Computer Science 1032 – Class Notes

- A computer is simply an box-shaped object with lights – it requires software to function to its full capacity
- Computers are only aware of electronic pulses (whether they are on or off – in computer science, on = 1 and off = 0)
- 8 bits = 1 byte
- 1 bit is one pulse
- Information is transmitted between computers by the Internet and the world wide web
- Computer is based off of this binary code: [1,0,1,0,0,1,1]

- The Internet is an interconnection (transmission of pulses from websites), whereas the world wide web contains web pages made from HTML and stores all information on websites
- Each site has a storage of web pages
- Web pages are written in a code called HTML – this is displayed through web browsers (such as Explorer or Firefox)
- HTML (Hyper Text Markup Language):
  - ➢ Document tags
  - ➢ Formatting tags
  - ➢ Anchors, tables, lists
- Information transmission in a computer: user components (**user** tier), DBMS (database system – **database** tier), web browser (**server** tier)
- When a purchase is made over the Internet, pulses are sent to an updated database that records the order
- HTML is the first markup language of the world wide web – it is used to deliver content on the web, is an open standard, readable, portable and easy to use
- Structure of HTML file: **content** and/or **instructions**
- A web page must be published in order to be viewable online
- Important types of HTML coding:
  - ➢ <p> = text in a paragraph
  - ➢ = center the text
  - ➢ <HR> = horizontal rule
  - ➢ <h1> = header (1-6 from largest to smallest headers)
  - ➢
 = break
  - ➢ <b> = bold
  - ➢ <u> = underline
  - ➢ <i> = italics
  - ➢ <img (SRC – source) / > = image

- **HTTP** (hyper text transfer protocol)
- **FTP** (file transfer protocol)
- **DBMS** (data base management system)
- **SMTP** (simple mail transfer protocol)
- How to create a table in HTML:
  - ➢ <table> = create table
  - ➢ <tr> = table row (add additional tags for more table rows)
  - ➢ <td> = describes each column
  - ➢ Table border = 1 or 2
- How to create a list in HTML:
  - ➢ = ordered list
  - ➢ <L1> = indicates the items on the list
  - ➢ = unordered list
- **XML** (extensible markup language)
- Electronic data interchange allows for documents to be exchanged (they can be used over the internet) – they group data together and send one file

- Creating files in XML is done using Textedit

- Example of a namespace:
- In XML Schema, "maxOccurs=*any value*" means something can occur an unlimited amount of times, whereas "minOccurs" means it will only occur once

- David Ticoll suggests that in 10 years, there will be an unlimited amount of storage for free
- There will be a collision between the **real** world and the **virtual** world
- Canadian economy is becoming electronically and technologically based
- Information is output and data is input
- Data is collected, recorded, stored and processed
- Knowledge is derived from data
- Data is processed by summing, ordering, averaging, grouping and comparing

<u>Chapter 1 Notes</u>

- An **information system** is a group of components that interact to produce information
- All information systems comprise a **five-component framework** that includes: **computer hardware** (the electronic or associated gadgetry that constitute a computer system)**, software** (programs or **applications**)**, data** (basic building blocks of information)**, procedures** (instructions or processes that must be followed to achieve the objective) and **people** (interact with the system)

- **Management information systems (MIS)** comprise the development and use of information systems that help organizations achieve their goals – the definition of MIS has the following three key elements: **development and use** (one must contribute to the development of the information system and learn how to use it effectively)**, information systems** and **goals and objectives** (it must be made sure that the information system is effectively meeting the needs of the entity)
- **Information technology (IT)** comprises the hardware, software and data of an information system
- The sector tied to the use of information systems is the **Information and Communications Technology (ICT) sector** (included over 32,000 companies and generated revenues of over $162B)
- **Moore's Law** notes that the density of circuits on an integrated chip was doubling approximately every two years or so (this has been generally accurate over the last four decades)
- IT is embedded in IS

Chapter 2

- A **business process** is a series of tasks or steps designed to produce a product or service (also known as a **business system**) – manufacturing, marketing, selling, purchasing, delivering, ordering, receiving, storing
- A business process consists of activities, resources, facilities and information
- **Activities** transform resources and information of one type into resources and information of another type (activities may be manual, automated, controlled by computers, or some combination of both)
- **Resources** are items of value (resources may be people as well, such as a supplier or purchaser)
- **Facilities** are structures used within the business process - resources may be stored in facilities
- **Information** is used to determine how to transform the inputs they receive into the outputs they produce
- Object Management Group (OMG), a software industry standards organization, created the **Business Process Modeling Notation (BPMN)** (a standard set of terms and graphical notations for documenting business processes)
- Information is derived from data, which represents recorded facts or figures (information is data presented in a meaningful context)
- **Characteristics of good information: accuracy** (good information is accurate), **timeliness** (good information must be produced in time for its intended use – it is late if it is past 6 weeks or the event), **relevant** (information should be relevant to the context and the subject), **sufficient and succinct** (information must be sufficient only to the extent that it can be easily read and identified by the user) and **worth its cost**

