended. Just as a person with integrity means what he or she says and can be trusted to consistently represent the truth, information integrity means information truly represents its intended meaning. Information can lose its integrity through malicious intent, such as when someone who is not authorized makes a change to intentionally misrepresent something. An example of this would be when a hacker is hired to go into the university's system and change a grade.

Integrity can also be lost unintentionally, such as when a computer power surge corrupts a file or someone authorized to make a change accidentally deletes a file or enters incorrect information.

Availability

Information availability is the third part of the CIA triad. *Availability* means that information can be accessed and modified by anyone authorized to do so in an appropriate timeframe. Depending on the type of information, *appropriate timeframe* can mean different things. For example, a stock trader needs information to be available immediately, while a sales person may be happy to get sales numbers for the day in a report the next morning. Companies such as Amazon.com will require their servers to be available twenty-four hours a day, seven days a week. Other companies may not suffer if their web servers are down for a few minutes once in a while.

12.6.3: Tools to Use

Tools for Information Security

In order to ensure the confidentiality, integrity, and availability of information, organizations can choose from a variety of tools. Each of these tools can be utilized as part of an overall information-security policy, which will be discussed in the next section.

12.6.3.1: Authentication

Authentication

The most common way to identify someone is through their physical appearance, but how do we identify someone sitting behind a computer screen or at the ATM? Tools for authentication are used to ensure that the person accessing the information is, indeed, who they present themselves to be.

Authentication can be accomplished by identifying someone through one or more of three factors: something they know, something they have, or something they are. For example, the most common form of authentication today is the user ID and password. In this case, the authentication is done by confirming something that the user knows (their ID and password). But this form of authentication is easy to compromise (see sidebar) and stronger forms of authentication are sometimes needed. Identifying someone only by something they have, such as a key or a card, can also be problematic. When that identifying token is lost or stolen, the identity can be easily stolen. The final factor, something you are, is much harder to compromise. This factor identifies a user through the use of a physical characteristic, such as an eye-scan or fingerprint. Identifying someone through their physical characteristics is called biometrics.

A more secure way to authenticate a user is to do multi-factor authentication. By combining two or more of the factors listed above, it becomes much more difficult for someone to misrepresent themselves. An example of this would be the use of an RSA SecurID token. The RSA device is something you have, and will generate a new access code every sixty seconds. To log in to an information resource using the RSA device, you combine something you know, a four-digit PIN, with the code generated by the device. The only way to properly authenticate is by both knowing the code *and* having the RSA device.

12.6.3.2: Access Control

Access Control

Once a user has been authenticated, the next step is to ensure that they can only access the information resources that are appropriate. This is done through the use of access control. Access control determines which users are authorized to read, modify, add, and/or delete information. Several different access control models exist. Here we will discuss two: the access control list (ACL) and role-based access control (RBAC).

For each information resource that an organization wishes to manage, a list of users who have the ability to take specific actions can be created. This is an access control list, or ACL. For each user, specific capabilities are assigned, such as *read*, *write*, *delete*, or *add*. Only users with those capabilities are allowed to perform those functions. If a user is not on the list, they have no ability to even know that the information resource exists.

ACLs are simple to understand and maintain. However, they have several drawbacks. The primary drawback is that each information resource is managed separately, so if a security administrator wanted to add or remove a user to a large set of information resources, it would be quite difficult. And as the number of users and resources increase, ACLs become harder to maintain. This has led to an improved method of access control, called role-based access control, or RBAC. With RBAC, instead of giving specific users access rights to an information resource, users are assigned to roles and then those roles are assigned the access. This allows the administrators to manage users and roles separately, simplifying administration and, by extension, improving security.

Comparison of ACL and RBAC (click to enlarge)

12.6.3.3: Encryption

Encryption

Many times, an organization needs to transmit information over the Internet or transfer it on external media such as a CD or flash drive. In these cases, even with proper authentication and access control, it is possible for an unauthorized person to get access to the data. Encryption is a process of encoding data upon its transmission or storage so that only authorized individuals can read it. This encoding is accomplished by a computer program, which encodes the plain text that needs to be transmitted; then the recipient receives the cipher text and decodes it (decryption). In order for this to work, the sender and receiver need to agree on the method of encoding so that both parties can communicate properly. Both parties share the encryption key, enabling them to encode and decode each other's messages. This is called symmetric key encryption. This type of encryption is problematic because the key is available in two different places.

An alternative to symmetric key encryption is public key encryption. In public key encryption, two keys are used: a public key and a private key. To send an encrypted message, you obtain the public key, encode the message, and send it. The recipient then uses the private key to decode it. The public key can be given to anyone who wishes to send the recipient a message. Each user simply needs one private key and one public key in order to secure messages. The private key is necessary in order to decrypt something sent with the public key.

Public key encryption (click for larger diagram)

Sidebar: Password Security

So why is using just a simple user ID/password not considered a secure method of authentication? It turns out that this single-factor authentication is extremely easy to compromise. Good password policies must be put in place in order to ensure that passwords cannot be compromised. Below are some of the more common policies that organizations should put in place.

- Require complex passwords. One reason passwords are compromised is that they can be easily guessed. A recent study found that the top three passwords people used in 2012 were *password*, *123456* and *12345678*.[1] A password should not be simple, or a word that can be found in a dictionary. One of the first things a hacker will do is try to crack a password by testing every term in the dictionary! Instead, a good password policy is one that requires the use of a minimum of eight characters, and at least one upper-case letter, one special character, and one number.
- Change passwords regularly. It is essential that users change their passwords on a regular basis. Users should change their passwords every sixty to ninety days, ensuring that any passwords that might have been stolen or guessed will not be able to be used against the company.
- Train employees not to give away passwords. One of the primary methods that is used to steal passwords is to simply figure them out by asking the users or administrators. *Pretexting* occurs when an attacker calls a helpdesk or security administrator and pretends to be a particular authorized user having trouble logging in. Then, by providing some personal information about the authorized user, the attacker convinces the security person to reset the password and tell him what it is. Another way that employees may be tricked into giving away passwords is through e-mail phishing. Phishing occurs when a user receives an e-mail that looks as if it is from a trusted source, such as their bank, or their employer. In the e-mail, the user is asked to click a link and log in to a website that mimics the genuine website and enter their ID and password, which are then captured by the attacker.

12.6.3.4: Backups

Backups

Another essential tool for information security is a comprehensive backup plan for the entire organization. Not only should the data on the corporate servers be backed up, but individual computers used throughout the organization should also be backed up. A good backup plan should consist of several components.

- A full understanding of the organizational information resources. What information does the organization actually have? Where is it stored? Some data may be stored on the organization's servers, other data on users' hard drives, some in the cloud, and some on third-party sites. An organization should make a full inventory of all of the information that needs to be backed up and determine the best way back it up.
- Regular backups of all data. The frequency of backups should be based on how important the data is to the company, combined with the ability of the company to replace any data that is lost. Critical data should be backed up daily, while less critical data could be backed up weekly.
- Offsite storage of backup data sets. If all of the backup data is being stored in the same facility as the original copies of the data, then a single event, such as an earthquake, fire, or tornado, would take out both the original data and the backup! It is essential that part of the backup plan is to store the data in an offsite location.
- Test of data restoration. On a regular basis, the backups should be put to the test by having some of the data restored. This will ensure that the process is working and will give the organization confidence in the backup plan.

Besides these considerations, organizations should also examine their operations to determine what effect downtime would have on their business. If their information technology were to be unavailable for any sustained period of time, how would it impact the business?

Additional concepts related to backup include the following:

- Universal Power Supply (UPS). A UPS is a device that provides battery backup to critical components of the system, allowing them to stay online longer and/or allowing the IT staff to shut them down using proper procedures in order to prevent the data loss that might occur from a power failure.
- Alternate, or "hot" sites. Some organizations choose to have an alternate site where an exact replica of their critical data is always kept up to date. When the primary site goes down, the alternate site is immediately brought online so that little or no downtime is experienced.

As information has become a strategic asset, a whole industry has sprung up around the technologies necessary for implementing a proper backup strategy. A company can contract with a service provider to back up all of their data or they can purchase large amounts of online storage space and do it themselves. Technologies such as storage area networks and archival systems are now used by most large businesses.

12.6.4: Firewalls

Firewalls

Network configuration with firewalls, IDS, and a DMZ. Click to enlarge.

Another method that an organization should use to increase security on its network is a firewall. A firewall can exist as hardware or software (or both). A hardware firewall is a device that is connected to the network and filters the packets based on a set of rules. A software firewall runs on the operating system and intercepts packets as they arrive to a computer. A firewall protects all company servers and computers by stopping packets from outside the organization's network that do not meet a strict set of criteria. A firewall may also be configured to restrict the flow of packets leaving the organization. This may be done to eliminate the possibility of employees watching YouTube videos or using Facebook from a company computer.

Some organizations may choose to implement multiple firewalls as part of their network security configuration, creating one or more sections of their network that are partially secured. This segment of the network is referred to as a DMZ, borrowing the term *demilitarized zone* from the military, and it is where an organization may place resources that need broader access but still need to be secured.

12.6.5: IDS

Intrusion Detection Systems

Another device that can be placed on the network for security purposes is an intrusion detection system, or IDS. An IDS does not add any additional security; instead, it provides the functionality to identify if the network is being attacked. An IDS can be configured to watch for specific types of activities and then alert security personnel if that activity occurs. An IDS also can log various types of traffic on the network for analysis later. An IDS is an essential part of any good security setup.

Sidebar: Virtual Private Networks

Using firewalls and other security technologies, organizations can effectively protect many of their information resources by making them invisible to the outside world. But what if an employee working from home requires access to some of these resources? What if a consultant is hired who needs to do work on the internal corporate network from a remote location? In these cases, a virtual private network (VPN) is called for.

A VPN allows a user who is outside of a corporate network to take a detour around the firewall and access the internal network from the outside. Through a combination of software and security measures, this lets an organization allow limited access to its networks while at the same time ensuring overall security.

12.6.6: Physical Security

Physical Security

An organization can implement the best authentication scheme in the world, develop the best access control, and install firewalls and intrusion prevention, but its security cannot be complete without implementation of physical security. Physical security is the protection of the actual hardware and networking components that store and transmit information resources. To implement physical security, an organization must identify all of the vulnerable resources and take measures to ensure that these resources cannot be physically tampered with or stolen. These measures include the following.

- Locked doors: It may seem obvious, but all the security in the world is useless if an intruder can simply walk in and physically remove a computing device. High-value information assets should be secured in a location with limited access.
- Physical intrusion detection: High-value information assets should be monitored through the use of security cameras and other means to detect unauthorized access to the physical locations where they exist.
- Secured equipment: Devices should be locked down to prevent them from being stolen. One employee's hard drive could contain all of your customer information, so it is essential that it be secured.
- Environmental monitoring: An organization's servers and other high-value equipment should always be kept in a room that is monitored for temperature, humidity, and airflow. The risk of a server failure rises when these factors go out of a specified range.
- Employee training: One of the most common ways thieves steal corporate information is to steal employee laptops while employees are traveling. Employees should be trained to secure their equipment whenever they are away from the office.

12.6.7: Security Policies

Security Policies

Besides the technical controls listed above, organizations also need to implement security policies as a form of administrative control. In fact, these policies should really be a starting point in developing an overall security plan. A good information-security policy lays out the guidelines for employee use of the information resources of the company and provides the company recourse in the case that an employee violates a policy.

According to the SANS Institute, a good policy is “a formal, brief, and high-level statement or plan that embraces an organization’s general beliefs, goals, objectives, and acceptable procedures for a specified subject area.” Policies require compliance; failure to comply with a policy will result in disciplinary action. A policy does not lay out the specific technical details, instead it focuses on the desired results. A security policy should be based on the guiding principles of confidentiality, integrity, and availability.[2]

A good example of a security policy that many will be familiar with is a web use policy. A web use policy lays out the responsibilities of company employees as they use company resources to access the Internet. A good example of a web use policy is included in Harvard University’s “Computer Rules and Responsibilities” policy, which can be found here.

A security policy should also address any governmental or industry regulations that apply to the organization. For example, if the organization is a university, it must be aware of the Family Educational Rights and Privacy Act (FERPA), which restricts who has access to student information. Health care organizations are obligated to follow several regulations, such as the Health Insurance Portability and Accountability Act (HIPAA).

A good resource for learning more about security policies is the SANS Institute’s Information Security Policy Page.

Usability

When looking to secure information resources, organizations must balance the need for security with users’ need to effectively access and use these resources. If a system’s security measures make it difficult to use, then users will find ways around the security, which may make the system more vulnerable than it would have been without the security measures! Take, for example, password policies. If the organization requires an extremely long password with several special characters, an employee may resort to writing it down and putting it in a drawer since it will be impossible to memorize.

12.6.7.1: Mobile Secuirty

Sidebar: Mobile Security

As the use of mobile devices such as smartphones and tablets proliferates, organizations must be ready to address the unique security concerns that the use of these devices bring. One of the first questions an organization must consider is whether to allow mobile devices in the workplace at all. Many employees already have these devices, so the question becomes: Should we allow employees to bring their own devices and use them as part of their employment activities? Or should we provide the devices to our employees? Creating a BYOD ("Bring Your Own Device") policy allows employees to integrate themselves more fully into their job and can bring higher employee satisfaction and productivity. In many cases, it may be virtually impossible to prevent employees from having their own smartphones or iPads in the workplace. If the organization provides the devices to its employees, it gains more control over use of the devices, but it also exposes itself to the possibility of an administrative (and costly) mess.

Mobile devices can pose many unique security challenges to an organization. Probably one of the biggest concerns is theft of intellectual property. For an employee with malicious intent, it would be a very simple process to connect a mobile device either to a computer via the USB port, or wirelessly to the corporate network, and download confidential data. It would also be easy to secretly take a high-quality picture using a built-in camera.

When an employee does have permission to access and save company data on his or her device, a different security threat emerges: that device now becomes a target for thieves. Theft of mobile devices (in this case, including laptops) is one of the primary methods that data thieves use.

So what can be done to secure mobile devices? It will start with a good policy regarding their use. According to a 2013 SANS study, organizations should consider developing a mobile device policy that addresses the following issues: use of the camera, use of voice recording, application purchases, encryption at rest, Wi-Fi autoconnect settings, bluetooth settings, VPN use, password settings, lost or stolen device reporting, and backup. [3]

Besides policies, there are several different tools that an organization can use to mitigate some of these risks. For example, if a device is stolen or lost, geolocation software can help the organization find it. In some cases, it may even make sense to install remote data-removal software, which will remove data from a device if it becomes a security risk.

12.6.8: Personal info Sec

Personal Information Security

Poster from Stop. Think. Connect. Click to enlarge. (Copyright: Stop. Think. Connect. http://stopthinkconnect.org/resources)

We will end this chapter with a discussion of what measures each of us, as individual users, can take to secure our computing technologies. There is no way to have 100% security, but there are several simple steps we, as individuals, can take to make ourselves more secure.

- Keep your software up to date. Whenever a software vendor determines that a security flaw has been found in their software, they will release an update to the software that you can download to fix the problem. Turn on automatic updating on your computer to automate this process.
- Install antivirus software and keep it up to date. There are many good antivirus software packages on the market today, including free ones.
- Be smart about your connections. You should be aware of your surroundings. When connecting to a Wi-Fi network in a public place, be aware that you could be at risk of being spied on by others sharing that network. It is advisable not to access your financial or personal data while attached to a Wi-Fi hotspot. You should also be aware that connecting USB flash drives to your device could also put you at risk. Do not attach an unfamiliar flash drive to your device unless you can scan it first with your security software.
- Back up your data. Just as organizations need to back up their data, individuals need to as well. And the same rules apply: do it regularly and keep a copy of it in another location. One simple solution for this is to set up an account with an online backup service, such as Mozy or Carbonite, to automate your backups.
- Secure your accounts with two-factor authentication. Most e-mail and social media providers now have a two-factor authentication option. The way this works is simple: when you log in to your account from an unfamiliar computer for the first time, it sends you a text message with a code that you must enter to confirm that you are really you. This means that no one else can log in to your accounts without knowing your password *and* having your mobile phone with them.
- Make your passwords long, strong, and unique. For your personal passwords, you should follow the same rules that are recommended for organizations. Your passwords should be long (eight or more characters) and contain at least two of the following: upper-case letters, numbers, and special characters. You also should use different passwords for different accounts, so that if someone steals your password for one account, they still are locked out of your other accounts.
- Be suspicious of strange links and attachments. When you receive an e-mail, tweet, or Facebook post, be suspicious of any links or attachments included there. Do not click on the link directly if you are at all suspicious. Instead, if you want to access the website, find it yourself and navigate to it directly.

You can find more about these steps and many other ways to be secure with your computing by going to Stop. Think. Connect. This website is part of a campaign that was launched in October of 2010 by the STOP. THINK. CONNECT. Messaging Convention in partnership with the U.S. government, including the White House.

12.6.S: Summary

Summary

As computing and networking resources have become more and more an integral part of business, they have also become a target of criminals. Organizations must be vigilant with the way they protect their resources. The same holds true for us personally: as digital devices become more and more intertwined with our lives, it becomes crucial for us to understand how to protect ourselves.

Study Questions

1. Briefly define each of the three members of the information security triad.
2. What does the term *authentication* mean?
3. What is multi-factor authentication?
4. What is role-based access control?
5. What is the purpose of encryption?
6. What are two good examples of a complex password?
7. What is pretexting?
8. What are the components of a good backup plan?
9. What is a firewall?
10. What does the term *physical security* mean?

Exercises

1. Describe one method of multi-factor authentication that you have experienced and discuss the pros and cons of using multi-factor authentication.
2. What are some of the latest advances in encryption technologies? Conduct some independent research on encryption using scholarly or practitioner resources, then write a two- to three-page paper that describes at least two new advances in encryption technology.
3. What is the password policy at your place of employment or study? Do you have to change passwords every so often? What are the minimum requirements for a password?
4. When was the last time you backed up your data? What method did you use? In one to two pages, describe a method for backing up your data. Ask your instructor if you can get extra credit for backing up your data.
5. Find the information security policy at your place of employment or study. Is it a good policy? Does it meet the standards outlined in the chapter?
6. How are you doing on keeping your own information secure? Review the steps listed in the chapter and comment on how well you are doing.

