Dalhousie University
Faculty of Computer Science

CSCI 2141 – Intro to Database Systems

Data and Information

Week 1

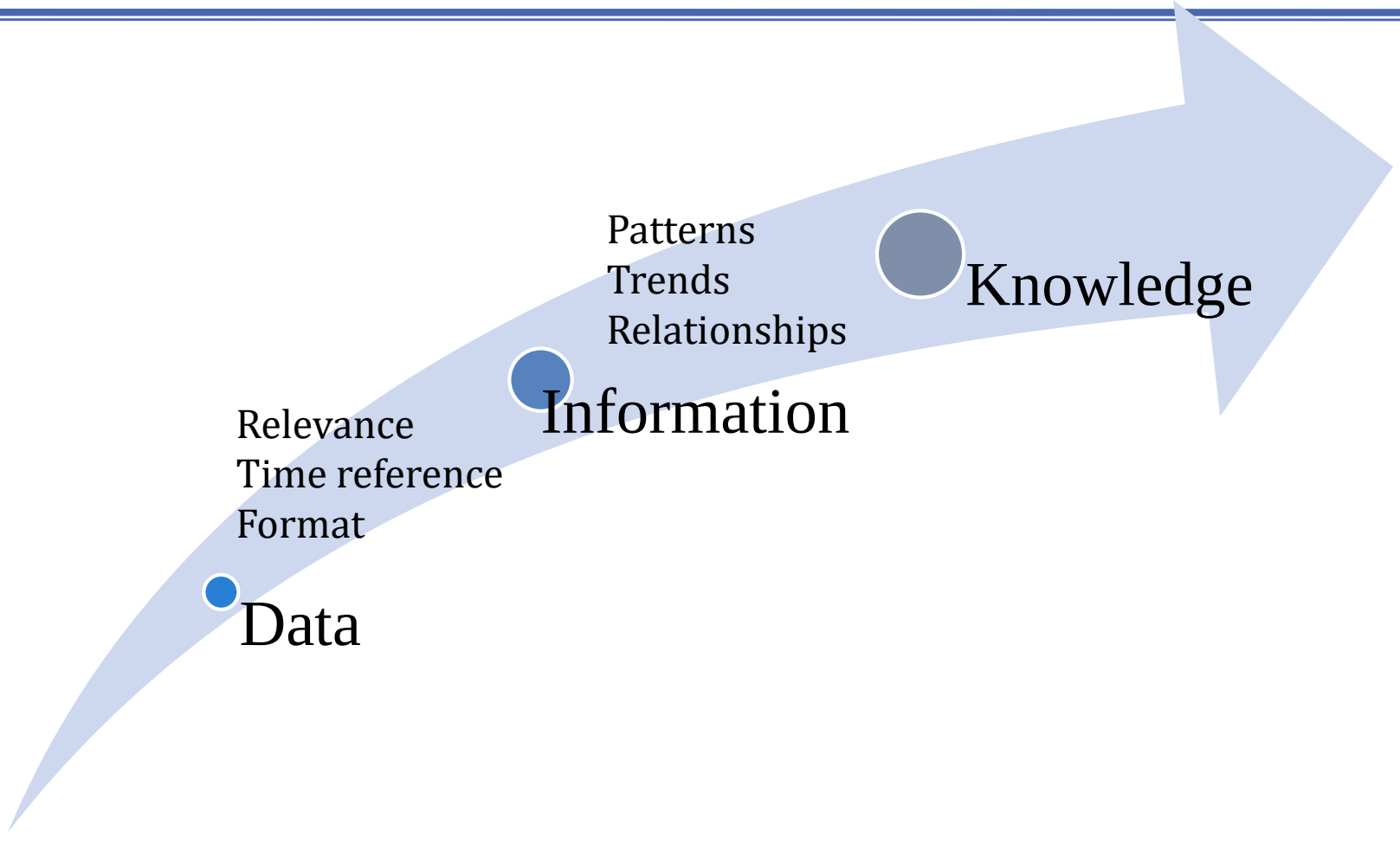
Data, Information and Knowledge

- **Data** is the raw representation of facts (as text, numbers, graphics, images, sound or video)

- **Information** is data put in context. Data that are processed to be useful; provides answers to "who", "what", "where", and "when" questions

- **Knowledge** is the application of data and information; answers “how” questions

Data, Information and Knowledge



Data → Information

a) Data entry screen

DO NOT append School ID (MT) to Member ID

Member ID * (<= 7 alphanumeric) (Password will be initially set to be the same as Member ID)