- **Business process management (BPM)** is work that involves determining the most efficient and profitable way to do business
- To support improvements in business processes, many business uses the following methods: **total quality management (TQM), six sigma** and **lean production**
- **Information system: Hardware – Software – Data – Procedures -- People**
- **Automated systems** are systems where computers do the work by following instructions in the software (it eliminates the work people must do) – automation works from the human side to the computer side
- **Manual systems** are operated by people (the implementation of a manual system may be due to complicated payment processes and/or exceptions)
- Decisions occur at three levels in organizations: operational, managerial and strategic
- Information system to support payment:
    - ➢ For a business (automated): hardware (cash register computer, database host computer), software (sales-recording program on cash register), data (sales data, inventory database), procedures (operate cash register) and people (cashier)
    - ➢ For an individual customer (manual): hardware (personal computer), software (Adobe Acrobat Reader, email), data (quantity received, shipping invoice), procedures (reconcile receipt document with invoice, issue payment authorization if needed, process exceptions) and people (accounts payable)
- Information system to support purchasing (balance between computer and human work): hardware (personal computer, database host computer), software (inventory application program, purchasing program), data (inventory database), procedures (issue Purchase Order according to inventory management practices and guidelines) and people (purchasing clerk)
- **Operational** decisions concern daily activities – information systems that support operational decision making are called **transaction processing systems (TPS)**
- **Managerial** decisions concern the allocation and utilization of resources – information systems that support managerial decision making are called **management information systems (MIS)**
- **Strategic** decisions concern broader organizational issues such as product development – information systems that support strategic decision making are called **executive information systems (EIS)**
- A **structured decision** is one for which there is an understood and accepted method for making the decision
- An **unstructured decision** is one for which there is no agreed-upon decision-making method
- The following are typically the steps in the decision-making process: intelligence gathering, alternatives formulation, choice, implementation and review
- **Intelligence gathering** involves determining what is to be decided, what the criteria for the decision will be and what data are available

- **Alternatives formulation** is the stage in which decision makers lay out various alternatives
- A **choice** must be selected, and then it must be **implemented**
- After the steps have taken place, an ongoing **review** must be conducted to monitor the effectiveness of the choice
- Microsoft Visio is used to create process models using the **data flow** diagramming technique
- **Business process modeling** is the creation of a diagram and/or text that provides an overview of the flow of events that occur in an organization as a result of stimulus (it is meant to take into consideration all aspects of the business)
- **Conceptual modeling** is a broader area of modeling that is used to understand how complicated issues can be simplified to help people identify, implement, use and maintain more valuable information systems
- **Process awareness** refers to the amount of knowledge the people have about the process in which they are involved
- **Modeling tools** (also called **modeling applications**) include: Microsoft Visio, IBM Rational Rose, Borland Together, iGrafx FlowCharter and freeware applications (such as IDS Scheer ARIS Express)
- Most common business **process modeling techniques**: **data flow diagrams (DFDs)**, system flowcharts, workflow diagrams, use cases from the Unified Modeling Language (UML), and business process diagrams from the Business Process Modeling Notation (BPMN)
- Microsoft Visio has the following **DFD shapes: process** (symbolizes a set of related activities), **interface** (also labeled **external entity**, symbolizes a person or thing that exists outside the process), **data store** (symbolizes a place where data are either stored or retrieved) and **connector** (represents the data that flow between processes, entities, and data stores)
- **Rules for combining shapes:**
  - ➢ Rule 1: Every process should have at least one data flow connector coming into it and at least one data flow connector out of it.
  - ➢ Rule 2: An external entity cannot be directly connected to a data store.
  - ➢ Rule 3: A data store cannot be directly connected to another data store. A data store must be connected to a process
- The DFD model uses the concept of **leveling** to slowly reveal increasing amounts of information – the least amount of detail is called a **level-0 diagram** or **top process** (shows a large single process)
- **See page 51 for instructions on how to draw a level-0 DFD**
- **See page 52 for instructions on how to create a level-1 DFD (a more detailed process description)**

Chapter 3

- **Productivity** is measured by dividing the Canadian GDP by the number total paid hours worked by Canadians
- Stephen Roach and Robert Solow argued that the widespread increase in computers does not positively contribute to productivity
- **Productivity paradox** is the question of how IT creates **business value** – this paradox exists in part because it is difficult to measure the productivity of an IT investment
- There are 3 ways to determine value derived from an IT investment:
  - ➢ Productivity (companies can produce better quality products faster)
  - ➢ Competition (technology may be necessary to compete with firms in the industry)
  - ➢ Benefits to the end consumer (processes are made easier and more cost efficient, and products and services are higher quality)
- **Canadian Coalition for Tomorrow's ICT Skills (CCICT)** was founded by Bell in 2007 and is an industry-led group of Canadian employers, educational institutions and industries that ensure the ability of Canadian organizations to hire ICT professionals in the 21st century
- The CCICT developed the **Business Technology Management (BTM)** program to train and certify people in the sector – one aspect of this is the **Skills Framework for the Information Age (SFIA)**
- **Efficiency** is when business processes can be done more quickly and/or with fewer resources
- **Effectiveness** involves delivering greater consumer value through new and improved products and services
- **Value chain** is a network of activities that improve the effectiveness of a good or service (first formalized by Michael Porter, a Harvard professor)
- **Margin** is the difference between the price the customer is willing to pay and the cost of producing the good
- When a company undertakes activities that affect raw materials (or the first phase of production), it is moving upstream in the value chain – when a company undertakes activities that are closer to the end consumer, it is moving downstream
- **Primary activities** are activities in which value is directly added to the product [inbound logistics (receiving, storing, and disseminating inputs to the product), operations (transforming inputs into the final product), outbound logistics (collecting, storing, and physically distributing the product to buyers), marketing and sales and service (assisting customer's use of the product)]
- **Support activities** are to support the primary activities – adds value indirectly (e.g. firm infrastructure, human resources, technology development and procurement)
- An industry structure can be assesses using Porter's **five forces model** – states that five forces affect industry profitability:
  1. Bargaining power of the customers
  2. Threat of substitutions