---

1. "Born to be breached" by Sean Gallagher on Nov 3 2012. *Arstechnica*. Retrieved from http://arstechnica.com/information-t...e-most-common/ on May 15, 2013.↵
2. SANS Institute. "A Short Primer for Developing Security Policies." Accessed from http://www.sans.org/security-resourc...icy_Primer.pdf on May 31, 2013.↵
3. Taken from SANS Institute's Mobile Device Checklist. You can review the full checklist at www.sans.org/score/checklists/mobile-device-checklist.xls. ↵

SECTION OVERVIEW

12.7: Does IT Matter?

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- define the productivity paradox and explain the current thinking on this topic;
- evaluate Carr's argument in "Does IT Matter?";
- describe the components of competitive advantage; and
- describe information systems that can provide businesses with competitive advantage.

Introduction

For over fifty years, computing technology has been a part of business. Organizations have spent trillions of dollars on information technologies. But has all this investment in IT made a difference? Have we seen increases in productivity? Are companies that invest in IT more competitive? In this chapter, we will look at the value IT can bring to an organization and try to answer these questions. We will begin by highlighting two important works from the past two decades.

12.7.1: The Paradox

The Productivity Paradox

In 1991, Erik Brynjolfsson wrote an article, published in the *Communications of the ACM*, entitled "The Productivity Paradox of Information Technology: Review and Assessment." By reviewing studies about the impact of IT investment on productivity, Brynjolfsson was able to conclude that the addition of information technology to business had not improved productivity at all – the "productivity paradox." From the article[1] He does not draw any specific conclusions from this finding, and provides the following analysis:

Although it is too early to conclude that IT's productivity contribution has been subpar, a paradox remains in our inability to unequivocally document any contribution after so much effort. The various explanations that have been proposed can be grouped into four categories:

1) Mismeasurement of outputs and inputs,

2) Lags due to learning and adjustment,

3) Redistribution and dissipation of profits,

4) Mismanagement of information and technology.

In 1998, Brynjolfsson and Lorin Hitt published a follow-up paper entitled " Beyond the Productivity Paradox ."[2] In this paper, the authors utilized new data that had been collected and found that IT did, indeed, provide a positive result for businesses. Further, they found that sometimes the true advantages in using technology were not directly relatable to higher productivity, but to "softer" measures, such as the impact on organizational structure. They also found that the impact of information technology can vary widely between companies.

12.7.2: No - IT does NOT Matter

IT Doesn't Matter

Just as a consensus was forming about the value of IT, the Internet stock market bubble burst. Just two years later, in 2003, Harvard professor Nicholas Carr wrote his article "IT Doesn't Matter" in the *Harvard Business Review.* In this article Carr asserts that as information technology has become more ubiquitous, it has also become less of a differentiator. In other words: because information technology is so readily available and the software used so easily copied, businesses cannot hope to implement these tools to provide any sort of competitive advantage. Carr goes on to suggest that since IT is essentially a commodity, it should be managed like one: low cost, low risk. Using the analogy of electricity, Carr describes how a firm should never be the first to try a new technology, thereby letting others take the risks. IT management should see themselves as a utility within the company and work to keep costs down . For IT, providing the best service with minimal downtime is the goal.

As you can imagine, this article caused quite an uproar, especially from IT companies. Many articles were written in defense of IT; many others in support of Carr. Carr released a book based on the article in 2004, entitled *Does IT Matter?* Click here to watch a video of Carr being interviewed about his book on CNET.

Probably the best thing to come out of the article and subsequent book was that it opened up discussion on the place of IT in a business strategy, and exactly what role IT could play in competitive advantage. It is that question that we want to address in the rest of the this chapter.

[Embed "IT Doesn't Matter" classroom video footage here: http://quickstream.biola.edu/distanc...n'tMatter.f4v]

12.7.3: The Value Chain

The Value Chain

Porter's value chain (click to enlarge)

In his book, Porter describes exactly how a company can create value (and therefore profit). Value is built through the value chain: a series of activities undertaken by the company to produce a product or service. Each step in the value chain contributes to the overall value of a product or service. While the value chain may not be a perfect model for every type of company, it does provide a way to analyze just how a company is producing value. The value chain is made up of two sets of activities: primary activities and support activities. We will briefly examine these activities and discuss how information technology can play a role in creating value by contributing to cost advantage or differentiation advantage, or both.

The primary activities are the functions that directly impact the creation of a product or service. The goal of the primary activities is to add more value than they cost. The primary activities are:

- Inbound logistics: These are the functions performed to bring in raw materials and other needed inputs. Information technology can be used here to make these processes more efficient, such as with supply-chain management systems, which allow the suppliers to manage their own inventory.
- Operations: Any part of a business that is involved in converting the raw materials into the final products or services is part of operations. From manufacturing to business process management (covered in chapter 8), information technology can be used to provide more efficient processes and increase innovation through flows of information.
- Outbound logistics: These are the functions required to get the product out to the customer. As with inbound logistics, IT can be used here to improve processes, such as allowing for real-time inventory checks. IT can also be a delivery mechanism itself.
- Sales/Marketing: The functions that will entice buyers to purchase the products are part of sales and marketing. Information technology is used in almost all aspects of this activity. From online advertising to online surveys, IT can be used to innovate product design and reach customers like never before. The company website can be a sales channel itself.
- Service: The functions a business performs after the product has been purchased to maintain and enhance the product's value are part of the service activity. Service can be enhanced via technology as well, including support services through websites and knowledge bases.

The support activities are the functions in an organization that support, and cut across, all of the primary activities. The support activities are:

- Firm infrastructure: This includes organizational functions such as finance, accounting, and quality control, all of which depend on information technology; the use of ERP systems (to be covered in chapter 9) is a good example of the impact that IT can have on these functions.
- Human resource management: This activity consists of recruiting, hiring, and other services needed to attract and retain employees. Using the Internet, HR departments can increase their reach when looking for candidates. There is also the possibility of allowing employees to use technology for a more flexible work environment.
- Technology development: Here we have the technological advances and innovations that support the primary activities. These advances are then integrated across the firm or within one of the primary activities to add value. Information technology would fall specifically under this activity.
- Procurement: The activities involved in acquiring the raw materials used in the creation of products and services are called procurement. Business-to-business e-commerce can be used to improve the acquisition of materials.

This analysis of the value chain provides some insight into how information technology can lead to competitive advantage. Let's now look at another tool that Porter developed – the "five forces" model.

12.7.4: Porter's Five Forces

Porter's Five Forces

Porter's five forces (click to enlarge)

Porter developed the "five forces" model as a framework for industry analysis. This model can be used to help understand just how competitive an industry is and to analyze its strengths and weaknesses. The model consists of five elements, each of which plays a role in determining the average profitability of an industry. In 2001, Porter wrote an article entitled "Strategy and the Internet," in which he takes this model and looks at how the Internet impacts the profitability of an industry. Below is a quick summary of each of the five forces and the impact of the Internet.

- **Threat of substitute products or services:** How easily can a product or service be replaced with something else? The more types of products or services there are that can meet a particular need, the less profitability there will be in an industry. For example, the advent of the mobile phone has replaced the need for pagers. The Internet has made people more aware of substitute products, driving down industry profits in those industries being substituted.
- **Bargaining power of suppliers:** When a company has several suppliers to choose from, it can demand a lower price. When a sole supplier exists, then the company is at the mercy of the supplier. For example, if only one company makes the controller chip for a car engine, that company can control the price, at least to some extent. The Internet has given companies access to more suppliers, driving down prices. On the other hand, suppliers now also have the ability to sell directly to customers.
- **Bargaining power of customers:** A company that is the sole provider of a unique product has the ability to control pricing. But the Internet has given customers many more options to choose from.
- B**arriers to entry:** The easier it is to enter an industry, the tougher it will be to make a profit in that industry. The Internet has an overall effect of making it easier to enter industries. It is also very easy to copy technology, so new innovations will not last that long.
- **Rivalry among existing competitors:** The more competitors there are in an industry, the bigger a factor price becomes. The advent of the Internet has increased competition by widening the geographic market and lowering the costs of doing business. For example, a manufacturer in Southern California may now have to compete against a manufacturer in the South, where wages are lower.

Porter's five forces are used to analyze an industry to determine the average profitability of a company within that industry. Adding in Porter's analysis of the Internet, we can see that the Internet (and by extension, information technology in general) has the effect of lowering overall profitability. [3] While the Internet has certainly produced many companies that are big winners, the overall winners have been the consumers, who have been given an ever-increasing market of products and services and lower prices.

12.7.5: Investing in IT

Investing in IT for Competitive Advantage

In 2008, Brynjolfsson and McAfee published a study in the *Harvard Business Review* on the role of IT in competitive advantage, entitled "Investing in the IT That Makes a Competitive Difference." Their study confirmed that IT *can* play a role in competitive advantage, if deployed wisely. In their study, they draw three conclusions[6]:

- First, the data show that IT has sharpened differences among companies instead of reducing them. This reflects the fact that while companies have always varied widely in their ability to select, adopt, and exploit innovations, technology has accelerated and amplified these differences.
- Second, good management matters: Highly qualified vendors, consultants, and IT departments might be necessary for the successful implementation of enterprise technologies themselves, but the real value comes from the process innovations that can now be delivered on those platforms. Fostering the right innovations and propagating them widely are both executive responsibilities – ones that can't be delegated.
- Finally, the competitive shakeup brought on by IT is not nearly complete, even in the IT-intensive US economy. We expect to see these altered competitive dynamics in other countries, as well, as their IT investments grow.

Information systems can be used for competitive advantage, but they must be used strategically. Organizations must understand how they want to differentiate themselves and then use all the elements of information systems (hardware, software, data, people, and process) to accomplish that differentiation.

12.7.6: Summary

Summary

Information systems are integrated into all components of business today, but can they bring competitive advantage? Over the years, there have been many answers to this question. Early research could not draw any connections between IT and profitability, but later research has shown that the impact can be positive. IT is not a panacea; just purchasing and installing the latest technology will not, by itself, make a company more successful. Instead, the combination of the right technologies and good management, together, will give a company the best chance of a positive result.

Study Questions

1. What is the productivity paradox?
2. Summarize Carr's argument in "Does IT Matter."
3. How is the 2008 study by Brynjolfsson and McAfee different from previous studies? How is it the same?
4. What does it mean for a business to have a competitive advantage?
5. What are the primary activities and support activities of the value chain?
6. What has been the overall impact of the Internet on industry profitability? Who has been the true winner?
7. How does EDI work?
8. Give an example of a semi-structured decision and explain what inputs would be necessary to provide assistance in making the decision.
9. What does a collaborative information system do?
10. How can IT play a role in competitive advantage, according to the 2008 article by Brynjolfsson and McAfee?

Exercises

1. Do some independent research on Nicholas Carr (the author of "IT Doesn't Matter") and explain his current position on the ability of IT to provide competitive advantage.
2. Review the WebEx website. What features of WebEx would contribute to good collaboration? What makes WebEx a better collaboration tool than something like Skype or Google Hangouts?
3. Think of a semi-structured decision that you make in your daily life and build your own DSS using a spreadsheet that would help you make that decision.

1. Brynjolfsson, Erik. "The Productivity Paradox of Information Technology: Review and Assessment" Copyright © 1993, 1994 Erik Brynjolfsson, All Rights Reserved Center for Coordination Science MIT Sloan School of Management Cambridge, Massachusetts Previous version: December 1991 This Version: September 1992 Published in Communications of the ACM, December, 1993; and Japan Management Research, June, 1994 (in Japanese).↵
2. Brynjolfsson, Erik and Lorin Hitt. "Beyond the Productivity Paradox", Communications of the ACM, August 1998, Vol. 41(8): pp. 49–55. Copyright © 1998 by Association for Computing Machinery, Inc. (ACM). ↵
3. Porter, Michael. "Strategy and the Internet," Harvard Business Review, Vol. 79, No. 3, March 2001. http://hbswk.hbs.edu/item/2165.html↵
4. Wiseman, C., & MacMillan, I. C. (1984). CREATING COMPETITIVE WEAPONS FROM INFORMATION SYSTEMS. Journal Of Business Strategy, 5(2), 42.↵
5. Taken from http://www.isabelhealthcare.com/home/ourmission. Accessed July 15, 2013. ↵
6. McAfee, Andrew and Brynjolfsson, Erik "Investing in the IT That Makes a Competitive Difference" Harvard Business Review, (July-August, 2008)↵

SECTION OVERVIEW

12.8: Business Processes

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- define the term *business process*;
- identify the different systems needed to support business processes in an organization;
- explain the value of an enterprise resource planning (ERP) system;
- explain how business process management and business process reengineering work; and
- understand how information technology combined with business processes can bring an organization competitive advantage.

Introduction

The fourth component of information systems is *process*. But what is a process and how does it tie into information systems? And in what ways do processes have a role in business? This chapter will look to answer those questions and also describe how business processes can be used for strategic advantage.

12.8.1: Business Process

What Is a Business Process?

We have all heard the term *process* before, but what exactly does it mean? A process is a series of tasks that are completed in order to accomplish a goal. A business process, therefore, is a process that is focused on achieving a goal for a business. If you have worked in a business setting, you have participated in a business process. Anything from a simple process for making a sandwich at Subway to building a space shuttle utilizes one or more business processes.

Processes are something that businesses go through every day in order to accomplish their mission. The better their processes, the more effective the business. Some businesses see their processes as a strategy for achieving competitive advantage. A process that achieves its goal in a unique way can set a company apart. A process that eliminates costs can allow a company to lower its prices (or retain more profit).

12.8.2: Documenting a Process

Documenting a Process

Every day, each of us will conduct many processes without even thinking about them: getting ready for work, using an ATM, reading our e-mail, etc. But as processes grow more complex, they need to be documented. For businesses, it is essential to do this, because it allows them to ensure control over how activities are undertaken in their organization. It also allows for standardization: McDonald’s has the same process for building a Big Mac in all of its restaurants.

The simplest way to document a process is to simply create a list. The list shows each step in the process; each step can be checked off upon completion. For example, a simple process, such as how to create an account on eBay, might look like this:

1. Go to ebay.com.
2. Click on “register.”
3. Enter your contact information in the “Tell us about you” box.
4. Choose your user ID and password.
5. Agree to User Agreement and Privacy Policy by clicking on “Submit.”

For processes that are not so straightforward, documenting the process as a checklist may not be sufficient. For example, here is the process for determining if an article for a term needs to be added to Wikipedia:

1. Search Wikipedia to determine if the term already exists.
2. If the term is found, then an article is already written, so you must think of another term. Go to 1.
3. If the term is not found, then look to see if there is a related term.
4. If there is a related term, then create a redirect.
5. If there is not a related term, then create a new article.

This procedure is relatively simple – in fact, it has the same number of steps as the previous example – but because it has some decision points, it is more difficult to track with as a simple list. In these cases, it may make more sense to use a diagram to document the process:

Process diagram for determining if a new term should be added to Wikipedia (click to enlarge). (Public Domain)

12.8.3: Managing Documentation

Managing Business Process Documentation

As organizations begin to document their processes, it becomes an administrative task to keep track of them. As processes change and improve, it is important to know which processes are the most recent. It is also important to manage the process so that it can be easily updated! The requirement to manage process documentation has been one of the driving forces behind the creation of the *document management system*. A document management system stores and tracks documents and supports the following functions:

- Versions and timestamps. The document management system will keep multiple versions of documents. The most recent version of a document is easy to identify and will be served up by default.
- Approvals and workflows. When a process needs to be changed, the system will manage both access to the documents for editing and the routing of the document for approvals.
- Communication. When a process changes, those who implement the process need to be made aware of the changes. A document management system will notify the appropriate people when a change to a document is approved.

Of course, document management systems are not only used for managing business process documentation. Many other types of documents are managed in these systems, such as legal documents or design documents.

12.8.4: ERP Systems

ERP Systems

An enterprise resource planning (ERP) system is a software application with a centralized database that can be used to run an entire company. Let's take a closer look at the definition of each of these components:

- An ERP system (click to enlarge)

  A software application: The system is a software application, which means that it has been developed with specific logic and rules behind it. It has to be installed and configured to work specifically for an individual organization.
- With a centralized database: All data in an ERP system is stored in a single, central database. This centralization is key to the success of an ERP – data entered in one part of the company can be immediately available to other parts of the company.
- That can be used to run an entire company: An ERP can be used to manage an entire organization's operations. If they so wish, companies can purchase modules for an ERP that represent different functions within the organization, such as finance, manufacturing, and sales. Some companies choose to purchase many modules, others choose a subset of the modules.

An ERP system not only centralizes an organization's data, but the processes it enforces are the processes the organization adopts. When an ERP vendor designs a module, it has to implement the rules for the associated business processes. A selling point of an ERP system is that it has best practices built right into it. In other words, when an organization implements an ERP, it also gets improved best practices as part of the deal!

For many organizations, the implementation of an ERP system is an excellent opportunity to improve their business practices and upgrade their software at the same time. But for others, an ERP brings them a challenge: Is the process embedded in the ERP really better than the process they are currently utilizing? And if they implement this ERP, and it happens to be the same one that all of their competitors have, will they simply become more like them, making it much more difficult to differentiate themselves?

Registered trademark of SAP

This has been one of the criticisms of ERP systems: that they commoditize business processes, driving all businesses to use the same processes and thereby lose their uniqueness. The good news is that ERP systems also have the capability to be configured with custom processes. For organizations that want to continue using their own processes or even design new ones, ERP systems offer ways to support this through the use of customizations.

But there is a drawback to customizing an ERP system: organizations have to maintain the changes themselves. Whenever an update to the ERP system comes out, any organization that has created a custom process will be required to add that change to their ERP. This will require someone to maintain a listing of these changes and will also require retesting the system every time an upgrade is made. Organizations will have to wrestle with this decision: When should they go ahead and accept the best-practice processes built into the ERP system and when should they spend the resources to develop their own processes? It makes the most sense to only customize those processes that are critical to the competitive advantage of the company.