First name or initial *

Middle name/initial

Last name *

☐ Chair/Head ☐ Inactive

☐ Bypass chair for evaluation

Department *

Area

Email *

Hire Term *

Member Default Status: *Changing the status here changes only the default that is pulled into the uploaded teaching schedules. To change the historical status of members and to see your changes reflected in the various reports, edit the teaching schedules themselves.*

Involvement:* ○ Participating ○ Supporting

Qualification:*

☐ Participates in the governance of the school

☐ Considered to be a long-term member

High Degree*

Year Awarded*

Rank* Assistant Professor

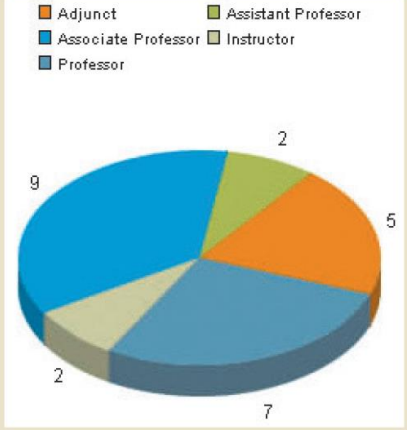
b) Raw data

Id	LastName	MidName	FirstName	DeptCode	Office	Email	Rank	HireYear	Degree
1	Washington	A.	George	MGMT	N135	gwashington@mtsu.edu	Professor	2001	Ph.D.
2	Adams		John	FIN	N313	jadams@mtsu.edu	Professor	1984	Ph.D.
3	Jefferson	L.	Thomas	ECON		tjefferson@mtsu.edu	Instructor	2002	M.B.A.
4	Madison	D.	James	FIN	N236	jmadison@mtsu.edu	Associate Professor	1994	Ph.D.
5	Monroe	N.	James	ACCT	N411	jmonroe@mtsu.edu	Assistant Professor	1995	Ph.D.
6	Adams	Q.	John	ACCT	N418	jqadams@mtsu.edu	Associate Professor	1989	Ph.D.
7	Jackson	C.	Andrew	ECON	N303	ajackson@mtsu.edu	Associate Professor	1999	Ph.D.
8	Van Buren	T.	Martin	FIN	N306	mvanburen@mtsu.edu	Professor	1988	Ph.D.
9	Harrison	R.	William	MKTG	N118	wharrison@mtsu.edu	Professor	1994	Ph.D.
10	Tyler	M.	John	MGMT		Jtyler@mtsu.edu	Assistant Professor	2000	Ed.D.
11	Polk		Cheryl	MKTG	N143	cpolk@mtsu.edu	Associate Professor	2002	Ph.D.
12	Taylor	G.	Zachary	ACCT	N415	ztaylor@mtsu.edu	Associate Professor	1996	Ph.D.
13	Fillmore		Millard	JCB	N219	mfillmore@mtsu.edu	Professor	1992	Ph.D.
14	Pierce	A.	Franklin	MKTG	N359	pfranklin@mtsu.edu	Instructor	2005	M.B.A.
15	Buchanan	T.	James	MGMT	N146	jbuchanan@mtsu.edu	Associate Professor	1996	D.B.A.
17	Lincoln	W.	Larry	MGMT	N150	llincoln@mtsu.edu	Associate Professor	1996	Ph.D.
18	Johnson		Andrew	ISYS	N360	ajohnson@mtsu.edu	Professor	1987	Ph.D.
19	Grant		Katie	MKTG	N120	kgrant@mtsu.edu	Assistant Professor	1989	D.B.A.
20	Rutherford		Hayes	ACCT	N408	hrutherford@mtsu.edu	Professor	1992	Ph.D.
21	Garfield	T.	Denise	ACCT		dgarfield@mtsu.edu	Assistant Professor	2018	Ph.D.
22	Arthur		Emily	ACCT	N413	earthur@mtsu.edu	Associate Professor	2003	J.D.
23	Cleveland	G.	Robert	ACCT	N401	rcleveland@mtsu.edu	Associate Professor	1997	Ph.D.
24	Harrison	X.	Patricia	BULA	N406	pharrison@mtsu.edu	Associate Professor	2001	J.D.
25	McKinley	B.	Priscilla	ISYS	N363	pmckinley@mtsu.edu	Adjunct	1994	M.S.
26	Roosevelt	F.	Hillary	MGMT	N104	hroosevelt@mtsu.edu	Associate Professor	2002	Ph.D.
27	Wilson		Laura	BCEN	N448	lwilson@mtsu.edu	Professor	1992	Ph.D.
28	Harding		Warren	MKTG	N114	wharding@mtsu.edu	Professor	1984	Ed.D.
29	Coolidge		Calvin	ECON	N316	ccoolidge@mtsu.edu	Professor	1975	Ph.D.
30	Hoover		Lisa	MGMT		lhoover@mtsu.edu	Adjunct	1978	M.B.A.
31	Truman		Betty	ACCT	N416	btruman@mtsu.edu	Professor	1971	Ed.D.
32	Johnson		Robert	BCEN	N240	rjohnsonr@mtsu.edu	Professor	2001	Ph.D.