3. Bargaining power of the suppliers
4. Threat of new entrants
5. Rivalry amongst existing firms

- After assessing the five forces, companies choose a competitive strategy (**cost, differentiation, focus** or **industry-wide**)
- **Sustaining technologies** are changes in technology that maintain the rate of improvement in customer value
- **Disruptive technologies** introduce a very new package of attributes to the accepted mainstream products
- **Diffusion of innovation** is the process of successfully releasing a new technology into society – this process has five steps/stages:
    1. Knowledge
    2. Persuasion
    3. Decision
    4. Implementation
    5. Confirmation
- **Switching costs** is when organizations retain customers by making it difficult or expensive for them to switch to another product
- **Sustained competitive advantage** is a distinctive way to compete that is effective over a long period of time
- A **spreadsheet** (also known as a **workbook**, which is a collection of one or more **worksheets**) is an electronic model built in Excel
- Spreadsheets store information in **cells** (the intersection of a column and a row)
- Spreadsheets allow you to enter the following information: letters, numbers and formulas (which always start with an "=" sign)
- To reference a cell, type in the equals sign and then the address of the cell
- **Relative cell addressing** keeps the relative position of the formula and the cells it references intact
- **Absolute cell addressing** consists of the column letter and row number surrounded by dollar signs – another method used is creating a **range name** (under the "Define Name" option of the cell, and can be used to reference the cell)
- To show data on a graph, go to design and select a type

Chapter 6

- **Collaboration** occurs when two or more people work together to achieve a common goal, result or product
- The effectiveness of a collaborative effort is driven by four critical factors: **communication skills and culture**, **communication systems**, **content management** and **workflow** (the process whereby content is created, edited, used and disposed of – used internally for organizations, as opposed to a business process) **control**
- **Network externality** is the effect of increased value as a network gets larger

- **Natural monopoly** occurs when one network can support all users and it is hard to switch for the consumer
- **Global positioning system (GPS)** uses a collection of dozens of satellites that orbit Earth and transmit precise micro-wave signals
- A **computer network** is a collection of computers that transmit and/or receive electronic signals through transmission media
- **Transmission media** might be physical or wireless media
- **Local area networks (LANs)** connect computers within a single physical site
- **Wide area networks (WANs)** connect computers at different geographical locations
- **Internet** is a network of networks
- **Protocol** is a set of rules that two communicating devices follow
- Elements of a LAN:
    - ➢ **Switch** is a special purpose computer that receives and transmits messages on the LAN
    - ➢ **Network interface card (NIC)** connects the device's circuitry to the network cable – **onboard NIC** is built into the computer
    - ➢ Each NIC has a unique identifier called a **media access control (MAC) address**
    - ➢ **Unshielded twisted pair (UTP) cable** is used to make connections – it consists of four pairs of twisted wires and uses an RJ-45 connector to connect to NIC devices on the LAN
    - ➢ **Optical fiber cables** may be used to replace UTP cables if they carry a lot of traffic or are far apart
    - ➢ **See page 197 for an outline of all of the elements of LANs and WANs**
- **Ethernet (IEEE 802.3 protocol)** is the world's most popular protocol for LANs – most personal computers are equipped with an onboard NIC that supports **10/100/1000 Ethernet**
- *k* stands for 1000 in communications equipment and 1024 in memory
- *M* stands for 1 000 000 in communications equipment and 1024 1024 in memory
- *G* stands for 1 000 000 000 in communications equipment and 1024 1024 1024 in memory
- 100 Mbps is 100 000 000 bits per second
- Communications speeds are measured in bits and memory sizes are expressed in bytes
- Wireless LANs require one or more **access points (APs)**
- **Wireless NIC (WNIC)** allows users to move around and stay connected without needing a plug in with a cable
- **Repeaters** and **reflectors** are sometimes used to amplify and reflect signals to extend the range
- **802.11 protocol** is Wi-Fi and is faster, with a speed of 248 Mbps
- **802.3 protocol** is another standard that is used

- **m-commerce** refers to mobile commerce, which speaks to users' ability to make transactions using their smartphones
- The Internet is a WAN
- **Routers** are special purpose computers that implement the protocol for WANs – routers are operated by your **Internet service provider (ISP)**
- The web is a subset of the internet, and consists of sites and users that process the **hypertext transfer protocol (HTTP) –** programs that implement the HTTP are called **browsers**
- **SMTP** is used for email and **FTP** is used for file transfer
- **Top-level domain (TLD)** are the last letters of an address (such as .ca, or .com)
- **Uniform resource locator (URL)** is an address on the Internet that is stated in a way humans can remember
- **IP address** is a logical address that consists only of numbers and periods – there are public and private IP addresses
- **Domain name system (DNS)** is used to convert URL into an IP address – this conversion is called domain name resolution
- Data in a computer must be converted to an **analog** before being sent, and then converted to a digital signal using a **modem**
- **Digital subscriber line (DSL) modems** operates on the same lines as voice telephones, but does not interfere with phone signals (DSL data transmissions and phone conversations can occur simultaneously)
- **Asymmetric digital subscriber lines (ADSLs)** are DSL lines that have different upload and download speeds
- **Symmetric digital subscriber lines (SDSLs)** are DSL lines that have the same speed in both directions
- **Cable modems** provide high speed data transmission using television lines
- **Narrowband** has communication speeds of less than 56 kbps
- **Broadband** has communications speeds of greater than 256 kbps
- **Dial-up modems** provide narrowband, while DSL and cable modems provide broadband
- **See page 2013 for a summary for the lines and speeds used to connect various networks**
- **Wireless WANs (WWNANs)** use cellular networks
- **Transmission control program/Internet protocol (TCP/IP)** has four layers: **network access layer** (transmission within a single network – this layer describes the equipment that is used for communications, the signaling used, and the protocols that will be used to communicate between machines), **Internet layer** (transmission across an internet – this layer works with IP addresses, and the standards control packet organization timing and constraints, as well as navigating packets from one IP address to another), **transport layer** (transmission across an internet – deals with opening connections and maintaining them) and **application layer** (programs for email, web browsing and file transfer – data is passed between programs)