Some of the best-known ERP vendors are SAP, Microsoft, and Oracle.

12.8.5: Process Management

Business Process Management

Organizations that are serious about improving their business processes will also create structures to manage those processes. Business process management (BPM) can be thought of as an intentional effort to plan, document, implement, and distribute an organization's business processes with the support of information technology.

BPM is more than just automating some simple steps. While automation can make a business more efficient, it cannot be used to provide a competitive advantage. BPM, on the other hand, can be an integral part of creating that advantage.

Not all of an organization's processes should be managed this way. An organization should look for processes that are essential to the functioning of the business and those that may be used to bring a competitive advantage. The best processes to look at are those that include employees from multiple departments, those that require decision-making that cannot be easily automated, and processes that change based on circumstances.

To make this clear, let's take a look at an example.

Suppose a large clothing retailer is looking to gain a competitive advantage through superior customer service. As part of this, they create a task force to develop a state-of-the-art returns policy that allows customers to return any article of clothing, no questions asked. The organization also decides that, in order to protect the competitive advantage that this returns policy will bring, they will develop their own customization to their ERP system to implement this returns policy. As they prepare to roll out the system, they invest in training for all of their customer-service employees, showing them how to use the new system and specifically how to process returns. Once the updated returns process is implemented, the organization will be able to measure several key indicators about returns that will allow them to adjust the policy as needed. For example, if they find that many women are returning their high-end dresses after wearing them once, they could implement a change to the process that limits – to, say, fourteen days – the time after the original purchase that an item can be returned. As changes to the returns policy are made, the changes are rolled out via internal communications, and updates to the returns processing on the system are made. In our example, the system would no longer allow a dress to be returned after fourteen days without an approved reason.

If done properly, business process management will provide several key benefits to an organization, which can be used to contribute to competitive advantage. These benefits include:

- Empowering employees. When a business process is designed correctly and supported with information technology, employees will be able to implement it on their own authority. In our returns-policy example, an employee would be able to accept returns made before fourteen days or use the system to make determinations on what returns would be allowed after fourteen days.
- Built-in reporting. By building measurement into the programming, the organization can keep up to date on key metrics regarding their processes. In our example, these can be used to improve the returns process and also, ideally, to reduce returns.
- Enforcing best practices. As an organization implements processes supported by information systems, it can work to implement the best practices for that class of business process. In our example, the organization may want to require that all customers returning a product without a receipt show a legal ID. This requirement can be built into the system so that the return will not be processed unless a valid ID number is entered.
- Enforcing consistency. By creating a process and enforcing it with information technology, it is possible to create a consistency across the entire organization. In our example, all stores in the retail chain can enforce the same returns policy. And if the returns policy changes, the change can be instantly enforced across the entire chain.

12.8.6: Process Re-engineering

Business Process Re-engineering

As organizations look to manage their processes to gain a competitive advantage, they also need to understand that their existing ways of doing things may not be the most effective or efficient. A process developed in the 1950s is not going to be better just because it is now supported by technology.

In 1990, Michael Hammer published an article in the *Harvard Business Review* entitled “Reengineering Work: Don’t Automate, Obliterate.” This article put forward the thought that simply automating a bad process does not make it better. Instead, companies should “blow up” their existing processes and develop new processes that take advantage of the new technologies and concepts. He states in the introduction to the article:[1]

Many of our job designs, work flows, control mechanisms, and organizational structures came of age in a different competitive environment and before the advent of the computer. They are geared towards greater efficiency and control. Yet the watchwords of the new decade are innovation and speed, service, and quality.

It is time to stop paving the cow paths. Instead of embedding outdated processes in silicon and software, we should obliterate them and start over. We should “reengineer” our businesses: use the power of modern information technology to radically redesign our business processes in order to achieve dramatic improvements in their performance.

Business process reengineering is not just taking an existing process and automating it. BPR is fully understanding the goals of a process and then dramatically redesigning it from the ground up to achieve dramatic improvements in productivity and quality. But this is easier said than done. Most of us think in terms of how to do small, local improvements to a process; complete redesign requires thinking on a larger scale. Hammer provides some guidelines for how to go about doing business process reengineering:

- Organize around outcomes, not tasks. This simply means to design the process so that, if possible, one person performs all the steps. Instead of repeating one step in the process over and over, the person stays involved in the process from start to finish.
- Have those who use the outcomes of the process perform the process. Using information technology, many simple tasks are now automated, so we can empower the person who needs the outcome of the process to perform it. The example Hammer gives here is purchasing: instead of having every department in the company use a purchasing department to order supplies, have the supplies ordered directly by those who need the supplies using an information system.
- Subsume information-processing work into the real work that produces the information. When one part of the company creates information (like sales information, or payment information), it should be processed by that same department. There is no need for one part of the company to process information created in another part of the company.
- Treat geographically dispersed resources as though they were centralized. With the communications technologies in place today, it becomes easier than ever to not worry about physical location. A multinational organization does not need separate support departments (such as IT, purchasing, etc.) for each location anymore.
- Link parallel activities instead of integrating their results. Departments that work in parallel should be sharing data and communicating with each other during their activities instead of waiting until each group is done and then comparing notes.
- Put the decision points where the work is performed, and build controls into the process. The people who do the work should have decision-making authority and the process itself should have built-in controls using information technology.
- Capture information once, at the source. Requiring information to be entered more than once causes delays and errors. With information technology, an organization can capture it once and then make it available whenever needed.

These principles may seem like common sense today, but in 1990 they took the business world by storm. Hammer gives example after example of how organizations improved their business processes by many orders of magnitude without adding any new employees, simply by changing how they did things (see sidebar).

Unfortunately, business process reengineering got a bad name in many organizations. This was because it was used as an excuse for cost cutting that really had nothing to do with BPR. For example, many companies simply used it as an excuse for laying off part of their workforce. Today, however, many of the principles of BPR have been integrated into businesses and are considered part of good business-process management.

12.8.7: Sample of Re-engineering

Reengineering the College Bookstore

The process of purchasing the correct textbooks in a timely manner for college classes has always been problematic. And now, with online bookstores such as Amazon competing directly with the college bookstore for students' purchases, the college bookstore is under pressure to justify its existence.

But college bookstores have one big advantage over their competitors: they have access to students' data. In other words, once a student has registered for classes, the bookstore knows exactly what books that student will need for the upcoming term. To leverage this advantage and take advantage of new technologies, the bookstore wants to implement a new process that will make purchasing books through the bookstore advantageous to students. Though they may not be able to compete on price, they can provide other advantages, such as reducing the time it takes to find the books and the ability to guarantee that the book is the correct one for the class. In order to do this, the bookstore will need to undertake a process redesign.

The goal of the process redesign is simple: capture a higher percentage of students as customers of the bookstore. After diagramming the existing process and meeting with student focus groups, the bookstore comes up with a new process. In the new process, the bookstore utilizes information technology to reduce the amount of work the students need to do in order to get their books. In this new process, the bookstore sends the students an e-mail with a list of all the books required for their upcoming classes. By clicking a link in this e-mail, the students can log into the bookstore, confirm their books, and purchase the books. The bookstore will then deliver the books to the students.

College bookstore process redesign (click to enlarge)

12.8.8: ISO 9001

ISO Certification

Many organizations now claim that they are using best practices when it comes to business processes. In order to set themselves apart and prove to their customers (and potential customers) that they are indeed doing this, these organizations are seeking out an ISO 9000 certification. ISO is an acronym for International Standards Organization (website here). This body defines quality standards that organizations can implement to show that they are, indeed, managing business processes in an effective way. The ISO 9000 certification is focused on quality management.

In order to receive ISO certification, an organization must be audited and found to meet specific criteria. In its most simple form, the auditors perform the following review:

- Tell me what you do *(describe the business process)*.
- Show me where it says that *(reference the process documentation)*.
- Prove that this is what happened *(exhibit evidence in documented records)*.

Over the years, this certification has evolved and many branches of the certification now exist. ISO certification is one way to separate an organization from others. You can find out more about the ISO 9000 standard here.

12.8.9: Summary

Summary

The advent of information technologies has had a huge impact on how organizations design, implement, and support business processes. From document management systems to ERP systems, information systems are tied into organizational processes. Using business process management, organizations can empower employees and leverage their processes for competitive advantage. Using business process reengineering, organizations can vastly improve their effectiveness and the quality of their products and services. Integrating information technology with business processes is one way that information systems can bring an organization lasting competitive advantage.

Study Questions

1. What does the term *business process* mean?
2. What are three examples of business process from a job you have had or an organization you have observed?
3. What is the value in documenting a business process?
4. What is an ERP system? How does an ERP system enforce best practices for an organization?
5. What is one of the criticisms of ERP systems?
6. What is business process reengineering? How is it different from incrementally improving a process?
7. Why did BPR get a bad name?
8. List the guidelines for redesigning a business process.
9. What is business process management? What role does it play in allowing a company to differentiate itself?
10. What does ISO certification signify?

Exercises

1. Think of a business process that you have had to perform in the past. How would you document this process? Would a diagram make more sense than a checklist? Document the process both as a checklist and as a diagram.
2. Review the return policies at your favorite retailer, then answer this question: What information systems do you think would need to be in place to support their return policy.
3. If you were implementing an ERP system, in which cases would you be more inclined to modify the ERP to match your business processes? What are the drawbacks of doing this?
4. Which ERP is the best? Do some original research and compare three leading ERP systems to each other. Write a two- to three-page paper that compares their features.

---

1. Hammer, Michael. "Reengineering work: don't automate, obliterate." *Harvard Business Review* 68.4 (1990): 104–112.↵

SECTION OVERVIEW

12.9: The People in Information Systems

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe each of the different roles that people play in the design, development, and use of information systems;
- understand the different career paths available to those who work with information systems;
- explain the importance of where the information-systems function is placed in an organization; and
- describe the different types of users of information systems.

Introduction

In the opening chapters of this text, we focused on the technology behind information systems: hardware, software, data, and networking. In the last chapter, we discussed business processes and the key role they can play in the success of a business. In this chapter, we will be discussing the last component of an information system: people.

People are involved in information systems in just about every way you can think of: people imagine information systems, people develop information systems, people support information systems, and, perhaps most importantly, people *use* information systems.

12.9.1: Creators

The Creators of Information Systems

The first group of people we are going to look at play a role in designing, developing, and building information systems. These people are generally very technical and have a background in programming and mathematics. Just about everyone who works in the creation of information systems has a minimum of a bachelor's degree in computer science or information systems, though that is not necessarily a requirement. We will be looking at the process of creating information systems in more detail in chapter 10.

12.9.1.1: System Analyst

Systems Analyst

The role of the systems analyst is to straddle the divide between identifying business needs and imagining a new or redesigned computer-based system to fulfill those needs. This individual will work with a person, team, or department with business requirements and identify the specific details of a system that needs to be built. Generally, this will require the analyst to have a good understanding of the business itself , the business processes involved, and the ability to document them well. The analyst will identify the different stakeholders in the system and work to involve the appropriate individuals in the process.

Once the requirements are determined, the analyst will begin the process of translating these requirements into an information-systems design. A good analyst will understand what different technological solutions will work and provide several different alternatives to the requester, based on the company's budgetary constraints, technology constraints, and culture. Once the solution is selected, the analyst will create a detailed document describing the new system. This new document will require that the analyst understand how to speak in the technical language of systems developers.

A systems analyst generally is not the one who does the actual development of the information system. The design document created by the systems analyst provides the detail needed to create the system and is handed off to a programmer (or team of programmers) to do the actual creation of the system. In some cases, however, a systems analyst may go ahead and create the system that he or she designed. This person is sometimes referred to as a programmer-analyst.

In other cases, the system may be assembled from off-the-shelf components by a person called a systems integrator. This is a specific type of systems analyst that understands how to get different software packages to work with each other.

To become a systems analyst, you should have a background both in the business and in systems design. Many analysts first worked as programmers and/or had experience in the business before becoming systems analysts.

12.9.1.2: System Programmer

Programmer

Programmers spend their time writing computer code in a programming language. In the case of systems development, programmers generally attempt to fulfill the design specifications given to them by a systems analyst. Many different styles of programming exist: a programmer may work alone for long stretches of time or may work in a team with other programmers. A programmer needs to be able to understand complex processes and also the intricacies of one or more programming languages. Generally, a programmer is very proficient in mathematics, as mathematical concepts underlie most programming code.

12.9.1.3: Computer Engineer

Computer Engineer

Computer engineers design the computing devices that we use every day. There are many types of computer engineers, who work on a variety of different types of devices and systems. Some of the more prominent engineering jobs are as follows:

- Hardware engineer. A hardware engineer designs hardware components, such as microprocessors. Many times, a hardware engineer is at the cutting edge of computing technology, creating something brand new. Other times, the hardware engineer's job is to engineer an existing component to work faster or use less power. Many times, a hardware engineer's job is to write code to create a program that will be implemented directly on a computer chip.
- Software engineer. Software engineers do not actually design devices; instead, they create new programming languages and operating systems, working at the lowest levels of the hardware to develop new kinds of software to run on the hardware.
- Systems engineer. A systems engineer takes the components designed by other engineers and makes them all work together. For example, to build a computer, the mother board, processor, memory, and hard disk all have to work together. A systems engineer has experience with many different types of hardware and software and knows how to integrate them to create new functionality.
- Network engineer. A network engineer's job is to understand the networking requirements of an organization and then design a communications system to meet those needs, using the networking hardware and software available.

There are many different types of computer engineers, and often the job descriptions overlap. While many may call themselves engineers based on a company job title, there is also a professional designation of "professional engineer," which has specific requirements behind it. In the US, each state has its own set of requirements for the use of this title, as do different countries around the world. Most often, it involves a professional licensing exam.

12.9.2: Operations and Administration

Information-Systems Operations and Administration

Another group of information-systems professionals are involved in the day-to-day operations and administration of IT. These people must keep the systems running and up-to-date so that the rest of the organization can make the most effective use of these resources.

12.9.2.1: Computer Operator

Computer Operator

A computer operator is the person who keeps the large computers running. This person's job is to oversee the mainframe computers and data centers in organizations. Some of their duties include keeping the operating systems up to date, ensuring available memory and disk storage, and overseeing the physical environment of the computer. Since mainframe computers increasingly have been replaced with servers, storage management systems, and other platforms, computer operators' jobs have grown broader and include working with these specialized systems.

12.9.2.2: Database Administrator

Database Administrator

A database administrator (DBA) is the person who manages the databases for an organization. This person creates and maintains databases that are used as part of applications or the data warehouse. The DBA also consults with systems analysts and programmers on projects that require access to or the creation of databases.

12.9.2.3: Support Desk

Help-Desk/Support Analyst

Most mid-size to large organizations have their own information-technology help desk. The help desk is the first line of support for computer users in the company. Computer users who are having problems or need information can contact the help desk for assistance. Many times, a help-desk worker is a junior-level employee who does not necessarily know how to answer all of the questions that come his or her way. In these cases, help-desk analysts work with senior-level support analysts or have a computer knowledgebase at their disposal to help them investigate the problem at hand. The help desk is a great place to break into working in IT because it exposes you to all of the different technologies within the company. A successful help-desk analyst should have good people and communications skills, as well as at least junior-level IT skills.

12.9.2.4: Trainer

Trainer

A computer trainer conducts classes to teach people specific computer skills. For example, if a new ERP system is being installed in an organization, one part of the implementation process is to teach all of the users how to use the new system. A trainer may work for a software company and be contracted to come in to conduct classes when needed; a trainer may work for a company that offers regular training sessions; or a trainer may be employed full time for an organization to handle all of their computer instruction needs. To be successful as a trainer, you need to be able to communicate technical concepts well and also have a lot of patience!

12.9.3: Managers

Managing Information Systems

The management of information-systems functions is critical to the success of information systems within the organization. Here are some of the jobs associated with the management of information systems.

12.9.3.1: CIO

CIO

The CIO, or chief information officer, is the head of the information-systems function. This person aligns the plans and operations of the information systems with the strategic goals of the organization. This includes tasks such as budgeting, strategic planning, and personnel decisions for the information-systems function. The CIO must also be the face of the IT department within the organization. This involves working with senior leaders in all parts of the organization to ensure good communication and planning.

Interestingly, the CIO position does not necessarily require a lot of technical expertise. While helpful, it is more important for this person to have good management skills and understand the business. Many organizations do not have someone with the title of CIO; instead, the head of the information-systems function is called vice president of information systems or director of information systems.

12.9.3.2: Functional Manager

Functional Manager

As an information-systems organization becomes larger, many of the different functions are grouped together and led by a manager. These functional managers report to the CIO and manage the employees specific to their function. For example, in a large organization, there is a group of systems analysts who report to a manager of the systems-analysis function. For more insight into how this might look, see the discussion later in the chapter of how information systems are organized.

12.9.3.3: ERP Manager

ERP Management

Organizations using an ERP require one or more individuals to manage these systems. These people make sure that the ERP system is completely up to date, work to implement any changes to the ERP that are needed, and consult with various user departments on needed reports or data extracts.

12.9.3.4: Project Managers

Project Managers

Information-systems projects are notorious for going over budget and being delivered late. In many cases, a failed IT project can spell doom for a company. A project manager is responsible for keeping projects on time and on budget. This person works with the stakeholders of the project to keep the team organized and communicates the status of the project to management. A project manager does not have authority over the project team; instead, the project manager coordinates schedules and resources in order to maximize the project outcomes. A project manager must be a good communicator and an extremely organized person. A project manager should also have good people skills. Many organizations require each of their project managers to become certified as a project management professional (PMP).

12.9.3.5: Info-Sec Officer

Information-Security Officer

An information-security officer is in charge of setting information-security policies for an organization, and then overseeing the implementation of those policies. This person may have one or more people reporting to them as part of the information-security team. As information has become a critical asset, this position has become highly valued. The information-security officer must ensure that the organization's information remains secure from both internal and external threats.

12.9.3.6: Emerging Roles

Emerging Roles

As technology evolves, many new roles are becoming more common as other roles fade. For example, as we enter the age of "big data," we are seeing the need for more data analysts and business-intelligence specialists. Many companies are now hiring social-media experts and mobile-technology specialists. The increased use of cloud computing and virtual-machine technologies also is breeding demand for expertise in those areas.