c) Information in summary format

Rank	COUNT	%/INFS	TOT/COL	%/COL. TOT.	%/COL. FAC.
Adjunct	5	20.00%	23	21.74%	3.27%
Assistant Professor	2	8.00%	28	7.14%	1.31%
Associate Professor	9	36.00%	37	24.32%	5.88%
Instructor	2	8.00%	18	11.11%	1.31%
Professor	7	28.00%	47	14.89%	4.58%

d) Information in graphical format



Data – Key to Business Survival

- Data constitutes the building blocks of information
- Information is produced by processing data
- Information is used to reveal the meaning of data
- Accurate, relevant, and timely information is the key to good decision making
- Good decision making is the key to organizational survival in a global environment

FIGURE 1.1 THE PERVASIVE NATURE OF DATABASES

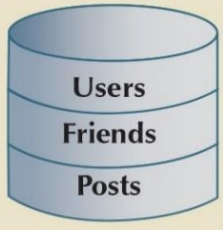
A Day In Susan's Life

See how many databases she interacts with each day

Before leaving for work, Susan checks her Facebook and Twitter accounts

Where is the data about the friends and groups stored?

Where are the "likes" stored and what would they be used for?



On her lunch break, she picks up her prescription at the pharmacy

Where is the pharmacy inventory data stored?

What data about each product will be in the inventory data?

What data is kept about each customer and where is it stored?

After work, Susan goes to the grocery store

Where is the product data stored?

Is the product quantity in stock updated at checkout?

Does she pay with a credit card?

At night, she plans for a trip and buys airline tickets and hotel reservations online

Where does the online travel website get the airline and hotel data from?

What customer data would be kept by the website?

Where would the customer data be stored?

Then she makes a few online purchases

Where are the product and stock data stored?

Where does the system get the data to generate product "recommendations" to the customer?

Where would credit card information be stored?