- **Transmission control program (TCP)** is used to break apart large messages in **segments**, which are then packaged by the IP layer program into **packets** (the router sends the packet to another router, and then confirms whether it is the right destination)
- **Firewall** is a computing device that prevents unauthorized network access
- **Port** is a number that is used to uniquely identify a transaction over a network
- **Access control list (ACL)** keeps track of which IP addresses are allowed and which are prohibited
- **Packet-filtering firewalls** examine each part of the message (source address, destination address and other data) and determine whether or not to let it pass
- **Encryption** is the process of transforming clear texted into coded, unintelligible text for secure storage or communication
- **Asymmetric encryption** is where different keys are used to encode and decode the message – **symmetric encryption** is much simpler
- The flow of symmetric encryption works as follows:
    1. Your computer obtains the public key of the website to which it will connect
    2. Your computer generates a key for symmetric encryption
    3. Your computer encodes that key using the website's public key, and then sends the encrypted symmetric key to the website
    4. The website decodes the symmetric key using its private key
    5. Your computer and the website communicate using symmetric encryption
- **Virtual private network** uses the Internet or a private internet to create a connection
- **Tunnel** is a virtual, private pathway over a public or shared network from the VPN client to the VPN server
- **Web crawler** is a software program that browses the web – the list of URLs created by the crawler is referred to as the **crawler frontier**
- **Search engine indexing** is the process of organizing the information retrieved

Chapter 4

- **Mainframes** were commercial digital computers (hardware only) that could add or subtract 16,000 numbers per second – they were mainly used by businesses and governments and were room-sized
- **Microcomputers** are run by microprocessors – these incorporated a CPU and short–term memory, and became critical to the development of personal computers
- Important changing elements of IT: **price and performance advances, "small is powerful"** and "**the network is the thing"**
- **Hardware** consists only of the physical electronic components and related gadgetry that input, process, output and store data according to instructions encoded in computer programs or software

- Four basic components of computers: **storage, input, process** and **output**
- **Processing devices** include the **central processing unit (CPU)** that selects instructions, processes them, performs arithmetic and logical comparisons, and stores results of operations in memory
- CPU performance is measured in Hertz (Hz) [kilohertz (kHz), megahertz (mHz) and gigahertz (GHz)]
- A computer's **main memory** is referred to as **random access memory (RAM)**
- **Storage hardware** stores data and programs
- **Special function cards** can be added to augment the different components of a computer
- Computers represent data using **binary digits** (**bits**)
- **Kilobyte** = 1024 bytes
- **Megabyte** = 1024K = 1 048 576 bytes
- **Gigabyte** = 1024MB = 1 073 741 824 bytes
- **Terabyte** = 1024GB = 1 099 511 627 776 bytes
- **Petabyte** = 1024TB = 1 X 10^{15}
- **Bytes** consist of groups of eight bits
- **Data channel** or **bus** is used to move an instruction from the main memory to the CPU
- **Cache** is an element of the CPU that is used to store a very small amount of fast memory
- **Operating system (OS)** is a program that controls the computer's resources as well as a block of data
- **Memory swapping** is when the computer will replace unused memory with a new program
- The cache and main memory are **volatile**, meaning that their contents are lost when power goes off
- **Solid-state devices** (such as USB sticks and magnetic disks) are **non-volatile**
- **Client** computers are used for word processing, spreadsheets and database access
- **Servers** include Facebook, Google and Amazon and are used to provide a variety of services
- **Server farm** is a collection of computers that is used to coordinate all activities of a server
- **Cloud computing** is where customers do not own they computer they use, but get the services of hardware, software and applications through a web browser
- **Grid computing** uses software to divide and apportion pieces of a program among several computers to address a single problem
- **See figure 4.10 to see what a manager needs to know about software**
- **Instruction set** are the commands that a CPU can process
- Four major operating systems:
    1. **Windows** (usually for business users)
    2. **Mac OS** (developed by Apple Computer Inc.
    3. **Unix** (typically used for science and engineering)

4. **Linux** (developed by the **open-source community**)