12.9.4: Careers

Career Paths in Information Systems

These job descriptions do not represent all possible jobs within an information-systems organization. Larger organizations will have more specialized roles; smaller organizations may combine some of these roles. Many of these roles may exist outside of a traditional information-systems organization, as we will discuss below.

Working with information systems can be a rewarding career choice. Whether you want to be involved in very technical jobs (programmer, database administrator), or you want to be involved in working with people (systems analyst, trainer), there are many different career paths available.

Many times, those in technical jobs who want career advancement find themselves in a dilemma: do they want to continue doing technical work, where sometimes their advancement options are limited, or do they want to become a manager of other employees and put themselves on a management career track? In many cases, those proficient in technical skills are not gifted with managerial skills. Some organizations, especially those that highly value their technically skilled employees, will create a technical track that exists in parallel to the management track so that they can retain employees who are contributing to the organization with their technical skills.

Sidebar: Are Certifications Worth Pursuing?

As technology is becoming more and more important to businesses, hiring employees with technical skills is becoming critical. But how can an organization ensure that the person they are hiring has the necessary skills? These days, many organizations are including technical certifications as a prerequisite for getting hired.

Certifications are designations given by a certifying body that someone has a specific level of knowledge in a specific technology. This certifying body is often the vendor of the product itself, though independent certifying organizations, such as CompTIA, also exist. Many of these organizations offer certification tracks, allowing a beginning certificate as a prerequisite to getting more advanced certificates. To get a certificate, you generally attend one or more training classes and then take one or more certification exams. Passing the exams with a certain score will qualify you for a certificate. In most cases, these classes and certificates are not free and, in fact, can run into the thousands of dollars. Some examples of the certifications in highest demand include Microsoft (software certifications), Cisco (networking), and SANS (security).

For many working in IT (or thinking about an IT career), determining whether to pursue one or more of these certifications is an important question. For many jobs, such as those involving networking or security, a certificate will be required by the employer as a way to determine which potential employees have a basic level of skill. For those who are already in an IT career, a more advanced certificate may lead to a promotion. There are other cases, however, when experience with a certain technology will negate the need for certification. For those wondering about the importance of certification, the best solution is to talk to potential employers and those already working in the field to determine the best choice.

12.9.5: Organization

Organizing the Information-Systems Function

In the early years of computing, the information-systems function (generally called data processing) was placed in the finance or accounting department of the organization. As computing became more important, a separate information-systems function was formed, but it still was generally placed under the CFO and considered to be an administrative function of the company. In the 1980s and 1990s, when companies began networking internally and then linking up to the Internet, the information-systems function was combined with the telecommunications functions and designated the information technology (IT) department. As the role of information technology continued to increase, its place in the organization also moved up the ladder. In many organizations today, the head of IT (the CIO) reports directly to the CEO.

12.9.5.1: Where are we?

Where in the Organization Should IS Be?

Before the advent of the personal computer, the information-systems function was centralized within organizations in order to maximize control over computing resources. When the PC began proliferating, many departments within organizations saw it as a chance to gain some computing resources for themselves. Some departments created an internal information-systems group, complete with systems analysts, programmers, and even database administrators. These departmental-IS groups were dedicated to the information needs of their own departments, providing quicker turnaround and higher levels of service than a centralized IT department. However, having several IS groups within an organization led to a lot of inefficiencies: there were now several people performing the same jobs in different departments. This decentralization also led to company data being stored in several places all over the company. In some organizations, a “matrix” reporting structure has developed, in which IT personnel are placed within a department and report to both the department management and the functional management within IS. The advantages of dedicated IS personnel for each department are weighed against the need for more control over the strategic information resources of the company.

For many companies, these questions are resolved by the implementation of the ERP system (see discussion of ERP in chapter 8). Because an ERP system consolidates most corporate data back into a single database, the implementation of an ERP system requires organizations to find “islands” of data so that they can integrate them back into the corporate system. The ERP allows organizations to regain control of their information and influences organizational decisions throughout the company.

12.9.5.2: New Thoughts

New Models of Organizations

The integration of information technology has influenced the structure of organizations. The increased ability to communicate and share information has led to a “flattening” of the organizational structure due to the removal of one or more layers of management.

Another organizational change enabled by information systems is the network-based organizational structure. In a networked-based organizational structure, groups of employees can work somewhat independently to accomplish a project. In a networked organization, people with the right skills are brought together for a project and then released to work on other projects when that project is over. These groups are somewhat informal and allow for all members of the group to maximize their effectiveness.

12.9.5.3: Outsourcing

Outsourcing

Many times, an organization needs a specific skill for a limited period of time. Instead of training an existing employee or hiring someone new, it may make more sense to outsource the job. Outsourcing can be used in many different situations within the information-systems function, such as the design and creation of a new website or the upgrade of an ERP system. Some organizations see outsourcing as a cost-cutting move, contracting out a whole group or department.

12.9.6: Users

Information-Systems Users – Types of Users

Besides the people who work to create, administer, and manage information systems, there is one more extremely important group of people: the users of information systems. This group represents a very large percentage of the people involved. If the user is not able to successfully learn and use an information system, the system is doomed to failure.

Technology adoption user types (click to enlarge). (Public Domain)

One tool that can be used to understand how users will adopt a new technology comes from a 1962 study by Everett Rogers. In his book, *Diffusion of Innovation,*[1] Rogers studied how farmers adopted new technologies, and he noticed that the adoption rate started slowly and then dramatically increased once adoption hit a certain point. He identified five specific types of technology adopters:

- Innovators. Innovators are the first individuals to adopt a new technology. Innovators are willing to take risks, are the youngest in age, have the highest social class, have great financial liquidity, are very social, and have the closest contact with scientific sources and interaction with other innovators. Risk tolerance has them adopting technologies that may ultimately fail. Financial resources help absorb these failures (Rogers 1962 5th ed, p. 282).
- Early adopters. The early adopters are those who adopt innovation after a technology has been introduced and proven. These individuals have the highest degree of opinion leadership among the other adopter categories, which means that they can influence the opinions of the largest majority. They are typically younger in age, have higher social status, more financial liquidity, more advanced education, and are more socially aware than later adopters. These people are more discrete in adoption choices than innovators, and realize judicious choice of adoption will help them maintain a central communication position (Rogers 1962 5th ed, p. 283).
- Early majority. Individuals in this category adopt an innovation after a varying degree of time. This time of adoption is significantly longer than the innovators and early adopters. This group tends to be slower in the adoption process, has above average social status, has contact with early adopters, and seldom holds positions of opinion leadership in a system (Rogers 1962 5th ed, p. 283).
- Late majority. The late majority will adopt an innovation after the average member of the society. These individuals approach an innovation with a high degree of skepticism, have below average social status, very little financial liquidity, are in contact with others in the late majority and the early majority, and show very little opinion leadership.
- Laggards. Individuals in this category are the last to adopt an innovation. Unlike those in the previous categories, individuals in this category show no opinion leadership. These individuals typically have an aversion to change-agents and tend to be advanced in age. Laggards typically tend to be focused on “traditions,” are likely to have the lowest social status and the lowest financial liquidity, be oldest of all other adopters, and be in contact with only family and close friends.

These five types of users can be translated into information-technology adopters as well, and provide additional insight into how to implement new information systems within an organization. For example, when rolling out a new system, IT may want to identify the innovators and early adopters within the organization and work with them first, then leverage their adoption to drive the rest of the implementation.

12.9.7: Summary

Summary

In this chapter, we have reviewed the many different categories of individuals who make up the people component of information systems. The world of information technology is changing so fast that new roles are being created all the time, and roles that existed for decades are being phased out. That said, this chapter should have given you a good idea of the importance of the people component of information systems.

Study Questions

1. Describe the role of a systems analyst.
2. What are some of the different roles for a computer engineer?
3. What are the duties of a computer operator?
4. What does the CIO do?
5. Describe the job of a project manager.
6. Explain the point of having two different career paths in information systems.
7. What are the advantages and disadvantages of centralizing the IT function?
8. What impact has information technology had on the way companies are organized?
9. What are the five types of information-systems users?
10. Why would an organization outsource?

Exercises

1. Which IT job would you like to have? Do some original research and write a two-page paper describing the duties of the job you are interested in.
2. Spend a few minutes on Dice or Monster to find IT jobs in your area. What IT jobs are currently available? Write up a two-page paper describing three jobs, their starting salary (if listed), and the skills and education needed for the job.
3. How is the IT function organized in your school or place of employment? Create an organization chart showing how the IT organization fits into your overall organization. Comment on how centralized or decentralized the IT function is.
4. What type of IT user are you? Take a look at the five types of technology adopters and then write a one-page summary of where you think you fit in this model.

1. Rogers, E. M. (1962). Diffusion of innovations. New York: Free Press↵

SECTION OVERVIEW

12.10: Information Systems Development

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- explain the overall process of developing a new software application;
- explain the differences between software development methodologies;
- understand the different types of programming languages used to develop software;
- understand some of the issues surrounding the development of websites and mobile applications; and
- identify the four primary implementation policies.

Introduction

When someone has an idea for a new function to be performed by a computer, how does that idea become reality? If a company wants to implement a new business process and needs new hardware or software to support it, how do they go about making it happen? In this chapter, we will discuss the different methods of taking those ideas and bringing them to reality, a process known as information systems development.

Programming

As we learned in chapter 2, software is created via programming. Programming is the process of creating a set of logical instructions for a digital device to follow using a programming language. The process of programming is sometimes called "coding" because the syntax of a programming language is not in a form that everyone can understand – it is in "code."

The process of developing good software is usually not as simple as sitting down and writing some code. True, sometimes a programmer can quickly write a short program to solve a need. But most of the time, the creation of software is a resource-intensive process that involves several different groups of people in an organization. In the following sections, we are going to review several different methodologies for software development.

12.10.1: SDLC

Systems-Development Life Cycle

The first development methodology we are going to review is the systems-development life cycle (SDLC). This methodology was first developed in the 1960s to manage the large software projects associated with corporate systems running on mainframes. It is a very structured and risk-averse methodology designed to manage large projects that included multiple programmers and systems that would have a large impact on the organization.

SDLC waterfall (click to enlarge)

Various definitions of the SDLC methodology exist, but most contain the following phases.

1. Preliminary Analysis. In this phase, a review is done of the request. Is creating a solution possible? What alternatives exist? What is currently being done about it? Is this project a good fit for our organization? A key part of this step is a feasibility analysis, which includes an analysis of the technical feasibility (is it possible to create this?), the economic feasibility (can we afford to do this?), and the legal feasibility (are we allowed to do this?). This step is important in determining if the project should even get started.
2. System Analysis. In this phase, one or more system analysts work with different stakeholder groups to determine the specific requirements for the new system. No programming is done in this step. Instead, procedures are documented, key players are interviewed, and data requirements are developed in order to get an overall picture of exactly what the system is supposed to do. The result of this phase is a system-requirements document.
3. System Design. In this phase, a designer takes the system-requirements document created in the previous phase and develops the specific technical details required for the system. It is in this phase that the business requirements are translated into specific technical requirements. The design for the user interface, database, data inputs and outputs, and reporting are developed here. The result of this phase is a system-design document. This document will have everything a programmer will need to actually create the system.
4. Programming. The code finally gets written in the programming phase. Using the system-design document as a guide, a programmer (or team of programmers) develop the program. The result of this phase is an initial working program that meets the requirements laid out in the system-analysis phase and the design developed in the system-design phase.
5. Testing. In the testing phase, the software program developed in the previous phase is put through a series of structured tests. The first is a unit test, which tests individual parts of the code for errors or bugs. Next is a system test, where the different components of the system are tested to ensure that they work together properly. Finally, the user-acceptance test allows those that will be using the software to test the system to ensure that it meets their standards. Any bugs, errors, or problems found during testing are addressed and then tested again.
6. Implementation. Once the new system is developed and tested, it has to be implemented in the organization. This phase includes training the users, providing documentation, and conversion from any previous system to the new system. Implementation can take many forms, depending on the type of system, the number and type of users, and how urgent it is that the system become operational. These different forms of implementation are covered later in the chapter.
7. Maintenance. This final phase takes place once the implementation phase is complete. In this phase, the system has a structured support process in place: reported bugs are fixed and requests for new features are evaluated and implemented; system updates and backups are performed on a regular basis.

The SDLC methodology is sometimes referred to as the waterfall methodology to represent how each step is a separate part of the process; only when one step is completed can another step begin. After each step, an organization must decide whether to move to the next step or not. This methodology has been criticized for being quite rigid. For example, changes to the requirements are not allowed once the process has begun. No software is available until after the programming phase.

Again, SDLC was developed for large, structured projects. Projects using SDLC can sometimes take months or years to complete. Because of its inflexibility and the availability of new programming techniques and tools, many other software-development methodologies have been developed. Many of these retain some of the underlying concepts of SDLC but are not as rigid.

12.10.2: Rapid App Dev

Rapid Application Development

The RAD methodology (Public Domain)

Rapid application development (RAD) is a software-development (or systems-development) methodology that focuses on quickly building a working model of the software, getting feedback from users, and then using that feedback to update the working model. After several iterations of development, a final version is developed and implemented.

The RAD methodology consists of four phases:

1. Requirements Planning. This phase is similar to the preliminary-analysis, system-analysis, and design phases of the SDLC. In this phase, the overall requirements for the system are defined, a team is identified, and feasibility is determined.
2. User Design. In this phase, representatives of the users work with the system analysts, designers, and programmers to interactively create the design of the system. One technique for working with all of these various stakeholders is the so-called JAD session. JAD is an acronym for joint application development. A JAD session gets all of the stakeholders together to have a structured discussion about the design of the system. Application developers also sit in on this meeting and observe, trying to understand the essence of the requirements.
3. Construction. In the construction phase, the application developers, working with the users, build the next version of the system.This is an interactive process, and changes can be made as developers are working on the program. This step is executed in parallel with the User Design step in an iterative fashion, until an acceptable version of the product is developed.
4. Cutover. In this step, which is similar to the implementation step of the SDLC, the system goes live. All steps required to move from the previous state to the use of the new system are completed here.

As you can see, the RAD methodology is much more compressed than SDLC. Many of the SDLC steps are combined and the focus is on user participation and iteration. This methodology is much better suited for smaller projects than SDLC and has the added advantage of giving users the ability to provide feedback throughout the process. SDLC requires more documentation and attention to detail and is well suited to large, resource-intensive projects. RAD makes more sense for smaller projects that are less resource-intensive and need to be developed quickly.

12.10.3: Agile and Lean Methods

Agile Methodologies

Agile methodologies are a group of methodologies that utilize incremental changes with a focus on quality and attention to detail. Each increment is released in a specified period of time (called a time box), creating a regular release schedule with very specific objectives. While considered a separate methodology from RAD, they share some of the same principles: iterative development, user interaction, ability to change. The agile methodologies are based on the "Agile Manifesto," first released in 2001.

The characteristics of agile methods include:

- small cross-functional teams that include development-team members and users;
- daily status meetings to discuss the current state of the project;
- short time-frame increments (from days to one or two weeks) for each change to be completed; and
- at the end of each iteration, a working project is completed to demonstrate to the stakeholders.

The goal of the agile methodologies is to provide the flexibility of an iterative approach while ensuring a quality product.

Lean Methodology

The lean methodology (click to enlarge)

One last methodology we will discuss is a relatively new concept taken from the business bestseller *The Lean Startup*, by Eric Reis. In this methodology, the focus is on taking an initial idea and developing a minimum viable product (MVP). The MVP is a working software application with just enough functionality to demonstrate the idea behind the project. Once the MVP is developed, it is given to potential users for review. Feedback on the MVP is generated in two forms: (1) direct observation and discussion with the users, and (2) usage statistics gathered from the software itself. Using these two forms of feedback, the team determines whether they should continue in the same direction or rethink the core idea behind the project, change the functions, and create a new MVP. This change in strategy is called a pivot. Several iterations of the MVP are developed, with new functions added each time based on the feedback, until a final product is completed.

The biggest difference between the lean methodology and the other methodologies is that the full set of requirements for the system are not known when the project is launched. As each iteration of the project is released, the statistics and feedback gathered are used to determine the requirements. The lean methodology works best in an entrepreneurial environment where a company is interested in determining if their idea for a software application is worth developing.

Sidebar: The Quality Triangle

The quality triangle

When developing software, or any sort of product or service, there exists a tension between the developers and the different stakeholder groups, such as management, users, and investors. This tension relates to how quickly the software can be developed (time), how much money will be spent (cost), and how well it will be built (quality). The quality triangle is a simple concept. It states that for any product or service being developed, you can only address two of the following: time, cost, and quality.

So what does it mean that you can only address two of the three? It means that you cannot complete a low-cost, high-quality project in a small amount of time. However, if you are willing or able to spend a lot of money, then a project can be completed quickly with high-quality results (through hiring more good programmers). If a project's completion date is not a priority, then it can be completed at a lower cost with higher-quality results. Of course, these are just generalizations, and different projects may not fit this model perfectly. But overall, this model helps us understand the tradeoffs that we must make when we are developing new products and services.

SECTION OVERVIEW

12.10.4: Programming Languages

Programming Languages

As I noted earlier, software developers create software using one of several programming languages. A programming language is an artificial language that provides a way for a programmer to create structured code to communicate logic in a format that can be executed by the computer hardware. Over the past few decades, many different types of programming languages have evolved to meet many different needs. One way to characterize programming languages is by their “generation.”

Featured articles

12.10.4.1: Language Generations

Generations of Programming Languages

Early languages were specific to the type of hardware that had to be programmed; each type of computer hardware had a different low-level programming language (in fact, even today there are differences at the lower level, though they are now obscured by higher-level programming languages). In these early languages, very specific instructions had to be entered line by line – a tedious process.