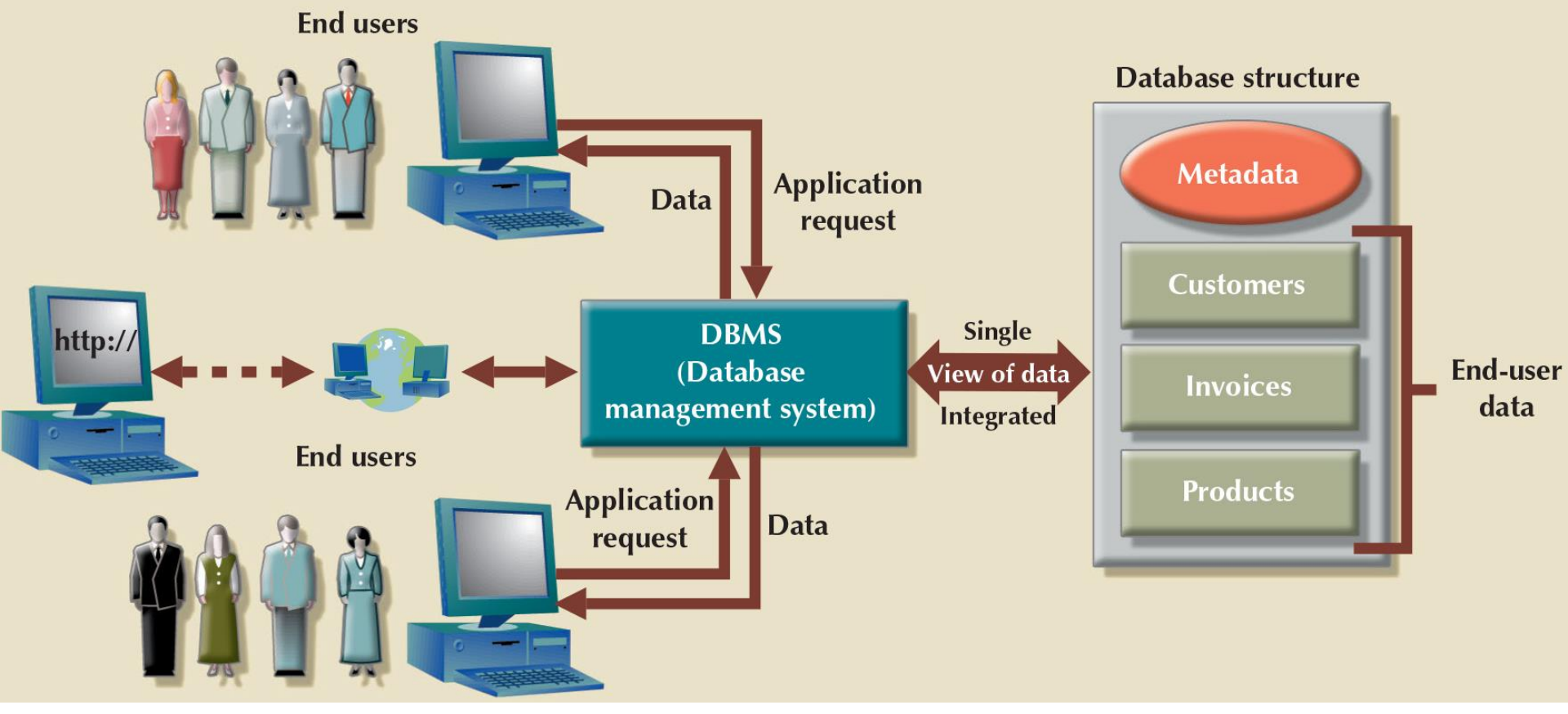
Introducing the Database

- Shared, integrated computer structure that stores a collection of:
  - End-user data - Raw facts of interest to end user
  - **Metadata**: Data about data, which the end-user data are integrated and managed
    - Describe data characteristics and relationships
- **Database management system (DBMS)**
  - Collection of programs
  - Manages the database structure
  - Controls access to data stored in the database

Role of the DBMS

- Intermediary between the user and the database
- Enables data to be shared
- Presents the end user with an integrated view of the data
- Receives and translates application requests into operations required to fulfill the requests
- Hides database's internal complexity from the application programs and users

The DBMS Manages the Interaction between the End User and the Database



Advantages of the DBMS

- Better data integration and less data inconsistency
  - **Data inconsistency**: Different versions of the same data appear in different places
- Increased end-user productivity
- Improved:
  - Data sharing
  - Data security
  - Data access
  - Decision making
- **Data quality**: Accuracy, validity, and timeliness of data

Types of Databases – User

- **Single-user database**: Supports one user at a time
  - **Desktop database**: Runs on PC
- **Multiuser database**: Supports multiple users at the same time
  - **Workgroup databases**: Supports a small number of users or a specific department
  - **Enterprise database**: Supports many users across many departments

Types of Databases – Location

- **Centralized database**: Data is located at a single site
- **Distributed database**: Data is distributed across different sites
- **Cloud database**: Created and maintained using cloud data services that provide defined performance measures for the database

Types of Databases – Purpose

- **General-purpose databases**: Contains a wide variety of data used in multiple disciplines
- **Discipline-specific databases**: Contains data focused on specific subject areas

Types of Databases – Time sensitivity and Use

- **Operational database**: Designed to support a company's day-to-day operations
  - also called online transaction processing (OLTP) database
- **Analytical database**: Stores historical data and business metrics used exclusively for tactical or strategic decision making
  - **Data warehouse**: Stores data in a format optimized for decision support
  - **Online analytical processing (OLAP):** Tools for retrieving, processing, and modeling data from the data warehouse
- **Business intelligence**: Captures and processes business data to generate information that support decision making

Types of Databases – Structure

- **Unstructured data**: It exists in their original state
- **Structured data**: It results from formatting
  - Structure is applied based on type of processing to be performed
- **Semi-structured data**: Processed to some extent
- **Extensible Markup Language (XML)**
  - Represents data elements in textual format

Database Design

- Focuses on the design of the database structure that will be used to store and manage end-user data
- Well-designed database
  - Facilitates data management
  - Generates accurate and valuable information
- Poorly designed database causes difficult-to-trace errors