- Software is purchased by buying a **license**
- **Application software** consists of programs that perform a business function
- **Horizontal market application** software provides capabilities common across all organizations and industries (off-the-shelf)
- **Vertical market application** software serves the needs of a specific industry (off-the-shelf and then customized)
- **One-of-a-kind application** software is developed for a specific, unique need (custom)
- **Firmware** is computer software that is installed into such devices as printers, print servers and various types of communications devices to be used as a medium for communication
- **Basic input/output system (BIOS)** is an important piece of firmware that is used to protect volatile memory when a computer is switched on and off
- **Thin clients** require nothing more than a browser to operate
- **Thick clients** requires programs other than a browser to operate (does not require a network)
- **See figure 4.14 on page 117 to see a business manager's role in software and hardware specification**
- **Virus** is a computer program that replicates itself
- **Payload** is the program code that causes unwanted activity
- **Macro-viruses** attach themselves to Word, Excel and other types of documents and appear in the startup files
- **Worm** is a virus that propagates using the Internet or other computer network
- **Zombies** are the computers that are affected by the worm
- **Botnet** is a set of computers and applications that are coordinated through a network and are used to perform malicious tasks
- **Patches** are used to cover up security holes
- **Antivirus programs** are used to protect computers and networks

Chapter 5

- **Content** is essentially property and is closely associated with **intellectual property** (a form of creative endeavor that can be protected through a trademark, patent, copyright, industrial design, or integrated circuit topography)
- Data management focuses on how to efficiently and effectively process bytes
- Web **content management systems (CMSs)** have been developed to help companies organize the process of presenting accurate web content
- **Opentext** is a leader in enterprise content management, having grown to over $1 billion per year (located in Waterloo, Ontario)
- Lists can be used to store data that involves a single theme
- Databases must be used to store data that involves multiple themes
- **Database** is a self-describing collection of integrated records

- **Bytes** (characters of data) are grouped into **columns** (also called **fields**) – these are typically headings such as "Student Name"
- Columns are in turn grouped into **rows** (also called **records**)
- **Table** or **file** is a group of similar rows
- **Metadata** is special data that describes the structure of the database (data that describes data)
- **Key** is a column or group of columns that identifies a unique row in a table
- **Foreign keys** are keys that are from a different table from the one in which they reside
- **Relational database** is a database that carries its data in the form of tables and that represents relationships using foreign keys
- Database = tables or files + relationships among rows in tables + metadata
- **Database management system (DBMS)** is a program used to create, process, and administer a database – some popular DBMS products include **Oracle** (Oracle Corporation), **SQL Server** (Microsoft), **Access** (Microsoft), **DB2** (IBM) and **MYSQL**
- User ⮀⮀ Database application (forms, reports, queries, application programs) ⮀⮀ DBMS (Database Management System) ⮀⮀ Database (tables, relationships, metadata)
- Applications use the DBMS for four operations: read, insert, modify or delete data
- **Structured Query Language (SQL)** is an international standard language used for processing a database
- DBMSs create the database and its structures, process the database and administer the database
- **Database application** is a collection of forms, reports, queries and application programs that process a database
- Data entry **forms** are used to read, insert, modify and delete data
- **Reports** are used to show data in a structured context
- **Query** is a search for a key word for which the user is looking
- **Multiuser processing** is where multiple users are processing the same database
- **Lost-update problem** is where incorrect data is processed due to multiuser processing
- **Enterprise DBMS** products process large organizational and workgroup databases
- **Personal DBMS** products are designed for smaller, simpler database applications – the only personal DBMS is Microsoft Access
- **Fields** are sections that store **attributes** of entries
- **Entity** is anything about which the organization wishes to store as data
- All fields containing data about one entry form a record
- **File** is a set of related records
- **Tables** consist of rows and columns and should have a primary key

- In an **entity-relationship (ER) diagram**, two lines indicates total participation, one line indicates partial participation, rectangles represent entities, diamonds represent relationships and ovals/circles represent attributes
- **Normalization** includes the *process* (converts one table into two or more tables and makes them well-structured), fixing *data integrity problems* (such as different names for the same entity and incorrect and inconsistent information, can be solved by eliminating duplicated data) and *normalized tables* (slower to process, eliminates duplicated data and every table has a single topic)

Chapter 7

- Key business functions in almost every business include accounting, finance, human resources, marketing & sales, operations and procurement
- **Functional systems** facilitate the work of a single department or business function (such as customer management for sales & marketing and accounts receivable for accounting)
- **Functional silos** are functional applications that are designed to work independently of one another – however, all functional systems are interrelated
- **Cross-developmental** (also known as **cross-functional**) **systems** integrate data and business processes across different departments and systems
- **Interorganizational systems** are functional systems that are used by two or more related companies (i.e. e-commerce applications)
- Major problems with isolated functional systems:
    - Data duplication
    - Data inconsistency
    - Disjointed processes
    - Limited information and lack of integrated information
    - Isolated decisions lead to inefficient overall activities
    - Increased expense
- **Business process design (modeling)** is the procedure whereby business analysts create a representation of the operational processes of a company in order to assess and improve those processes
- Two approaches to integration:
    1. **Enterprise application integration (EAI)** is an approach to combining functional systems, which uses layers of software as a bridge to connect different functional systems
    2. **Enterprise resource planning (ERP) systems** support many or all of the primary business processes as well as the human resources and accounting support processes – it is a central database that is combined with a set of standard business processes built on top of the database to ensure integration between functional areas ¥ this is very expensive
- Business process designs must be tailored to **industry standard processes**, as one organization alone cannot bore all of the costs of the design project