First-generation languages are called machine code. In machine code, programming is done by directly setting actual ones and zeroes (the bits) in the program using binary code. Here is an example program that adds 1234 and 4321 using machine language:

```
10111001 00000000
11010010 10100001
00000100 00000000
10001001 00000000
00001110 10001011
00000000 00011110
00000000 00011110
00000000 00000010
10111001 00000000
11100001 00000011
00010000 11000011
10001001 10100011
00001110 00000100
00000010 00000000
```

Assembly language is the second-generation language. Assembly language gives english-like phrases to the machine-code instructions, making it easier to program. An assembly-language program must be run through an assembler, which converts it into machine code. Here is an example program that adds 1234 and 4321 using assembly language:

```
MOV CX,1234
MOV DS:[0],CX
MOV CX,4321
MOV AX,DS:[0]
MOV BX,DS:[2]
ADD AX,BX
MOV DS:[4],AX
```

Third-generation languages are not specific to the type of hardware on which they run and are much more like spoken languages. Most third-generation languages must be compiled, a process that converts them into machine code. Well-known third-generation languages include BASIC, C, Pascal, and Java. Here is an example using BASIC:

```
A=1234
B=4321
C=A+B
END
```

Fourth-generation languages are a class of programming tools that enable fast application development using intuitive interfaces and environments. Many times, a fourth-generation language has a very specific purpose, such as database interaction or report-writing. These tools can be used by those with very little formal training in programming and allow for the quick development of applications and/or functionality. Examples of fourth-generation languages include: Clipper, FOCUS, FoxPro, SQL, and SPSS.

Why would anyone want to program in a lower-level language when they require so much more work? The answer is similar to why some prefer to drive stick-shift automobiles instead of automatic transmission: control and efficiency. Lower-level languages, such as assembly language, are much more efficient and execute much more quickly. You have finer control over the hardware as well. Sometimes, a combination of higher- and lower-level languages are mixed together to get the best of both worlds: the programmer will create the overall structure and interface using a higher-level language but will use lower-level languages for the parts of the program that are used many times or require more precision.

The programming language spectrum (click to enlarge)

Compiled vs. Interpreted

Besides classifying a program language based on its generation, it can also be classified by whether it is compiled or interpreted. As we have learned, a computer language is written in a human-readable form. In a compiled language, the program code is translated into a machine-readable form called an executable that can be run on the hardware. Some well-known compiled languages include C, C++, and COBOL.

An interpreted language is one that requires a runtime program to be installed in order to execute. This runtime program then interprets the program code line by line and runs it. Interpreted languages are generally easier to work with but also are slower and require more system resources. Examples of popular interpreted languages include BASIC, PHP, PERL, and Python. The web languages of HTML and Javascript would also be considered interpreted because they require a browser in order to run.

The Java programming language is an interesting exception to this classification, as it is actually a hybrid of the two. A program written in Java is partially compiled to create a program that can be understood by the Java Virtual Machine (JVM). Each type of operating system has its own JVM which must be installed, which is what allows Java programs to run on many different types of operating systems.

12.10.4.2: Compiled vs Interpreted

Compiled vs. Interpreted

Besides classifying a program language based on its generation, it can also be classified by whether it is compiled or interpreted. As we have learned, a computer language is written in a human-readable form. In a compiled language, the program code is translated into a machine-readable form called an executable that can be run on the hardware. Some well-known compiled languages include C, C++, and COBOL.

An interpreted language is one that requires a runtime program to be installed in order to execute. This runtime program then interprets the program code line by line and runs it. Interpreted languages are generally easier to work with but also are slower and require more system resources. Examples of popular interpreted languages include BASIC, PHP, PERL, and Python. The web languages of HTML and Javascript would also be considered interpreted because they require a browser in order to run.

The Java programming language is an interesting exception to this classification, as it is actually a hybrid of the two. A program written in Java is partially compiled to create a program that can be understood by the Java Virtual Machine (JVM). Each type of operating system has its own JVM which must be installed, which is what allows Java programs to run on many different types of operating systems.

12.10.4.3: Procedural / Object Oriented

Procedural vs. Object-Oriented

A procedural programming language is designed to allow a programmer to define a specific starting point for the program and then execute sequentially. All early programming languages worked this way. As user interfaces became more interactive and graphical, it made sense for programming languages to evolve to allow the user to define the flow of the program. The object-oriented programming language is set up so that the programmer defines "objects" that can take certain actions based on input from the user. In other words, a procedural program focuses on the sequence of activities to be performed; an object-oriented program focuses on the different items being manipulated.

For example, in a human-resources system, an "EMPLOYEE" object would be needed. If the program needed to retrieve or set data regarding an employee, it would first create an employee object in the program and then set or retrieve the values needed. Every object has properties, which are descriptive fields associated with the object. In the example below, an employee object has the properties "Name", "Employee number", "Birthdate" and "Date of hire". An object also has "methods", which can take actions related to the object. In the example, there are two methods. The first is "ComputePay()", which will return the current amount owed the employee. The second is "ListEmployees()", which will retrieve a list of employees who report to this employee.

Object: EMPLOYEE
Name Employee number Birthdate Date of hire
ComputePay() ListEmployees()

Figure: An example of an object

Sidebar: What is COBOL?

If you have been around business programming very long, you may have heard about the COBOL programming language. COBOL is a procedural, compiled language that at one time was the primary programming language for business applications. Invented in 1959 for use on large mainframe computers, COBOL is an abbreviation of common business-oriented language. With the advent of more efficient programming languages, COBOL is now rarely seen outside of old, legacy applications.

12.10.5: IDE / CASE

Programming Tools

To write a program, a programmer needs little more than a text editor and a good idea. However, to be productive, he or she must be able to check the syntax of the code, and, in some cases, compile the code. To be more efficient at programming, additional tools, such as an integrated development environment (IDE) or computer-aided software-engineering (CASE) tools, can be used.

Integrated Development Environment

For most programming languages, an IDE can be used. An IDE provides a variety of tools for the programmer, and usually includes:

- an editor for writing the program that will color-code or highlight keywords from the programming language;
- a help system that gives detailed documentation regarding the programming language;
- a compiler/interpreter, which will allow the programmer to run the program;
- a debugging tool, which will provide the programmer details about the execution of the program in order to resolve problems in the code; and
- a check-in/check-out mechanism, which allows for a team of programmers to work together on a project and not write over each other's code changes.

Probably the most popular IDE software package right now is Microsoft's Visual Studio. Visual Studio is the IDE for all of Microsoft's programming languages, including Visual Basic, Visual C++, and Visual C#.

CASE Tools

While an IDE provides several tools to assist the programmer in writing the program, the code still must be written. Computer-aided software-engineering (CASE) tools allow a designer to develop software with little or no programming. Instead, the CASE tool writes the code for the designer. CASE tools come in many varieties, but their goal is to generate quality code based on input created by the designer.

12.10.6: Home Grown or Purchased

Build vs. Buy

When an organization decides that a new software program needs to be developed, they must determine if it makes more sense to build it themselves or to purchase it from an outside company. This is the “build vs. buy” decision.

There are many advantages to purchasing software from an outside company. First, it is generally less expensive to purchase a software package than to build it. Second, when a software package is purchased, it is available much more quickly than if the package is built in-house. Software applications can take months or years to build; a purchased package can be up and running within a month. A purchased package has already been tested and many of the bugs have already been worked out. It is the role of a systems integrator to make various purchased systems and the existing systems at the organization work together.

There are also disadvantages to purchasing software. First, the same software you are using can be used by your competitors. If a company is trying to differentiate itself based on a business process that is in that purchased software, it will have a hard time doing so if its competitors use the same software. Another disadvantage to purchasing software is the process of customization. If you purchase a software package from a vendor and then customize it, you will have to manage those customizations every time the vendor provides an upgrade. This can become an administrative headache, to say the least!

Even if an organization determines to buy software, it still makes sense to go through many of the same analyses that they would do if they were going to build it themselves. This is an important decision that could have a long-term strategic impact on the organization.

12.10.7: Web Service

Web Services

As we saw in chapter 3, the move to cloud computing has allowed software to be looked at as a service. One option companies have these days is to license functions provided by other companies instead of writing the code themselves. These are called web services, and they can greatly simplify the addition of functionality to a website.

For example, suppose a company wishes to provide a map showing the location of someone who has called their support line. By utilizing Google Maps API web services, they can build a Google Map right into their application. Or a shoe company could make it easier for its retailers to sell shoes online by providing a shoe-size web service that the retailers could embed right into their website.

Web services can blur the lines between "build vs. buy." Companies can choose to build a software application themselves but then purchase functionality from vendors to supplement their system.

12.10.8: End User Dev

End-User Computing

In many organizations, application development is not limited to the programmers and analysts in the information-technology department. Especially in larger organizations, other departments develop their own department-specific applications. The people who build these are not necessarily trained in programming or application development, but they tend to be adept with computers. A person, for example, who is skilled in a particular software package, such as a spreadsheet or database package, may be called upon to build smaller applications for use by his or her own department. This phenomenon is referred to as end-user development, or end-user computing.

End-user computing can have many advantages for an organization. First, it brings the development of applications closer to those who will use them. Because IT departments are sometimes quite backlogged, it also provides a means to have software created more quickly. Many organizations encourage end-user computing to reduce the strain on the IT department.

End-user computing does have its disadvantages as well. If departments within an organization are developing their own applications, the organization may end up with several applications that perform similar functions, which is inefficient, since it is a duplication of effort. Sometimes, these different versions of the same application end up providing different results, bringing confusion when departments interact. These applications are often developed by someone with little or no formal training in programming. In these cases, the software developed can have problems that then have to be resolved by the IT department.

End-user computing can be beneficial to an organization, but it should be managed. The IT department should set guidelines and provide tools for the departments who want to create their own solutions. Communication between departments will go a long way towards successful use of end-user computing.

Sidebar: Building a Mobile App

In many ways, building an application for a mobile device is exactly the same as building an application for a traditional computer. Understanding the requirements for the application, designing the interface, working with users – all of these steps still need to be carried out.

So what's different about building an application for a mobile device? In some ways, mobile applications are more limited. An application running on a mobile device must be designed to be functional on a smaller screen. Mobile applications should be designed to use fingers as the primary pointing device. Mobile devices generally have less available memory, storage space, and processing power.

Mobile applications also have many advantages over applications built for traditional computers. Mobile applications have access to the functionality of the mobile device, which usually includes features such as geolocation data, messaging, the camera, and even a gyroscope.

One of the most important questions regarding development for mobile devices is this: Do we want to develop an app at all? A mobile app is an expensive proposition, and it will only run on one type of mobile device at a time. For example, if you create an iPhone app, users with Android phones are out of luck. Each app takes several thousand dollars to create, so this may not be the best use of your funds.

Many organizations are moving away from developing a specific app for a mobile device and are instead making their websites more functional on mobile devices. Using a web-design framework called responsive design, a website can be made highly functional no matter what type of device is browsing it. With a responsive website, images resize themselves based on the size of the device's screen, and text flows and sizes itself properly for optimal viewing. You can find out more about responsive design here.

12.10.9: Project Implementation

Implementation Methodologies

Once a new system is developed (or purchased), the organization must determine the best method for implementing it. Convincing a group of people to learn and use a new system can be a very difficult process. Using new software, and the business processes it gives rise to, can have far-reaching effects within the organization.

There are several different methodologies an organization can adopt to implement a new system. Four of the most popular are listed below.

- Direct cutover. In the direct-cutover implementation methodology, the organization selects a particular date that the old system is not going to be used anymore. On that date, the users begin using the new system and the old system is unavailable. The advantages to using this methodology are that it is very fast and the least expensive. However, this method is the riskiest as well. If the new system has an operational problem or if the users are not properly prepared, it could prove disastrous for the organization.
- Pilot implementation. In this methodology, a subset of the organization (called a pilot group) starts using the new system before the rest of the organization. This has a smaller impact on the company and allows the support team to focus on a smaller group of individuals.
- Parallel operation. With parallel operation, the old and new systems are used simultaneously for a limited period of time. This method is the least risky because the old system is still being used while the new system is essentially being tested. However, this is by far the most expensive methodology since work is duplicated and support is needed for both systems in full.
- Phased implementation. In phased implementation, different functions of the new application are used as functions from the old system are turned off. This approach allows an organization to slowly move from one system to another.

Which of these implementation methodologies to use depends on the complexity and importance of the old and new systems.

Change Management

As new systems are brought online and old systems are phased out, it becomes important to manage the way change is implemented in the organization. Change should never be introduced in a vacuum. The organization should be sure to communicate proposed changes before they happen and plan to minimize the impact of the change that will occur after implementation. Change management is a critical component of IT oversight.

Maintenance

Once a new system has been introduced, it enters the maintenance phase. In this phase, the system is in production and is being used by the organization. While the system is no longer actively being developed, changes need to be made when bugs are found or new features are requested. During the maintenance phase, IT management must ensure that the system continues to stay aligned with business priorities and continues to run well.

12.10.10: Summary

Summary

Software development is about so much more than programming. Developing new software applications requires several steps, from the formal SDLC process to more informal processes such as agile programming or lean methodologies. Programming languages have evolved from very low-level machine-specific languages to higher-level languages that allow a programmer to write software for a wide variety of machines. Most programmers work with software development tools that provide them with integrated components to make the software development process more efficient. For some organizations, building their own software applications does not make the most sense; instead, they choose to purchase software built by a third party to save development costs and speed implementation. In end-user computing, software development happens outside the information technology department. When implementing new software applications, there are several different types of implementation methodologies that must be considered.

Study Questions

1. What are the steps in the SDLC methodology?
2. What is RAD software development?
3. What makes the lean methodology unique?
4. What are three differences between second-generation and third-generation languages?
5. Why would an organization consider building its own software application if it is cheaper to buy one?
6. What is responsive design?
7. What is the relationship between HTML and CSS in website design?
8. What is the difference between the pilot implementation methodology and the parallel implementation methodology?
9. What is change management?
10. What are the four different implementation methodologies?

Exercises

1. Which software-development methodology would be best if an organization needed to develop a software tool for a small group of users in the marketing department? Why? Which implementation methodology should they use? Why?
2. Doing your own research, find three programming languages and categorize them in these areas: generation, compiled vs. interpreted, procedural vs. object-oriented.
3. Some argue that HTML is not a programming language. Doing your own research, find three arguments for why it is not a programming language and three arguments for why it is.
4. Read more about responsive design using the link given in the text. Provide the links to three websites that use responsive design and explain how they demonstrate responsive-design behavior.

SECTION OVERVIEW

12.11: Globalization and the Digital Divide

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- explain the concept of globalization;
- describe the role of information technology in globalization;
- identify the issues experienced by firms as they face a global economy; and
- define the digital divide and explain Nielsen's three stages of the digital divide.

Introduction

The Internet has wired the world. Today it is just as simple to communicate with someone on the other side of the world as it is to talk to someone next door. In this chapter, we will look at the implications of globalization and the impact it is having on the world.

12.11.1: Globalization?

What Is Globalization?

Globalization is the term used to refer to the integration of goods, services, and culture among the nations of the world. Globalization is not necessarily a new phenomenon; in many ways, we have been experiencing globalization since the days of European colonization. Further advances in telecommunication and transportation technologies accelerated globalization. The advent of the the worldwide Internet has made all nations next-door neighbors.

The Internet is truly a worldwide phenomenon. As of 2012, the Internet was being used in over 150 countries by a staggering 2.4 billion people worldwide, and growing.[1] From its initial beginnings in the United States in the 1970s to the development of the World Wide Web in the 1990s to the social networks and e-commerce of today, the Internet has continued to increase the integration between countries, making globalization a fact of life for citizens all over the world.

WORLD INTERNET USAGE AND POPULATION STATISTICS JUNE, 2019 - Updated						
World Regions	Population (2019 Est.)	Population % of World	Internet Users 30 June 2019	Penetration Rate (% Pop.)	Growth 2000-2019	Internet World %
Africa	1,320,038,716	17.1 %	525,148,631	39.8 %	11,533 %	11.9 %
Asia	4,241,972,790	55.0 %	2,200,658,148	51.9 %	1,825 %	49.8 %
Europe	829,173,007	10.7 %	719,413,014	86.8 %	585 %	16.3 %
Latin America / Caribbean	658,345,826	8.5 %	447,495,130	68.0 %	2,377 %	10.1 %
Middle East	258,356,867	3.3 %	173,576,793	67.2 %	5,184 %	3.9 %
North America	366,496,802	4.7 %	327,568,628	89.4 %	203 %	7.4 %
Oceania / Australia	41,839,201	0.5 %	28,634,278	68.4 %	276 %	0.6 %
WORLD TOTAL	7,716,223,209	100.0 %	4,422,494,622	57.3 %	1,125 %	100.0 %

Source: Internet World Stats

12.11.1.1: The World...

The Network Society

In 1996, social-sciences researcher Manuel Castells published *The Rise of the Network Society*, in which he identified new ways in which economic activity was being organized around the networks that the new telecommunication technologies have provided. This new, global economic activity was different from the past, because "it is an economy with the capacity to work as a unit in real time on a planetary scale."[2] We are now into this network society, where we are all connected on a global scale.

The World Is Flat

In 2005, Thomas Friedman's seminal book, *The World Is Flat*, was published. In this book, Friedman unpacks the impacts that the personal computer, the Internet, and communication software have had on business, specifically the impact they have had on globalization. He begins the book by defining the three eras of globalization[3]:

- "Globalization 1.0″ occurred from 1492 until about 1800. In this era, globalization was centered around countries. It was about how much horsepower, wind power, and steam power a country had and how creatively it was deployed. The world shrank from size "large" to size "medium."
- "Globalization 2.0″ occurred from about 1800 until 2000, interrupted only by the two World Wars. In this era, the dynamic force driving change was multinational companies. The world shrank from size "medium" to size "small."
- "Globalization 3.0″ is our current era, beginning in the year 2000. The convergence of the personal computer, fiber-optic Internet connections, and software has created a "flat-world platform" that allows small groups and even individuals to go global. The world has shrunk from size "small" to size "tiny."

According to Friedman, this third era of globalization was brought about, in many respects, by information technology. Some of the specific technologies he lists include:

- *The graphical user interface for the personal computer popularized in the late 1980s.* Before the graphical user interface, using a computer was relatively difficult. By making the personal computer something that anyone could use, it became commonplace very quickly. Friedman points out that this digital storage of content made people much more productive and, as the Internet evolved, made it simpler to communicate content worldwide.
- *The build-out of the Internet infrastructure during the dot-com boom during the late-1990s.* During the late 1990s, telecommunications companies laid thousands of miles of fiber-optic cable all over the world, turning network communications into a commodity. At the same time, the Internet protocols, such as SMTP (e-mail), HTML (web pages), and TCP/IP (network communications) became standards that were available for free and used by everyone.
- *The introduction of software to automate and integrate business processes.* As the Internet continued to grow and become the dominant form of communication, it became essential to build on the standards developed earlier so that the websites and applications running on the Internet would work well together. Friedman calls this "workflow software," by which he means software that allows people to work together more easily, and allows different software packages and databases to integrate with each other more easily. Examples include payment-processing systems and shipping calculators.

These three technologies came together in the late 1990s to create a "platform for global collaboration." Once these technologies were in place, they continued to evolve. Friedman also points out a couple more technologies that have contributed to the flat-world platform – the open-source movement (see chapter 10) and the advent of mobile technologies.

The World Is Flat was published in 2005. Since then, we have seen even more growth in information technologies that have contributed to global collaborations. We will discuss current and future trends in chapter 13.

12.11.2: Going Global

The Global Firm

The new era of globalization allows any business to become international. By accessing this new platform of technologies, Castells's vision of working as a unit in real time on a planetary scale can be a reality. Some of the advantages of this include the following:

- The ability to locate expertise and labor around the world. Instead of drawing employees from their local area, organizations can now hire people from the global labor pool. This also allows organizations to pay a lower labor cost for the same work based on the prevailing wage in different countries.
- The ability to operate 24 hours a day. With employees in different time zones all around the world, an organization can literally operate around the clock, handing off work on projects from one part of the world to another. Businesses can also keep their digital storefront (their website) open all the time.
- A larger market for their products. Once a product is being sold online, it is available for purchase from a worldwide consumer base. Even if a company's products do not appeal beyond its own country's borders, being online has also made the product more visible to consumers within that country.

In order to fully take advantage of these new capabilities, companies need to understand that there are also challenges in dealing with employees and customers from different cultures. Some of these challenges include:

- Infrastructure differences. Each country has its own infrastructure, many of which are not of the same quality as the US infrastructure (average 4.60 MBps). For every South Korea (16 MBps average speed) there is an Egypt (0.83 MBps) or an India (0.82 MBps). A business cannot depend on every country it deals with having the same Internet speeds. See the sidebar called "How Does My Internet Speed Compare?"
- Labor laws and regulations. Different countries (even different states in the United States) have different laws and regulations. A company that wants to hire employees from other countries must understand the different regulations and concerns.
- Legal restrictions. Many countries have restrictions on what can be sold or how a product can be advertised. It is important for a business to understand what is allowed. For example, in Germany, it is illegal to sell anything Nazi related; in China, it is illegal to put anything sexually suggestive online.
- Language, customs, and preferences. Every country has its own (or several) unique culture(s), which a business must consider when trying to market a product there. Additionally, different countries have different preferences. For example, in some parts of the world, people prefer to eat their french fries with mayonnaise instead of ketchup; in other parts of the world, specific hand gestures (such as the thumbs-up) are offensive.
- International shipping. Shipping products between countries in a timely manner can be challenging. Inconsistent address formats, dishonest customs agents, and prohibitive shipping costs are all factors that must be considered when trying to deliver products internationally.

Because of these challenges, many businesses choose not to expand globally, either for labor or for customers. Whether a business has its own website or relies on a third-party, such as Amazon or eBay, the question of whether or not to globalize must be carefully considered.

Sidebar: How Does My Internet Speed Compare?

How does your Internet speed compare with others in your state, country, or around the world? The chart below shows how Internet speeds compare in different countries. You can find the full list of countries by going to this article (http://royal.pingdom.com/2010/11/12/real-connection-speeds-for-internet-users-across-the-world/). You can also compare the evolution of Internet speeds among countries by using this tool (http://www.akamai.com/stateoftheinternet/).

Internet speeds by country. Click to enlarge.

So how does your own Internet speed compare? There are many online tools you can use to determine the speed at which you are connected. One of the most trusted sites is speedtest.net, where you can test both your download speeds and upload speeds.

12.11.3: Digital Divide

The Digital Divide

As the Internet continues to make inroads across the world, it is also creating a separation between those who have access to this global network and those who do not. This separation is called the "digital divide" and is of great concern. An article in *Crossroads* puts it this way[4]:

> Adopted by the ACM Council in 1992, the ACM Code of Ethics and Professional Conduct focuses on issues involving the Digital Divide that could prevent certain categories of people — those from low-income households, senior citizens, single-parent children, the undereducated, minorities, and residents of rural areas — from receiving adequate access to the wide variety of resources offered by computer technology. This Code of Ethics positions the use of computers as a fundamental ethical consideration: "In a fair society, all individuals would have equal opportunity to participate in, or benefit from, the use of computer resources regardless of race, sex, religion, age, disability, national origin, or other similar factors." This article summarizes the digital divide in its various forms, and analyzes reasons for the growing inequality in people's access to Internet services. It also describes how society can bridge the digital divide: the serious social gap between information "haves" and "have-nots."

The digital divide can occur between countries, regions, or even neighborhoods. In many US cities, there are pockets with little or no Internet access, while just a few miles away high-speed broadband is common.

Solutions to the digital divide have had mixed success over the years. Many times, just providing Internet access and/or computing devices is not enough to bring true Internet access to a country, region, or neighborhood.

12.11.4: OLPC

One Laptop per Child

One attempt to repair the digital divide was the One Laptop per Child effort. As stated on the organization's website, "The mission of One Laptop per Child (OLPC) is to empower the children of developing countries to learn by providing one connected laptop to every school-age child. In order to accomplish our goal, we need people who believe in what we're doing and want to help make education for the world's children a priority, not a privilege."[5] Announced to great fanfare in 2005 by Nicholas Negroponte, the OLPC project seemed destined for success.

The XO laptop. Click to enlarge. (CC-BY: Mike McGregor)

The centerpiece of the project was the laptop itself: an inexpensive computer designed to withstand a lot of punishment. It utilized a revolutionary "mesh" network, allowing the laptops to act as repeaters, extending a Wi-Fi network far beyond their normal range. They also used minimal power, making them practical for remote areas with limited access to the electrical grid.

Unfortunately, the OLPC project failed to live up to expectations, running into many of the problems related to globalization discussed above: different cultures, corruption, and competition. In an article that examined the success and failures of OLPC, the authors state, "Expecting a laptop to cause such a revolutionary change showed a degree of naivete, even for an organization with the best of intentions and the smartest of people."[6] Today, OLPC is evolving their methods and their technology, trying to deliver an OLPC tablet computer.

12.11.5: Changing the Divide

A New Understanding of the Digital Divide

In 2006, web-usability consultant Jakob Nielsen wrote an article that got to the heart of our understanding of this problem. In his article, he breaks the digital divide up into three stages: the economic divide, the usability divide, and the empowerment divide[7].

What is usually called the digital divide is, in Nielsen's terms, the *economic divide*: the idea that some people can afford to have a computer and Internet access while others cannot. Because of Moore's Law (see chapter 2), the price of hardware has continued to drop and, at this point, we can now access digital technologies, such as smartphones, for very little. This fact, Nielsen asserts, means that for all intents and purposes, the economic divide is a moot point and we should not focus our resources on solving it.

The *usability divide* is concerned with the fact that "technology remains so complicated that many people couldn't use a computer even if they got one for free." And even for those who can use a computer, accessing all the benefits of having one is beyond their understanding. Included in this group are those with low literacy and seniors. According to Nielsen, we know how to help these users, but we are not doing it because there is little profit in doing so.

The *empowerment divide* is the most difficult to solve. It is concerned with how we use technology to empower ourselves. Very few users truly understand the power that digital technologies can give them. In his article, Nielsen explains that his (and others') research has shown that very few users contribute content to the Internet, use advanced search, or can even distinguish paid search ads from organic search results. Many people will limit what they can do online by accepting the basic, default settings of their computer and not work to understand how they can truly be empowered.

Understanding the digital divide using these three stages provides a more nuanced view of how we can work to alleviate it. While efforts such as One Laptop per Child are an excellent start, more work needs to be done to address the second and third stages of the digital divide for a more holistic solution.

Sidebar: Using Gaming to Bridge the Digital Divide

Paul Kim, the Assistant Dean and Chief Technology Officer of the Stanford Graduate School of Education, designed a project to address the digital divide for children in developing countries. [8] In their project, the researchers wanted to understand if children can adopt and teach themselves mobile learning technology, without help from teachers or other adults, and the processes and factors involved in this phenomenon. The researchers developed a mobile device called TeacherMate, which contained a game designed to help children learn math. The unique part of this research was that the researchers interacted directly with the children; they did not channel the mobile devices through the teachers or the schools. Another important factor to consider: in order to understand the context of the children's educational environment, the researchers began the project by working with parents and local nonprofits six months before their visit. While the results of this research are too detailed to go into here, it can be said that the researchers found that children can, indeed, adopt and teach themselves mobile learning technologies.

What makes this research so interesting when thinking about the digital divide is that the researchers found that, in order to be effective, they had to customize their technology and tailor their implementation to the specific group they were trying to reach. One of their conclusions stated the following:

Considering the rapid advancement of technology today, mobile learning options for future projects will only increase. Consequently, researchers must continue to investigate their impact; we believe there is a specific need for more in-depth studies on ICT [information and communication technology] design variations to meet different challenges of different localities.

To read more about Dr. Kim's project, locate the paper referenced in this sidebar.

12.11.6: Summary

Summary

Information technology has driven change on a global scale. As documented by Castells and Friedman, technology has given us the ability to integrate with people all over the world using digital tools. These tools have allowed businesses to broaden their labor pools, their markets, and even their operating hours. But they have also brought many new complications for businesses, which now must understand regulations, preferences, and cultures from many different nations. This new globalization has also exacerbated the digital divide. Nielsen has suggested that the digital divide consists of three stages (economic, usability, and empowerment), of which the economic stage is virtually solved.

Study Questions

1. What does the term *globalization* mean?
2. How does Friedman define the three eras of globalization?
3. Which technologies have had the biggest effect on globalization?
4. What are some of the advantages brought about by globalization?
5. What are the challenges of globalization?
6. What does the term *digital divide* mean?
7. What are Jakob Nielsen's three stages of the digital divide?
8. What was one of the key points of *The Rise of the Network Society*?
9. Which country has the highest average Internet speed? How does your country compare?
10. What is the OLPC project? Has it been successful?

Exercises

1. Compare the concept of Friedman's "Globalization 3.0″ with Nielsen empowerment stage of the digital divide.
2. Do some original research to determine some of the regulations that a US company may have to consider before doing business in one of the following countries: China, Germany, Saudi Arabia, Turkey.
3. Go to speedtest.net to determine your Internet speed. Compare your speed at home to the Internet speed at two other locations, such as your school, place of employment, or local coffee shop. Write up a one-page summary that compares these locations.
4. Give one example of the digital divide and describe what you would do to address it.
5. How did the research conducted by Paul Kim address the three levels of the digital divide?

---

1. http://internetworldstats.com/↵
2. Manuel Castells. 2000. The Rise of the Network Society (2nd ed.). Blackwell Publishers, Inc., Cambridge, MA, USA.↵
3. Friedman, T. L. (2005). The world is flat: A brief history of the twenty-first century. New York: Farrar, Straus and Giroux.↵
4. Kibum Kim. 2005. Challenges in HCI: digital divide. *Crossroads* 12, 2 (December 2005), 2-2. DOI=10.1145/1144375.1144377 http://doi.acm.org/10.1145/1144375.1144377↵
5. http://laptop.org/en/vision/mission/↵
6. One Laptop Per Child: Vision vs. Reality By Kenneth L. Kraemer, Jason Dedrick, Prakul Sharma Communications of the ACM, Vol. 52 No. 6, Pages 66-73↵
7. http://www.nngroup.com/articles/digi...-three-stages/↵
8. Kim, P., Buckner, E., Makany, T., & Kim, H. (2011). A comparative analysis of a game-based mobile learning model in low-socioeconomic communities of India. International Journal of Educational Development. doi:10.1016/j.ijedudev.2011.05.008.↵

SECTION OVERVIEW

12.12: The Ethical and Legal Implications of Information Systems

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe what the term *information systems ethics* means;
- explain what a code of ethics is and describe the advantages and disadvantages;
- define the term *intellectual property* and explain the protections provided by copyright, patent, and trademark; and
- describe the challenges that information technology brings to individual privacy.

Introduction

Information systems have had an impact far beyond the world of business. New technologies create new situations that we have never dealt with before. How do we handle the new capabilities that these devices empower us with? What new laws are going to be needed to protect us from ourselves? This chapter will kick off with a discussion of the impact of information systems on how we behave (ethics). This will be followed with the new legal structures being put in place, with a focus on intellectual property and privacy.

12.12.1: Ethics

Information Systems Ethics

The term *ethics* is defined as “a set of moral principles” or “the principles of conduct governing an individual or a group.”[1] Since the dawn of civilization, the study of ethics and their impact has fascinated mankind. But what do ethics have to do with information systems?

The introduction of new technology can have a profound effect on human behavior. New technologies give us capabilities that we did not have before, which in turn create environments and situations that have not been specifically addressed in ethical terms. Those who master new technologies gain new power; those who cannot or do not master them may lose power. In 1913, Henry Ford implemented the first moving assembly line to create his Model T cars. While this was a great step forward technologically (and economically), the assembly line reduced the value of human beings in the production process. The development of the atomic bomb concentrated unimaginable power in the hands of one government, who then had to wrestle with the decision to use it. Today’s digital technologies have created new categories of ethical dilemmas.

For example, the ability to anonymously make perfect copies of digital music has tempted many music fans to download copyrighted music for their own use without making payment to the music’s owner. Many of those who would never have walked into a music store and stolen a CD find themselves with dozens of illegally downloaded albums.

Digital technologies have given us the ability to aggregate information from multiple sources to create profiles of people. What would have taken weeks of work in the past can now be done in seconds, allowing private organizations and governments to know more about individuals than at any time in history. This information has value, but also chips away at the privacy of consumers and citizens.

12.12.1.1: Code of Ethics

Code of Ethics

One method for navigating new ethical waters is a code of ethics. A code of ethics is a document that outlines a set of acceptable behaviors for a professional or social group; generally, it is agreed to by all members of the group. The document details different actions that are considered appropriate and inappropriate.

A good example of a code of ethics is the *Code of Ethics and Professional Conduct* of the Association for Computing Machinery,[2] an organization of computing professionals that includes academics, researchers, and practitioners. Here is a quote from the preamble:

Commitment to ethical professional conduct is expected of every member (voting members, associate members, and student members) of the Association for Computing Machinery (ACM).

This Code, consisting of 24 imperatives formulated as statements of personal responsibility, identifies the elements of such a commitment. It contains many, but not all, issues professionals are likely to face. Section 1 outlines fundamental ethical considerations, while Section 2 addresses additional, more specific considerations of professional conduct. Statements in Section 3 pertain more specifically to individuals who have a leadership role, whether in the workplace or in a volunteer capacity such as with organizations like ACM. Principles involving compliance with this Code are given in Section 4.

In the ACM's code, you will find many straightforward ethical instructions, such as the admonition to be honest and trustworthy. But because this is also an organization of professionals that focuses on computing, there are more specific admonitions that relate directly to information technology:

- No one should enter or use another's computer system, software, or data files without permission. One must always have appropriate approval before using system resources, including communication ports, file space, other system peripherals, and computer time.
- Designing or implementing systems that deliberately or inadvertently demean individuals or groups is ethically unacceptable.
- Organizational leaders are responsible for ensuring that computer systems enhance, not degrade, the quality of working life. When implementing a computer system, organizations must consider the personal and professional development, physical safety, and human dignity of all workers. Appropriate human-computer ergonomic standards should be considered in system design and in the workplace.

One of the major advantages of creating a code of ethics is that it clarifies the acceptable standards of behavior for a professional group. The varied backgrounds and experiences of the members of a group lead to a variety of ideas regarding what is acceptable behavior. While to many the guidelines may seem obvious, having these items detailed provides clarity and consistency. Explicitly stating standards communicates the common guidelines to everyone in a clear manner.

Having a code of ethics can also have some drawbacks. First of all, a code of ethics does not have legal authority; in other words, breaking a code of ethics is not a crime in itself. So what happens if someone violates one of the guidelines? Many codes of ethics include a section that describes how such situations will be handled. In many cases, repeated violations of the code result in expulsion from the group.

In the case of ACM: "Adherence of professionals to a code of ethics is largely a voluntary matter. However, if a member does not follow this code by engaging in gross misconduct, membership in ACM may be terminated." Expulsion from ACM may not have much of an impact on many individuals, since membership in ACM is usually not a requirement for employment. However, expulsion from other organizations, such as a state bar organization or medical board, could carry a huge impact.

Another possible disadvantage of a code of ethics is that there is always a chance that important issues will arise that are not specifically addressed in the code. Technology is quickly changing, and a code of ethics might not be updated often enough to keep up with all of the changes. A good code of ethics, however, is written in a broad enough fashion that it can address the ethical issues of potential changes to technology while the organization behind the code makes revisions.

Finally, a code of ethics could have also be a disadvantage in that it may not entirely reflect the ethics or morals of every member of the group. Organizations with a diverse membership may have internal conflicts as to what is acceptable behavior. For example, there may be a difference of opinion on the consumption of alcoholic beverages at company events. In such cases, the organization must make a choice about the importance of addressing a specific behavior in the code.

12.12.2: Acceptable Use

Acceptable Use Policy

Many organizations that provide technology services to a group of constituents or the public require agreement to an acceptable use policy (AUP) before those services can be accessed. Similar to a code of ethics, this policy outlines what is allowed and what is not allowed while someone is using the organization's services. An everyday example of this is the terms of service that must be agreed to before using the public Wi-Fi at Starbucks, McDonald's, or even a university. Here is an example of an acceptable use policy from Virginia Tech.