- **Process blueprint** is a comprehensive set of inherent processes for all organizational activities
- Benefits of an ERP include:
    - Efficient business processes
    - Inventory reduction
    - Lead-time reduction
    - Improved customer service
    - Greater, real-time insight into an organization
    - Higher profitability
- **Customer relationship management (CRM) systems** support the business processes of attracting, selling, managing, delivering, and supporting customers (support all direct value-chain activities)
- **Customer life cycle** includes 4 phases:
    i. Marketing
    ii. Customer acquisition
    iii. Relationship management
    iv. Loss/churn
- CRM has 3 components which include solicitation, lead-tracking (pre-sale) and relationship management (post-sale) [this includes sales management applications and customer support applications]
- **Sales management applications** support sales to existing customers by prioritizing customers according to their purchase history
- **Integrated CRM applications** store data in a single database
- **Supply chain management (SCM) systems** are interorganizational systems that enable companies to efficiently handle the flow of goods from suppliers to customers
- **Supply chain** is a network of organizations and facilities that transforms raw materials into products delivered to customers
- **Disintermediation** is the reduction in the use of intermediaries between producers and consumers
- Four drivers that affect supply chain performance: facilities, inventory, transportation and information
- Three factors of information: purpose availability and means
- Three fundamental information systems involved with **supply chain management:** supplier relationship management (SRM), inventory and customer relationship management (CRM)
- **Supplier relationship management (SRM)** is a business process for managing al contacts between an organization and its suppliers [illegible] have three basic processes: source, purchase and settle
- Benefits of information systems on supply chain performance include:
    - Reduce costs of buying and selling
    - Increase supply chain speed
    - Reduce size and cost of inventories

- Improve delivery scheduling – enable just-in-time inventory

Chapter 8

- **Information overload** is the issue that arises when managers have an overabundance of information to use when making decisions
- **Russell Ackoff** makes three assumptions as to how managers make decisions:
    1. Managers will have no problem making decisions if they have the data that they need
    2. Managers may make poor decisions either because of a lack of information or because of an abundance of information
    3. Managers know the data that they require
- **Kilobyte (KB)** = 1,000 bytes (100 KB = A low-resolution photograph)
- **Megabyte (MB)** = 1,000,000 bytes (500 MB = CD-ROM)
- **Gigabyte (GB)** = 1,000,000,000 bytes (100 GB = A library floor of academic journals)
- **Terabyte (TB)** = 1,000,000,000,000 bytes (400 TB = National Climactic Data Center (NOAA) database)
- **Petabyte (PB)** = 1,000,000,000,000,000 bytes (200 PB = All printed material)
- **Exabyte (EB)** = 1,000,000,000,000,000,000 bytes (5 EB = All words ever spoken by human beings)
- Experts suggest that data is growing at a rate of 30% each year
- Examples of issues with data:
    - **Dirty data** are problematic data (data that does not make sense)
    - **Missing values**
    - **Inconsistent data**
    - **Granularity** refers to the degree of summarization or detail (**course data** are highly summarized, **fine data** express details that are too precise) ☒ wrong granularity is too fine, or not fine enough, and/or too many attributes or too many data points
    - **Clickstream data** is fine data and includes everything on which the customer clicks (tracking their behavior)
- **Online transaction processing (OLTP)** is a system that is used to process transactions online, as opposed to paper (transactions can either be processed immediately or they can be grouped and then process all together)
- **Data resource challenge** is the challenge of using collected data effectively in order to improve decision making
- **Decision support systems (DSSs)** [or **Online Analytical Processing (OLAP)**] focus on making OLTP-collected data useful for decision-making
- **OLAP cube** is an array of data understood in terms of its 0 or more dimensions
- OLAP has the ability to sum, count, average and perform other simple arithmetic operations on groups of data
- "**Drill down**" means to further divide the data into more detail

- **Business intelligence (BI) system** is a system that provides information for improving decision making
- **Group decision support systems (GDSSs)** allow multiple parties to participate in decision making and improve outcomes by reducing often inherent biases
- Five categories of BI systems:
    1. **Group decision support systems (GDSSs)** allow multiple parties to participate in decision making and improve outcomes by reducing often inherent biases [illegible] improve decision outcomes by reducing many of the biases inherent in group discussion and option evaluation
    2. **Reporting systems** integrate and process data by summing, sorting, grouping and formatting, and produce, administer and deliver reports [illegible] improve decisions by providing relevant, accurate, and timely information to the right person
    3. **Data-mining systems** use sophisticated statistical techniques to find patterns and relationships [illegible] improve decisions by discovering patterns and relationships in data to predict future outcomes
    4. **Knowledge management systems** share knowledge of products, product uses and best practices among employees, managers, customers and others [illegible] improve decisions by publishing employee and others' knowledge and create value from existing intellectual capital, thereby fostering innovation, improving customer service, increasing organizational responsiveness and reducing costs
    5. **Expert systems** encode human knowledge in the form of If/Then rules and process those rules to make a diagnosis or recommendation [illegible] improve decision making by non-experts by encoding, saving, and processing expert knowledge
- **Market-basket analysis** is a data-mining system that computes correlations of items on past orders to determine items that are frequently purchased together
- **Tools** allow data to be processed into information
- **Systems** allow people to effectively utilize that processed data
- **RFM analysis** is a way of analyzing and ranking customers according to their purchasing patterns (sorts customers based on their most recent purchases, the frequency of their purchases and the money that they spend on each purchase)
- **Data warehouses** are used to extract and clean data from operational systems and other sources and to store and catalogue that data for processing by BI tools
- **Metadata** are data about data
- **Data mart** is a data collection that is created to address the needs of a particular business function, problem or opportunity
- **Think of the data warehouse as a distributor and a data mart as a retailer in a supply chain
- **Data mining** is the application of statistical techniques to find patterns and relationships among data and to make classifications and predictions