Just as with a code of ethics, these acceptable use policies specify what is allowed and what is not allowed. Again, while some of the items listed are obvious to most, others are not so obvious:

- "Borrowing" someone else's login ID and password is prohibited.
- Using the provided access for commercial purposes, such as hosting your own business website, is not allowed.
- Sending out unsolicited email to a large group of people is prohibited.

Also as with codes of ethics, violations of these policies have various consequences. In most cases, such as with Wi-Fi, violating the acceptable use policy will mean that you will lose your access to the resource. While losing access to Wi-Fi at Starbucks may not have a lasting impact, a university student getting banned from the university's Wi-Fi (or possibly all network resources) could have a large impact.

12.12.3: Intellectual Property

Intellectual Property

One of the domains that have been deeply impacted by digital technologies is the domain of intellectual property. Digital technologies have driven a rise in new intellectual property claims and made it much more difficult to defend intellectual property.

Intellectual property is defined as "property (as an idea, invention, or process) that derives from the work of the mind or intellect."[3] This could include creations such as song lyrics, a computer program, a new type of toaster, or even a sculpture.

Practically speaking, it is very difficult to protect an idea. Instead, intellectual property laws are written to protect the tangible results of an idea. In other words, just coming up with a song in your head is not protected, but if you write it down it can be protected.

Protection of intellectual property is important because it gives people an incentive to be creative. Innovators with great ideas will be more likely to pursue those ideas if they have a clear understanding of how they will benefit. In the US Constitution, Article 8, Section 8, the authors saw fit to recognize the importance of protecting creative works:

Congress shall have the power . . . To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.

An important point to note here is the "limited time" qualification. While protecting intellectual property is important because of the incentives it provides, it is also necessary to limit the amount of benefit that can be received and allow the results of ideas to become part of the public domain.

Outside of the US, intellectual property protections vary. You can find out more about a specific country's intellectual property laws by visiting the World Intellectual Property Organization.

In the following sections we will review three of the best-known intellectual property protections: copyright, patent, and trademark.

12.12.3.1: Copyright

Copyright

Copyright is the protection given to songs, computer programs, books, and other creative works; any work that has an "author" can be copyrighted. Under the terms of copyright, the author of a work controls what can be done with the work, including:

- Who can make copies of the work.
- Who can make derivative works from the original work.
- Who can perform the work publicly.
- Who can display the work publicly.
- Who can distribute the work.

Many times, a work is not owned by an individual but is instead owned by a publisher with whom the original author has an agreement. In return for the rights to the work, the publisher will market and distribute the work and then pay the original author a portion of the proceeds.

Copyright protection lasts for the life of the original author plus seventy years. In the case of a copyrighted work owned by a publisher or another third party, the protection lasts for ninety-five years from the original creation date. For works created before 1978, the protections vary slightly. You can see the full details on copyright protections by reviewing the Copyright Basics document available at the US Copyright Office's website.

12.12.3.1.1: Copyright details

Obtaining Copyright Protection

In the United States, a copyright is obtained by the simple act of creating the original work. In other words, when an author writes down that song, makes that film, or designs that program, he or she automatically has the copyright. However, for a work that will be used commercially, it is advisable to register for a copyright with the US Copyright Office. A registered copyright is needed in order to bring legal action against someone who has used a work without permission.

First Sale Doctrine

If an artist creates a painting and sells it to a collector who then, for whatever reason, proceeds to destroy it, does the original artist have any recourse? What if the collector, instead of destroying it, begins making copies of it and sells them? Is this allowed? The first sale doctrine is a part of copyright law that addresses this, as shown below[4]:

The first sale doctrine, codified at 17 U.S.C. § 109, provides that an individual who knowingly purchases a copy of a copyrighted work from the copyright holder receives the right to sell, display or otherwise dispose of *that particular copy*, notwithstanding the interests of the copyright owner.

So, in our examples, the copyright owner has no recourse if the collector destroys her artwork. But the collector does not have the right to make copies of the artwork.

Fair Use

Another important provision within copyright law is that of fair use. Fair use is a limitation on copyright law that allows for the use of protected works without prior authorization in specific cases. For example, if a teacher wanted to discuss a current event in her class, she could pass out copies of a copyrighted news story to her students without first getting permission. Fair use is also what allows a student to quote a small portion of a copyrighted work in a research paper.

Unfortunately, the specific guidelines for what is considered fair use and what constitutes copyright violation are not well defined. Fair use is a well-known and respected concept and will only be challenged when copyright holders feel that the integrity or market value of their work is being threatened. The following four factors are considered when determining if something constitutes fair use: [5]

1. The purpose and character of the use, including whether such use is of commercial nature or is for nonprofit educational purposes;
2. The nature of the copyrighted work;
3. The amount and substantiality of the portion used in relation to the copyrighted work as a whole;
4. The effect of the use upon the potential market for, or value of, the copyrighted work.

If you are ever considering using a copyrighted work as part of something you are creating, you may be able to do so under fair use. However, it is always best to check with the copyright owner to be sure you are staying within your rights and not infringing upon theirs.

12.12.3.1.2: Digital Millennium Copyright

The Digital Millennium Copyright Act

As digital technologies have changed what it means to create, copy, and distribute media, a policy vacuum has been created. In 1998, the US Congress passed the Digital Millennium Copyright Act (DMCA), which extended copyright law to take into consideration digital technologies. Two of the best-known provisions from the DMCA are the anti-circumvention provision and the "safe harbor" provision.

- The anti-circumvention provision makes it illegal to create technology to circumvent technology that has been put in place to protect a copyrighted work. This provision includes not just the creation of the technology but also the publishing of information that describes how to do it. While this provision does allow for some exceptions, it has become quite controversial and has led to a movement to have it modified.
- The "safe harbor" provision limits the liability of online service providers when someone using their services commits copyright infringement. This is the provision that allows YouTube, for example, not to be held liable when someone posts a clip from a copyrighted movie. The provision does require the online service provider to take action when they are notified of the violation (a "takedown" notice). For an example of how takedown works, here's how YouTube handles these requests: YouTube Copyright Infringement Notification.

Many think that the DMCA goes too far and ends up limiting our freedom of speech. The Electronic Frontier Foundation (EFF) is at the forefront of this battle. For example, in discussing the anti-circumvention provision, the EFF states:

Yet the DMCA has become a serious threat that jeopardizes fair use, impedes competition and innovation, chills free expression and scientific research, and interferes with computer intrusion laws. If you circumvent DRM [digital rights management] locks for non-infringing fair uses or create the tools to do so you might be on the receiving end of a lawsuit.

Sidebar: Creative Commons

In chapter 2, we learned about open-source software. Open-source software has few or no copyright restrictions; the creators of the software publish their code and make their software available for others to use and distribute for free. This is great for software, but what about other forms of copyrighted works? If an artist or writer wants to make their works available, how can they go about doing so while still protecting the integrity of their work? Creative Commons is the solution to this problem.

Creative Commons is a nonprofit organization that provides legal tools for artists and authors. The tools offered make it simple to license artistic or literary work for others to use or distribute in a manner consistent with the author's intentions. Creative Commons licenses are indicated with the symbol . It is important to note that Creative Commons and public domain are not the same. When something is in the public domain, it has absolutely no restrictions on its use or distribution. Works whose copyrights have expired, for example, are in the public domain.

By using a Creative Commons license, authors can control the use of their work while still making it widely accessible. By attaching a Creative Commons license to their work, a legally binding license is created. Here are some examples of these licenses:

- CC-BY: This is the least restrictive license. It lets others distribute and build upon the work, even commercially, as long as they give the author credit for the original work.
- CC-BY-SA: This license restricts the distribution of the work via the "share-alike" clause. This means that others can freely distribute and build upon the work, but they must give credit to the original author *and* they must share using the same Creative Commons license.
- CC-BY-NC: This license is the same as CC-BY but adds the restriction that no one can make money with this work. NC stands for "non-commercial."
- CC-BY-NC-ND: This license is the same as CC-BY-NC but also adds the ND restriction, which means that no derivative works may be made from the original.

These are a few of the more common licenses that can be created using the tools that Creative Commons makes available. For a full listing of the licenses and to learn much more about Creative Commons, visit their web site.

12.12.3.2: Patent

Patent

Another important form of intellectual property protection is the patent. A patent creates protection for someone who invents a new product or process. The definition of invention is quite broad and covers many different fields. Here are some examples of items receiving patents:

- circuit designs in semiconductors;
- prescription drug formulas;
- firearms;
- locks;
- plumbing;
- engines;
- coating processes; and
- business processes.

Once a patent is granted, it provides the inventor with protection from others infringing on his or her patent. A patent holder has the right to "exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States for a limited time in exchange for public disclosure of the invention when the patent is granted."[6]

As with copyright, patent protection lasts for a limited period of time before the invention or process enters the public domain. In the US, a patent lasts twenty years. This is why generic drugs are available to replace brand-name drugs after twenty years.

12.12.3.2.1: Obtain Protection

Obtaining Patent Protection

Unlike copyright, a patent is not automatically granted when someone has an interesting idea and writes it down. In most countries, a patent application must be submitted to a government patent office. A patent will only be granted if the invention or process being submitted meets certain conditions:

- It must be original. The invention being submitted must not have been submitted before.
- It must be non-obvious. You cannot patent something that anyone could think of. For example, you could not put a pencil on a chair and try to get a patent for a pencil-holding chair.
- It must be useful. The invention being submitted must serve some purpose or have some use that would be desired.

The job of the patent office is to review patent applications to ensure that the item being submitted meets these requirements. This is not an easy job: in 2012, the US Patent Office received 576,763 patent applications and granted 276,788 patents. The current backlog for a patent approval is 18.1 months. Over the past fifty years, the number of patent applications has risen from just 100,000 a year to almost 600,000; digital technologies are driving much of this innovation.

Increase in patent applications, 1963–2012 (Source: US Patent and Trademark Office)

Sidebar: What Is a Patent Troll?

The advent of digital technologies has led to a large increase in patent filings and therefore a large number of patents being granted. Once a patent is granted, it is up to the owner of the patent to enforce it; if someone is found to be using the invention without permission, the patent holder has the right to sue to force that person to stop and to collect damages.

The rise in patents has led to a new form of profiteering called patent trolling. A patent troll is a person or organization who gains the rights to a patent but does not actually make the invention that the patent protects. Instead, the patent troll searches for those who are illegally using the invention in some way and sues them. In many cases, the infringement being alleged is questionable at best. For example, companies have been sued for using Wi-Fi or for scanning documents, technologies that have been on the market for many years.

Recently, the US government has begun taking action against patent trolls. Several pieces of legislation are working their way through Congress that will, if enacted, limit the ability of patent trolls to threaten innovation. You can learn a lot more about patent trolls by listening to a detailed investigation conducted by the radio program *This American Life*, by clicking this link.

12.12.3.3: Trademark

Trademark

A trademark is a word, phrase, logo, shape or sound that identifies a source of goods or services. For example, the Nike "Swoosh," the Facebook "f", and Apple's apple (with a bite taken out of it) are all trademarked. The concept behind trademarks is to protect the consumer. Imagine going to the local shopping center to purchase a specific item from a specific store and finding that there are several stores all with the same name!

Two types of trademarks exist – a common-law trademark and a registered trademark. As with copyright, an organization will automatically receive a trademark if a word, phrase, or logo is being used in the normal course of business (subject to some restrictions, discussed below). A common-law trademark is designated by placing "TM" next to the trademark. A registered trademark is one that has been examined, approved, and registered with the trademark office, such as the Patent and Trademark Office in the US. A registered trademark has the circle-R (®) placed next to the trademark.

While most any word, phrase, logo, shape, or sound can be trademarked, there are a few limitations. A trademark will not hold up legally if it meets one or more of the following conditions:

- The trademark is likely to cause confusion with a mark in a registration or prior application.
- The trademark is merely descriptive for the goods/services. For example, trying to register the trademark "blue" for a blue product you are selling will not pass muster.
- The trademark is a geographic term.
- The trademark is a surname. You will not be allowed to trademark "Smith's Bookstore."
- The trademark is ornamental as applied to the goods. For example, a repeating flower pattern that is a design on a plate cannot be trademarked.

As long as an organization uses its trademark and defends it against infringement, the protection afforded by it does not expire. Because of this, many organizations defend their trademark against other companies whose branding even only slightly copies their trademark. For example, Chick-fil-A has trademarked the phrase "Eat Mor Chikin" and has vigorously defended it against a small business using the slogan "Eat More Kale." Coca-Cola has trademarked the contour shape of its bottle and will bring legal action against any company using a bottle design similar to theirs. As an example of trademarks that have been diluted and have now lost their protection in the US are "aspirin" (originally trademarked by Bayer), "escalator" (originally trademarked by Otis), and "yo-yo" (originally trademarked by Duncan).

Information Systems and Intellectual Property

The rise of information systems has forced us to rethink how we deal with intellectual property. From the increase in patent applications swamping the government's patent office to the new laws that must be put in place to enforce copyright protection, digital technologies have impacted our behavior.

12.12.4: Privacy

Privacy

The term *privacy* has many definitions, but for our purposes, privacy will mean the ability to control information about oneself. Our ability to maintain our privacy has eroded substantially in the past decades, due to information systems.

Personally Identifiable Information

Information about a person that can be used to uniquely establish that person's identify is called personally identifiable information, or PII. This is a broad category that includes information such as:

- name;
- social security number;
- date of birth;
- place of birth;
- mother's maiden name;
- biometric records (fingerprint, face, etc.);
- medical records;
- educational records;
- financial information; and
- employment information.

Organizations that collect PII are responsible to protect it. The Department of Commerce recommends that "organizations minimize the use, collection, and retention of PII to what is strictly necessary to accomplish their business purpose and mission." They go on to state that "the likelihood of harm caused by a breach involving PII is greatly reduced if an organization minimizes the amount of PII it uses, collects, and stores."[7] Organizations that do not protect PII can face penalties, lawsuits, and loss of business. In the US, most states now have laws in place requiring organizations that have had security breaches related to PII to notify potential victims, as does the European Union.

Just because companies are required to protect your information does not mean they are restricted from sharing it. In the US, companies can share your information without your explicit consent (see sidebar below), though not all do so. Companies that collect PII are urged by the FTC to create a privacy policy and post it on their website. The state of California requires a privacy policy for any website that does business with a resident of the state (see http://www.privacy.ca.gov/lawenforcement/laws.htm).

While the privacy laws in the US seek to balance consumer protection with promoting commerce, in the European Union privacy is considered a fundamental right that outweighs the interests of commerce. This has led to much stricter privacy protection in the EU, but also makes commerce more difficult between the US and the EU.

12.12.4.1: Other Privacy Laws

Restrictions on Record Collecting

In the US, the government has strict guidelines on how much information can be collected about its citizens. Certain classes of information have been restricted by laws over time, and the advent of digital tools has made these restrictions more important than ever.

Children's Online Privacy Protection Act

Websites that are collecting information from children under the age of thirteen are required to comply with the Children's Online Privacy Protection Act (COPPA), which is enforced by the Federal Trade Commission (FTC). To comply with COPPA, organizations must make a good-faith effort to determine the age of those accessing their websites and, if users are under thirteen years old, must obtain parental consent before collecting any information.

Family Educational Rights and Privacy Act

The Family Educational Rights and Privacy Act (FERPA) is a US law that protects the privacy of student education records. In brief, this law specifies that parents have a right to their child's educational information until the child reaches either the age of eighteen or begins attending school beyond the high school level. At that point, control of the information is given to the child. While this law is not specifically about the digital collection of information on the Internet, the educational institutions that are collecting student information are at a higher risk for disclosing it improperly because of digital technologies.

Health Insurance Portability and Accountability Act

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) is the law the specifically singles out records related to health care as a special class of personally identifiable information. This law gives patients specific rights to control their medical records, requires health care providers and others who maintain this information to get specific permission in order to share it, and imposes penalties on the institutions that breach this trust. Since much of this information is now shared via electronic medical records, the protection of those systems becomes paramount.

Sidebar: Do Not Track

When it comes to getting permission to share personal information, the US and the EU have different approaches. In the US, the "opt-out" model is prevalent; in this model, the default agreement is that you have agreed to share your information with the organization and must explicitly tell them that you do not want your information shared. There are no laws prohibiting the sharing of your data (beyond some specific categories of data, such as medical records). In the European Union, the "opt-in" model is required to be the default. In this case, you must give your explicit permission before an organization can share your information.

To combat this sharing of information, the Do Not Track initiative was created. As its creators explain[8]:

Do Not Track is a technology and policy proposal that enables users to opt out of tracking by websites they do not visit, including analytics services, advertising networks, and social platforms. At present few of these third parties offer a reliable tracking opt out, and tools for blocking them are neither user-friendly nor comprehensive. Much like the popular Do Not Call registry, Do Not Track provides users with a single, simple, persistent choice to opt out of third-party web tracking.

12.12.4.2: Non-Obvious Relationship

Non-Obvious Relationship Awareness

Digital technologies have given us many new capabilities that simplify and expedite the collection of personal information. Every time we come into contact with digital technologies, information about us is being made available. From our location to our web-surfing habits, our criminal record to our credit report, we are constantly being monitored. This information can then be aggregated to create profiles of each and every one of us. While much of the information collected was available in the past, collecting it and combining it took time and effort. Today, detailed information about us is available for purchase from different companies. Even information not categorized as PII can be aggregated in such a way that an individual can be identified.

This process of collecting large quantities of a variety of information and then combining it to create profiles of individuals is known as non-obvious relationship awareness, or NORA. First commercialized by big casinos looking to find cheaters, NORA is used by both government agencies and private organizations, and it is big business.

Non-obvious relationship awareness (NORA)

In some settings, NORA can bring many benefits, such as in law enforcement. By being able to identify potential criminals more quickly, crimes can be solved more quickly or even prevented before they happen. But these advantages come at a price: our privacy.

12.12.5: Section 2 Summary

Summary

The rapid changes in information technology in the past few decades have brought a broad array of new capabilities and powers to governments, organizations, and individuals alike. These new capabilities have required thoughtful analysis and the creation of new norms, regulations, and laws. In this chapter, we have seen how the areas of intellectual property and privacy have been affected by these new capabilities and how the regulatory environment has been changed to address them.

Study Questions

1. What does the term *information systems ethics* mean?
2. What is a code of ethics? What is one advantage and one disadvantage of a code of ethics?
3. What does the term *intellectual property* mean? Give an example.
4. What protections are provided by a copyright? How do you obtain one?
5. What is fair use?
6. What protections are provided by a patent? How do you obtain one?
7. What does a trademark protect? How do you obtain one?
8. What does the term *personally identifiable information* mean?
9. What protections are provided by HIPAA, COPPA, and FERPA?
10. How would you explain the concept of NORA?

Exercises

1. Provide one example of how information technology has created an ethical dilemma that would not have existed before the advent of information technology.
2. Find an example of a code of ethics or acceptable use policy related to information technology and highlight five points that you think are important.
3. Do some original research on the effort to combat patent trolls. Write a two-page paper that discusses this legislation.
4. Give an example of how NORA could be used to identify an individual.
5. How are intellectual property protections different across the world? Pick two countries and do some original research, then compare the patent and copyright protections offered in those countries to those in the US. Write a two- to three-page paper describing the differences.

---

1. http://www.merriam-webster.com/dictionary/ethics↵
2. ACM Code of Ethics and Professional Conduct Adopted by ACM Council 10/16/92.↵
3. http://www.merriam-webster.com/dicti...ual%20property↵
4. http://www.justice.gov/usao/eousa/fo...9/crm01854.htm↵
5. http://www.copyright.gov/fls/fl102.html↵
6. From the US Patent and Trademark Office, "What Is A Patent?" http://www.uspto.gov/patents/↵
7. *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)*. National Institute of Standards and Technology, US Department of Commerce Special Publication 800-122. http://csrc.nist.gov/publications/ni.../sp800-122.pdf↵
8. http://donottrack.us/↵

12.13: Future Trends in Information Systems

Learning Objectives

Upon successful completion of this chapter, you will be able to:

- describe future trends in information systems.

Introduction

Information systems have evolved at a rapid pace ever since their introduction in the 1950s. Today, devices that we can hold in one hand are more powerful than the computers used to land a man on the moon. The Internet has made the entire world accessible to us, allowing us to communicate and collaborate with each other like never before. In this chapter, we will examine current trends and look ahead to what is coming next.

12.13.1: Global

Global

The first trend to note is the continuing expansion of globalization. The use of the Internet is growing all over the world, and with it the use of digital devices. The growth is coming from some unexpected places; countries such as Indonesia and Iran are leading the way in Internet growth. Look at the Growth column in the following table - how much the Internet has grown in the last 20 years.

WORLD INTERNET USAGE AND POPULATION STATISTICS
JUNE, 2019 - Updated

World Regions	Population (2019 Est.)	Population % of World	Internet Users 30 June 2019	Penetration Rate (% Pop.)	Growth 2000-2019	Internet World %
Africa	1,320,038,716	17.1 %	525,148,631	39.8 %	11,533 %	11.9 %
Asia	4,241,972,790	55.0 %	2,200,658,148	51.9 %	1,825 %	49.8 %
Europe	829,173,007	10.7 %	719,413,014	86.8 %	585 %	16.3 %
Latin America / Caribbean	658,345,826	8.5 %	447,495,130	68.0 %	2,377 %	10.1 %
Middle East	258,356,867	3.3 %	173,576,793	67.2 %	5,184 %	3.9 %
North America	366,496,802	4.7 %	327,568,628	89.4 %	203 %	7.4 %
Oceania / Australia	41,839,201	0.5 %	28,634,278	68.4 %	276 %	0.6 %
WORLD TOTAL	7,716,223,209	100.0 %	4,422,494,622	57.3 %	1,125 %	100.0 %

Global Internet growth, 2008–2012. Click to enlarge. (Source: Internet World Stats)

12.13.2: Social and Personal

Social

Social media growth is another trend that continues. Facebook now has over one billion users! In 2013, 80% of Facebook users were outside of the US and Canada.[1] Countries where Facebook is growing rapidly include Indonesia, Mexico, and the Philippines. [2]

Besides Facebook, other social media sites are also seeing tremendous growth. Over 70% of YouTube's users are outside the US, with the UK, India, Germany, Canada, France, South Korea, and Russia leading the way.[3] Pinterest gets over 50% of its users from outside the US, with over 9% from India. [4] Twitter now has over 230 million active users. [5] Social media sites not based in the US are also growing. China's QQ instant-messaging service is the eighth most-visited site in the world.[6]

Personal

Ever since the advent of Web 2.0 and e-commerce, users of information systems have expected to be able to modify their experiences to meet their personal tastes. From custom backgrounds on computer desktops to unique ringtones on mobile phones, makers of digital devices provide the ability to personalize how we use them. More recently, companies such as Netflix have begun assisting their users with personalizations by making suggestions. In the future, we will begin seeing devices perfectly matched to our personal preferences, based upon information collected about us in the past.

12.13.3: Mobile

Mobile

Perhaps the most impactful trend in digital technologies in the last decade has been the advent of mobile technologies. Beginning with the simple cellphone in the 1990s and evolving into the smartphones and tablets of today, the growth of mobile has been overwhelming. Here are some key indicators of this trend:

- Mobile device sales. In 2011, smartphones began outselling personal computers.[7]
- The number of smartphone subscribers grew at 31% in 2013, with China leading the way at 354 million smartphone users.
- Internet access via mobile. In May of 2013, mobile accounted for 15% of all Internet traffic. In China, 75% of Internet users used their smartphone to access it. Facebook reported that 68% of its active users used their mobile platform to access the social network.
- The rise of tablets. While Apple defined the smartphone with the iPhone, the iPad sold more than three times as many units in its first twelve months as the iPhone did in its first twelve months. Tablet shipments now outpace notebook PCs and desktop PCs. The research firm IDC predicts that 87% of all connected devices will be either smartphones or tablets by 2017. [8]

12.13.4: Wearable

Wearable

The average smartphone user looks at his or her smartphone 150 times a day for functions such as messaging (23 times), phone calls (22), listening to music (13), and social media (9).[9] Many of these functions would be much better served if the technology was worn on, or even physically integrated into, our bodies. This technology is known as a “wearable.”

Google Glass. Click to enlarge. (CC-BY: Flickr, user Tedeytan)

Wearables have been around for a long time, with technologies such as hearing aids and, later, bluetooth earpieces. But now, we are seeing an explosion of new wearable technologies. Perhaps the best known of these is Google Glass, an augmented reality device that you wear over your eyes like a pair of eyeglasses. Visible only to you, Google Glass will project images into your field of vision based on your context and voice commands. You can find out much more about Google Glass at http://www.google.com/glass/start/.

Another class of wearables are those related to health care. The UP by Jawbone consists of a wristband and an app that tracks how you sleep, move, and eat, then helps you use that information to feel your best. [10] It can be used to track your sleep patterns, moods, eating patterns, and other aspects of daily life, and then report back to you via an app on your smartphone or tablet. As our population ages and technology continues to evolve, there will be a large increase in wearables like this.

12.13.5: Collaborative

Collaborative

As more of us use smartphones and wearables, it will be simpler than ever to share data with each other for mutual benefit. Some of this sharing can be done passively, such as reporting our location in order to update traffic statistics. Other data can be reported actively, such as adding our rating of a restaurant to a review site.

The smartphone app Waze is a community-based tool that keeps track of the route you are traveling and how fast you are making your way to your destination. In return for providing your data, you can benefit from the data being sent from all of the other users of the app. Waze will route you around traffic and accidents based upon real-time reports from other users.

Yelp! allows consumers to post ratings and reviews of local businesses into a database, and then it provides that data back to consumers via its website or mobile phone app. By compiling ratings of restaurants, shopping centers, and services, and then allowing consumers to search through its directory, Yelp! has become a huge source of business for many companies. Unlike data collected passively however, Yelp! relies on its users to take the time to provide honest ratings and reviews.

Printable

One of the most amazing innovations to be developed recently is the 3-D printer. A 3-D printer allows you to print virtually any 3-D object based on a model of that object designed on a computer. 3-D printers work by creating layer upon layer of the model using malleable materials, such as different types of glass, metals, or even wax.

3-D printing is quite useful for prototyping the designs of products to determine their feasibility and marketability. 3-D printing has also been used to create working prosthetic legs, handguns, and even an ear that can hear beyond the range of normal hearing. The US Air Force now uses 3-D printed parts on the F-18 fighter jet.[11]

3-D printing is one of many technologies embraced by the "maker" movement. Chris Anderson, editor of *Wired* magazine, puts it this way[12]:

In a nutshell, the term "Maker" refers to a new category of builders who are using open-source methods and the latest technology to bring manufacturing out of its traditional factory context, and into the realm of the personal desktop computer. Until recently, the ability to manufacture was reserved for those who owned factories. What's happened over the last five years is that we've brought the Web's democratizing power to manufacturing. Today, you can manufacture with the push of a button.

12.13.6: Findable

Findable

The "Internet of Things" refers to the idea of physical objects being connected to the Internet. Advances in wireless technologies and sensors will allow physical objects to send and receive data about themselves. Many of the technologies to enable this are already available – it is just a matter of integrating them together.

In a 2010 report by McKinsey & Company on the Internet of Things[13], six broad applications are identified:

- Tracking behavior. When products are embedded with sensors, companies can track the movements of these products and even monitor interactions with them. Business models can be fine-tuned to take advantage of this behavioral data. Some insurance companies, for example, are offering to install location sensors in customers' cars. That allows these companies to base the price of policies on how a car is driven as well as where it travels.
- Enhanced situational awareness. Data from large numbers of sensors deployed, for example, in infrastructure (such as roads and buildings), or to report on environmental conditions (including soil moisture, ocean currents, or weather), can give decision makers a heightened awareness of real-time events, particularly when the sensors are used with advanced display or visualization technologies. Security personnel, for instance, can use sensor networks that combine video, audio, and vibration detectors to spot unauthorized individuals who enter restricted areas.
- Sensor-driven decision analysis. The Internet of Things also can support longer-range, more complex human planning and decision making. The technology requirements – tremendous storage and computing resources linked with advanced software systems that generate a variety of graphical displays for analyzing data – rise accordingly.
- Process optimization. Some industries, such as chemical production, are installing legions of sensors to bring much greater granularity to monitoring. These sensors feed data to computers, which in turn analyze the data and then send signals to actuators that adjust processes – for example, by modifying ingredient mixtures, temperatures, or pressures.
- Optimized resource consumption. Networked sensors and automated feedback mechanisms can change usage patterns for scarce resources, such as energy and water. This can be accomplished by dynamically changing the price of these goods to increase or reduce demand.
- Complex autonomous systems. The most demanding use of the Internet of Things involves the rapid, real-time sensing of unpredictable conditions and instantaneous responses guided by automated systems. This kind of machine decision-making mimics human reactions, though at vastly enhanced performance levels. The automobile industry, for instance, is stepping up the development of systems that can detect imminent collisions and take evasive action.

12.13.7: New Ideas

Autonomous

A final trend that is emerging is an extension of the Internet of Things: autonomous robots and vehicles. By combining software, sensors, and location technologies, devices that can operate themselves to perform specific functions are being developed. These take the form of creations such as medical nanotechnology robots (nanobots), self-driving cars, or unmanned aerial vehicles (UAVs).

A nanobot is a robot whose components are on the scale of about a nanometer, which is one-billionth of a meter. While still an emerging field, it is showing promise for applications in the medical field. For example, a set of nanobots could be introduced into the human body to combat cancer or a specific disease.

In March of 2012, Google introduced the world to their driverless car by releasing a video on YouTube showing a blind man driving the car around the San Francisco area. The car combines several technologies, including a laser radar system, worth about $150,000. While the car is not available commercially yet, three US states (Nevada, Florida, and California) have already passed legislation making driverless cars legal.

A UAV, often referred to as a “drone,” is a small airplane or helicopter that can fly without a pilot. Instead of a pilot, they are either run autonomously by computers in the vehicle or operated by a person using a remote control. While most drones today are used for military or civil applications, there is a growing market for personal drones. For around $300, a consumer can purchase a drone for personal use.

12.13.8: Summary

Summary

As the world of information technology moves forward, we will be constantly challenged by new capabilities and innovations that will both amaze and disgust us. As we learned in chapter 12, many times the new capabilities and powers that come with these new technologies will test us and require a new way of thinking about the world. Businesses and individuals alike need to be aware of these coming changes and prepare for them.

Study Questions

1. Which countries are the biggest users of the Internet? Social media? Mobile?
2. Which country had the largest Internet growth (in %) between 2008 and 2012?
3. How will most people connect to the Internet in the future?
4. What are two different applications of wearable technologies?
5. What are two different applications of collaborative technologies?
6. What capabilities do printable technologies have?
7. How will advances in wireless technologies and sensors make objects "findable"?
8. What is enhanced situational awareness?
9. What is a nanobot?
10. What is a UAV?

Exercises

1. If you were going to start a new technology business, which of the emerging trends do you think would be the biggest opportunity? Do some original research to estimate the market size.
2. What privacy concerns could be raised by collaborative technologies such as Waze?
3. Do some research about the first handgun printed using a 3-D printer and report on some of the concerns raised.
4. Write up an example of how the Internet of Things might provide a business with a competitive advantage.
5. How do you think wearable technologies could improve overall healthcare?
6. What potential problems do you see with a rise in the number of driverless cars? Do some independent research and write a two-page paper that describes where driverless cars are legal and what problems may occur.
7. Seek out the latest presentation by Mary Meeker on "Internet Trends" (if you cannot find it, the video from 2013 is available at http://allthingsd.com/20130529/mary-meekers-2013-internet-trends-deck-the-full-video/). Write a one-page paper describing what the top three trends are, in your opinion.

---

1. http://newsroom.fb.com/Key-Facts↵
2. http://www.socialbakers.com/blog/38-...ast-six-months↵
3. http://newmediarockstars.com/2013/03...a-infographic/↵
4. http://www.alexa.com/siteinfo/pinterest.com↵
5. https://about.twitter.com/company↵
6. http://www.alexa.com/siteinfo/qq.com↵
7. http://mashable.com/2012/02/03/smart...-overtake-pcs/↵
8. http://www.forbes.com/sites/louiscol...d-smartphones/↵
9. http://communities-dominate.blogs.co...le-moment.html↵
10. https://jawbone.com/up↵
11. http://www.economist.com/news/techno...inting-it-fast↵
12. Makers: The New Industrial Revolution by Chris Anderson. Crown Business; 2012.↵
13. http://www.mckinsey.com/insights/hig...rnet_of_things↵

CHAPTER OVERVIEW

Back Matter

Index

Detailed Licensing

Overview

Title: INT 1010: Concepts in Computing

Webpages: 238

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-SA 4.0: 59.2% (141 pages)
- Undeclared: 23.1% (55 pages)
- CC BY 4.0: 11.3% (27 pages)
- CC BY-NC-SA 4.0: 5.9% (14 pages)
- CC BY-NC 4.0: 0.4% (1 page)

By Page

- INT 1010: Concepts in Computing - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - About the Book - *CC BY-NC-SA 4.0*
    - Licensing - *Undeclared*
    - Acknowledgements - *Undeclared*
  - 1: Introduction to INT 1010 - Introduction to Information Technology - *Undeclared*
    - 1.1: Introduction - *Undeclared*
    - 1.3: Communication Etiquette - *CC BY-NC-SA 4.0*
    - 1.4: Summary - *Undeclared*
  - 2: Introduction to Information Systems - *CC BY 4.0*
    - 2.1: Information Systems - *CC BY 4.0*
    - 2.2: What is Competitive Advantage - *CC BY-SA 4.0*
    - 2.3: Components of an Information System - *CC BY-NC 4.0*
    - 2.4: Summary - *Undeclared*
  - 3: Hardware Components of an Information System - *Undeclared*
    - 3.1: Hardware - *Undeclared*
    - 3.2: Summary - *Undeclared*
  - 4: Software Component of an Information System - *CC BY-NC-SA 4.0*
    - 4.1: Computer Software - *CC BY-NC-SA 4.0*
    - 4.2: Operating Systems - *CC BY 4.0*
    - 4.3: File systems - *CC BY-NC-SA 4.0*
    - 4.4: Downloading Files - *CC BY 4.0*
    - 4.5: File Management - *CC BY 4.0*
    - 4.6: Summary - *Undeclared*
  - 5: Issues in Computing - *Undeclared*
    - 5.1: Information Systems Security - *Undeclared*
    - 5.2: The Ethical and Legal Implications of Information Systems - *CC BY-SA 4.0*
    - 5.3: Windows Security - *Undeclared*
    - 5.4: Summary - *Undeclared*
  - 6: Networking and Communication - *CC BY-NC-SA 4.0*
    - 6.1: Introduction to Networking and Communication - *CC BY-NC-SA 4.0*
    - 6.2: Network Basics - *Undeclared*
    - 6.3: Providing Resources in a Network - *CC BY-NC-SA 4.0*
    - 6.4: Internet Connections - *CC BY-NC-SA 4.0*
    - 6.5: The Changing Network Environment Network Trends - *CC BY-NC-SA 4.0*
    - 6.6: Network Security - *CC BY-NC-SA 4.0*
    - 6.7: Summary - *CC BY-NC-SA 4.0*
  - 7: Databases - *CC BY-NC-SA 4.0*
    - 7.1: Data and Databases - *CC BY 4.0*
    - 7.2: Before Database Systems - *Undeclared*
    - 7.3: The Relational Data Model and others - *CC BY 4.0*
    - 7.4: Databases and Security Issues - *CC BY-NC-SA 4.0*
    - 7.5: Fundamental Database Concepts - *CC BY 4.0*
    - 7.6: Appendix A - Designing a Database - *CC BY 4.0*
    - 7.7: Summary - *Undeclared*
  - 8: The People in Information Systems - *CC BY 4.0*
    - 8.1: Creators - *CC BY 4.0*
    - 8.2: Operations and Administration - *CC BY 4.0*
    - 8.3: Managers - *Undeclared*
    - 8.4: Computer and Information Technology Occupations - *Undeclared*
    - 8.5: Summary - *CC BY 4.0*
  - 9: Introduction to Web Development - *Undeclared*
    - 9.1: An Introduction - *Undeclared*