- **Unsupervised data mining** is when analysts do not create a model or hypothesis before running the analysis – instead they apply the data mining technique to the data and observe the results one technique for unsupervised data mining is **clutter analysis** (statistical techniques identify groups of entities that have similar characteristics
- **Supervised data mining** is when data miners develop a model prior to the analysis and apply statistical techniques to data to estimate the parameters of the model
- **Regression analysis** is analysis that measures the impact of a set of variables on another variable
- **Neutral networks** are a data mining technique used to predict values and make classifications
- **Lift** shows how much the base probability increases or decreases when other products are purchased

Chapter 9

- **E-commerce** is the buying and selling of goods and services over public and private computer networks
- **Merchant companies** are companies that buy and resell their goods
- **Nonmerchant companies** are those that arrange for the purchase and sale of goods without ever owning or taking title to those goods (i.e. auctions, exchanges or clearinghouses)
- Three types of merchant companies:
    1. **Business-to-consumer (B2C)** are e-commerce sales between a supplier and a retail customer – this has a **web storefront** where customers can manage their orders
    2. **Business-to-business (B2B)** are sales between companies
    3. **Business-to-government (B2G)** are sales between companies and governmental organizations
- **E-commerce auctions** match buyers and sellers by using an e-commerce version of a standard auction
- **Clearinghouses** provide goods and services at a stated price and arrange for the delivery of the goods, but they never take title
- **Electronic exchanges** are also examples of clearinghouses that match buyers and sellers (i.e. the Toronto Stock Exchange)
- **Disintegration** is the removal of intermediaries between parties
- **Intermediation** is when new parties insert themselves into the distribution chain
- **Price elasticity** measures how much demand rises or falls with changes in its price
- **Issues with e-commerce** include:
    - Channel conflict
    - Price conflict

    - Logistics expense
    - Customer service expense
    - Showrooming
    - Taxation
- **Showrooming** occurs when a customer learns about or tries a product or service in the high cost bricks-and-mortar retail store while completing the sales transaction at the low-cost Internet sales channel of another retailer
- Three types of capital:
    1. **Physical capital** ⚿ the investment of resources for future profit
    2. **Human capital** ⚿ the investment in human knowledge and skills for future profit
    3. **Social capital** ⚿ the investment in social relations with the expectation of returns in the marketplace – social capital adds value with information, influence, social credentials and personal reinforcement
- **Web 2.0** refers to the integration and interaction of products and services such as smartphones, user-created content, social networking and dynamic marketplaces ⚿ companies such as Google provide software as a free service as opposed to selling it
- **User-generated content** refers to the website content that is contributed by users
- **Crowdsourcing** is the practice of obtaining needed services, ideas, or content by soliciting contributions from a large number of people, especially from an online community
- **Mashup** is a webpage or application that uses content from more than one source to create a single new service displayed in a single graphical interface ⚿ control is necessary to prevent mashups in some industries such as banking (to ensure that credit card transactions are not mashed up with public web applications such as Maps)

Chapter 10

- **Five basic ways to acquire software:**
    - Buy it and use it as is
    - Buy it and customize it (most common)
    - Rent or lease it
    - Build it yourself
    - Outsource it
- **Project management body of knowledge (PMBOK)** is a guide published (1996) by the Project Management Institute that outlines the process of undertaking an IT project ⚿ suggests the five process groups in any project are:
    1. Initiating
    2. Planning
    3. Executing
    4. Controlling and monitoring

5. Closing

- **IT projects** are business projects that have a large IT component
- **Information technology project management (ITPM)** is the collection of techniques and methods that project managers use to plan, coordinate and complete IT projects
- **Nine project knowledge areas:**
    1. Integration management
    2. Scope management
    3. Time management
    4. Cost management
    5. Quality management
    6. Human resources management
    7. Communications management
    8. Risk management
    9. Procurement management
- **Project management professionals (PMPs)** are professionals with certified project management skills
- **IT operations** are the elements of a business that provide the delivery of service, maintenance and protection of the IT infrastructure
- **Production system** is a computer program typically used to provide some form of artificial intelligence (tasks that can be performed by computers by usually require humans), which consists primarily of a set of rules about behavior
- **Information technology infrastructure library (ITIL)** is a well-recognized collection of books that provide a framework of best-practice approaches to IT operations
- **IT project risk** is the inherent risk in undertaking a new IT project
- **Systems development life cycle (SDLC)** is the classic process used to acquire information systems ¥ five phase process:
    1. System definition
    2. Requirements analysis
    3. Component design
    4. Implementation
    5. System maintenance
- **System development** is done by those basic tasks
- **Agile methods** are more nimble methods of system development
- **System analysis** consists of the first two phases in SDLC
- **System definition phase** is required to define the goals and scope of the new information system ¥ in this, firms must also assess **cost feasibility, schedule feasibility, technical feasibility** and **organizational feasibility**
- **Requirements analysis phase** is used to form the project team and define the necessary requirements ¥ the work of analysis and design is done by **business analysts** and **systems analysts**

- ****Commercial on-the-shelf (COTS)** software never exactly fits an organization's requirements
- **Component design** consists of each of the five components of the information system being properly designed includes **normal processing, backup** and **failure recovery**
- **Implementation** occurs once all design is complete this has two main elements:
    - ➢ **Test plan:** includes **product quality assurance (PQA)** testing and **beta testing** (the process of allowing future system users to try out the new system on their own)
    - ➢ **System conversion** (installation): includes **pilot installation** (the system is installed in only part of the business), **phased installation** (the system is installed across the whole organization), **parallel installation** (the new system runs in parallel with the old one until the new system is tested and fully operational) and **plunge installation** (the new system is installed and the old system is shut down)
- **Maintenance phase** is used to track failures and discover any new enhancements
- **Patches** are fixes that can be applied to systems
- **Service packs** are bundle fixes for low-priority problems
- **Waterfall method** is a sequential design process in which progress is seen as flowing steadily downwards through the SDLC
- **Analysis paralysis** happens when analysts overanalyze a situation to the point where no action is taken
- **Outsourcing** is the process of hiring another organization to perform a service
- **Application service providers (ASPs)** are a particular form of outsourcing where an organization contracts with a vendor to "rent" applications from the company on a fee-for-service basis

Chapter 11

- **Chief Information Officer (CIO)** acts as the principal manager of the IT department
- **Chief Technology Officer (CTO)** heads the technology department
- IT department includes **technology, operations** and **development**
- **Project manager** is responsible for interacting with the client and moving the project successfully towards completion
- **Lead designer/analyst** is responsible for understanding client needs and developing the overall look and feel of the site and all design elements
- **Developer** is responsible for taking the design and creating the functioning site
- **Technical architect** is responsible for making decisions about technical issues related to the site
- **Competitive strategy** is a plan that aims to give a business a long-term advantage over its competitors

- **IT architecture** includes the complexities of choosing the specifications and design for an information system
- **Enterprise architects** manage and document the technological infrastructure of the firm
- **Zachman framework** is an IT architectural framework that divides systems into two dimensions:
    - **Six reasons for communication** (what – data, how – function, where – network, who – people, when – time, why – motivation)
    - **Stakeholder groups** (planner, owner, designer, builder, implementer, worker)
- **Alignment** is the process of matching organizational objectives with IT architecture
- **Information systems governance** is the development of consistent, cohesive management policies and verifiable internal processes for information technology and related services
- **Sarbanes-Oxley Act (SOX)** in the U.S. and the **Budget Measures Act** in Ontario are in place to force companies to comply with governance standards
- **Information systems audit** is an audit focused on information resources that are used to collect, store, process and retrieve information
- **Information Systems Audit and Control Association (ISACA)** is an organization that develops standards for IT audit and governance
- **Certified Information Systems Auditor (CISA)** is a professional certification for IT audit
- **Control Objectives for Information and Related Technology (COBIT)** is a framework of best practices designed for IT management
- **Information systems ethics** is an issue that is constantly addressed from a moral and legal standpoint
- **Green IT** (environmentally friendly IT, such as computers that automatically turned off when not in use) contributes to the **triple bottom line** for organizations
- **e-cycling** is recycling electronic computing devices
- **e-waste** is old technology that is no longer usable (obsolete) and must be disposed of

Chapter 12

- **Identity theft** occurs when there is a personal security threat
- **Personal Information Protection and Electronic Documents Act (PIPEDA)** is intended to balance an individual's right to privacy of his or her personal information, which organizations need to collect, use and share personal information for business purposes [illegible] governs how data are collected and used
- Three sources of **security threats**:
    1. Human error and mistakes
    2. Malicious human activity (includes **spam**)

3. Natural events and disasters

- **Unauthorized data disclosure** are data that are no supposed to be disclosed
- **Pretexting** is when someone pretends to be someone else to obtain information
- **Phishing** uses pretexting via email (**email spoofing**)
- **Spoofing** is another term for someone pretending to be someone else – **IP spoofing** occurs when an intruder uses another site's IP address as if it were that other site
- **Sniffing** is a technique used to intercept computer communications to gain access to private information
- **Drive-by sniffers** simply take computers with wireless connections through an area and search for unprotected wireless networks
- Typical security threats:
    - **Undisclosed data disclosure**
    - **Incorrect data modification**
    - **Faulty service**
    - **Denial of service**
    - **Loss of infrastructure**
- **Hacking** is when a person gains unauthorized access to a computer system
- **Denial of service (DOS)** is when servers or gateways are shut down
- Three components of a security system:
    1. Senior management involvement
    2. Safeguards → **technical safeguards** (such as encryption and firewalls), **data safeguards** (such as passwords and physical security) and **human safeguards** (includes hiring, training and education)
    3. Incident response
- Username is the **identification** and the password is the **authorization**
- **Smart card** is a card with a magnetic strip that is used for identification [such as a **personal identification number (PIN)**]
- **Biometric authentication** uses personal physical characteristics such as fingerprints
- **Malware** includes viruses, worms, spyware and adware
- **Spyware** is installed on a user's computer with the user's knowledge → this is used to spy on the actions and computer activity of the user
- **Adware** is installed without the user's permission and resides in the background to observe user behavior
- **Malware definitions** are patterns in malware code
- **Data safeguards** protect databases and other organizational data
- **Data administration** refers to an organization-wide function that is in charge of developing data policies and enforcing data standards
- **Database administration** refers to a function that pertains to a particular database

- **Key escrow** is an arrangement in which the keys need to decrypt encrypted data that are in escrow (an arrangement between parties) to allow a third party access to those keys
- **Hardening** a site means to take extraordinary measure to reduce a system's vulnerability
- Primary system procedures include **normal operation, backup** and **recovery**
- **Hot sites** are remote processing centres run by commercial disaster-recovery services (they provide all necessary equipment after a disaster on a monthly basis for a fee)
- **Cold sites** provide office space but the customers provide and install all necessary equipment to continue operations