

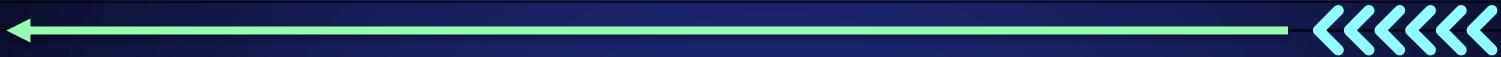
Intro to Databases



Introduction to Database

Lecture 1

Lecture outlines



After completing this chapter, you will be able to:

1. Define the difference between data and information
2. Describe what a database is, the various types of databases, and why they are valuable assets for decision making
3. Explain the importance of database design
4. See how modern databases evolved from file systems
5. Understand flaws in file system data management
6. Outline the main components of the database system
7. Describe the main functions of a database management system (DBMS)

01

Data versus Information

Data vs. Information

Data – Raw facts

- Must be formatted for storage and processing

Information – Process data to determine meaning

- Requires context!
- Can become knowledge

Data Management

- Generation, storage and retrieval of data to create information

Decision Making

- Accurate, relevant and timely information retrieval

FIGURE 1.2 TRANSFORMING RAW DATA INTO INFORMATION

a) Data entry screen

Middle Tennessee State University

You are viewing the College Contacts
Faculty/Monitors

[Home](#) | [Reports](#) | [Academic Agreements](#) | [Summaries](#) | [Maintenance](#) | [Wages & Accommodations](#) | [Wages](#) | [Ratings & Times](#) | [GCD](#) | [Calendar](#) | [Monitors](#) | [Logs](#) | [Settings](#) | [Issue](#) | [Sign Out](#)

[Batch](#) | [Save](#)

DO NOT append School ID (#) to Member ID

Member ID: (#) cannot be exactly zero (or the same as Member ID)

First name or initial:

Last name:

Rank:

Qualification:

Department:

Area:

Email:

Hire Term:

Chair/HOD Vice-Chair Typeless chair for evaluation

Participates in the governance of the school
 Considered to be a long-term member

High Degree:
Year awarded:
Rank:

b) Raw data

Id	FirstName	MidName	LastName	DeptCode	Office	Email	Rank	HireYear	Degress
1	Washington	A.	George	MGT	N135	gASHINGTON@mtsu.edu	Professor	2901-FbD	
2	Adams	J.	John	FIN	N132	jADAMS@mtsu.edu	Professor	1904-FbD	
3	Jefferson	L.	Thomas	ECON		tJEFFERSON@mtsu.edu	Instructor	2302-M.E.A.	
4	Madison	D.	James	FIN	N123	jMADISON@mtsu.edu	Associate Professor	1904-FbD	
5	Monroe	N.	James	ACCT	N401	jMONROE@mtsu.edu	Associate Professor	1805-FbD	
6	Adams	O.	John	ACCT	N403	jADAMS@mtsu.edu	Associate Professor	1809-FbD	
7	Jackson	C.	Andreas	ECON	N125	aJACKSON@mtsu.edu	Associate Professor	1908-FbD	
8	VanBuren	T.	Milton	FIN	N128	mVANBUREN@mtsu.edu	Professor	1808-FbD	
9	Harrison	R.	William	MKTG	N110	wHARRISON@mtsu.edu	Professor	1904-FbD	
10	Eaton	M.	John	MKT	N147	jEATON@mtsu.edu	Assistant Professor	2308-Ed.D	
11	McKinley	J.	Franklin	MKT	N148	fMCKINLEY@mtsu.edu	Professor	1896-FbD	
12	Taylor	G.	Zachary	ACCT	N415	zTAYLOR@mtsu.edu	Associate Professor	1919-Ed.D	
13	Patterson	M.	Miller-J	JOB	N129	mPATTERSON@mtsu.edu	Associate Professor	1906-FbD	
14	Pierce	A.	Frances	MKTG	N153	fPIERCE@mtsu.edu	Professor	1912-FbD	
15	Buckner	T.	James	MKT	N146	jBUCKNER@mtsu.edu	Instructor	2015-M.E.A.	
16	Lincoln	W.	Loyd	MKT	N150	lLINCOLN@mtsu.edu	Associate Professor	1896-FbD	
17	Johnson	A.	Andreas	SYG	N160	aJOHNSON@mtsu.edu	Professor	1807-FbD	
18	Grant	K.	Mitch	MKTG	N126	mGRANT@mtsu.edu	Assistant Professor	1919-C.B.A.	
19	Rutherford	J.	Hayes	ACCT	N408	hRUTHERFORD@mtsu.edu	Professor	1902-FbD	
20	Croftfield	T.	Charles	ACCT	N409	cCROFTFIELD@mtsu.edu	Assistant Professor	2316-FbD	
21	Astolfi	E.	Emily	ACCT	N413	eASTOLFI@mtsu.edu	Associate Professor	2303-JD	
22	Chenverfield	G.	Pelton	ACCT	N401	pCHENVERFIELD@mtsu.edu	Associate Professor	1897-FbD	
23	Harrison	X.	Patricia	BUSA	N404	pHARRISON@mtsu.edu	Associate Professor	2001-JD	
24	McCloskey	B.	Freddie	SYG	N161	fMCCLOSKEY@mtsu.edu	Adjunct	1904-M.D.	
25	Patterson	F.	Friday	MKT	N164	fPATTERSON@mtsu.edu	Associate Professor	2302-FbD	
26	Wallace	E.	Ernest	ROEN	N446	eWALLACE@mtsu.edu	Professor	1912-Ed.C.T.	
27	Harding		Wesley	MKTG	N114	wHARDING@mtsu.edu	Professor	1904-FbD	
28	Coolidge	S.	Samuel	ECON	N176	sCOOLIDGE@mtsu.edu	Professor	1895-FbD	
29	Hood		Lee	MKT	N165	lHOOD@mtsu.edu	Adjunct	1826-M.E.A.	
30	Thurman	B.	Betty	ACCT	N405	bTHURMAN@mtsu.edu	Professor	1801-Ed.D	
31	Jeffreese		Robert	ROEN	N140	rJEFFREES@mtsu.edu	Professor	2301-FbD	

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

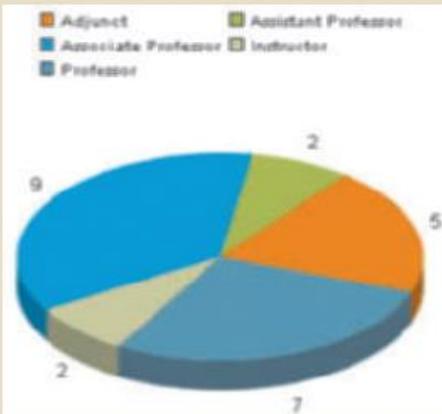


FIGURE 1.3 SMARTPHONE TRACKING

Smartphone Tracking is provided by GPS Location Services. GPS Location Services are enabled by default in most smartphones. When you install and run an application the first time, it will ask you if you want to enable location services for such application.

How does the phone know where "home" is? The phone knows where you spent most of your night hours and assumes that is your "home" location.

In the same way, it knows where you spent most of your work-day hours and assumes that is your "work" location.

How does the phone know where you last "parked" your car? If you use Bluetooth to pair your phone to your car audio system then, the phone knows where and when you are in your car. The phone remembers the last location you paired your phone to your car entertainment system.

If you enable "Frequent Locations" in your phone, it will track the places that you visit most frequently and keep an history of them.

Disclaimer: According to Apple, the GPS data is collected "in order to learn places that are significant to the individual". Data are only stored locally in the phone and never shared.

Frequent Locations: Murfreesboro

- Home 65 visits recorded since December 31, 2016
- Work 48 visits recorded since January 3, 2017
- Murfreesboro 11 visits recorded since January 4, 2017
- Park Ave 5 visits recorded since January 4, 2017
- Warrior Dr 4 visits recorded since February 11, 2017
- Memorial Blvd 3 visits recorded since January 21, 2017

02

»»» Introducing the Database



>>>> Database

A **database**: is a shared, integrated computer structure that stores a collection of the following:

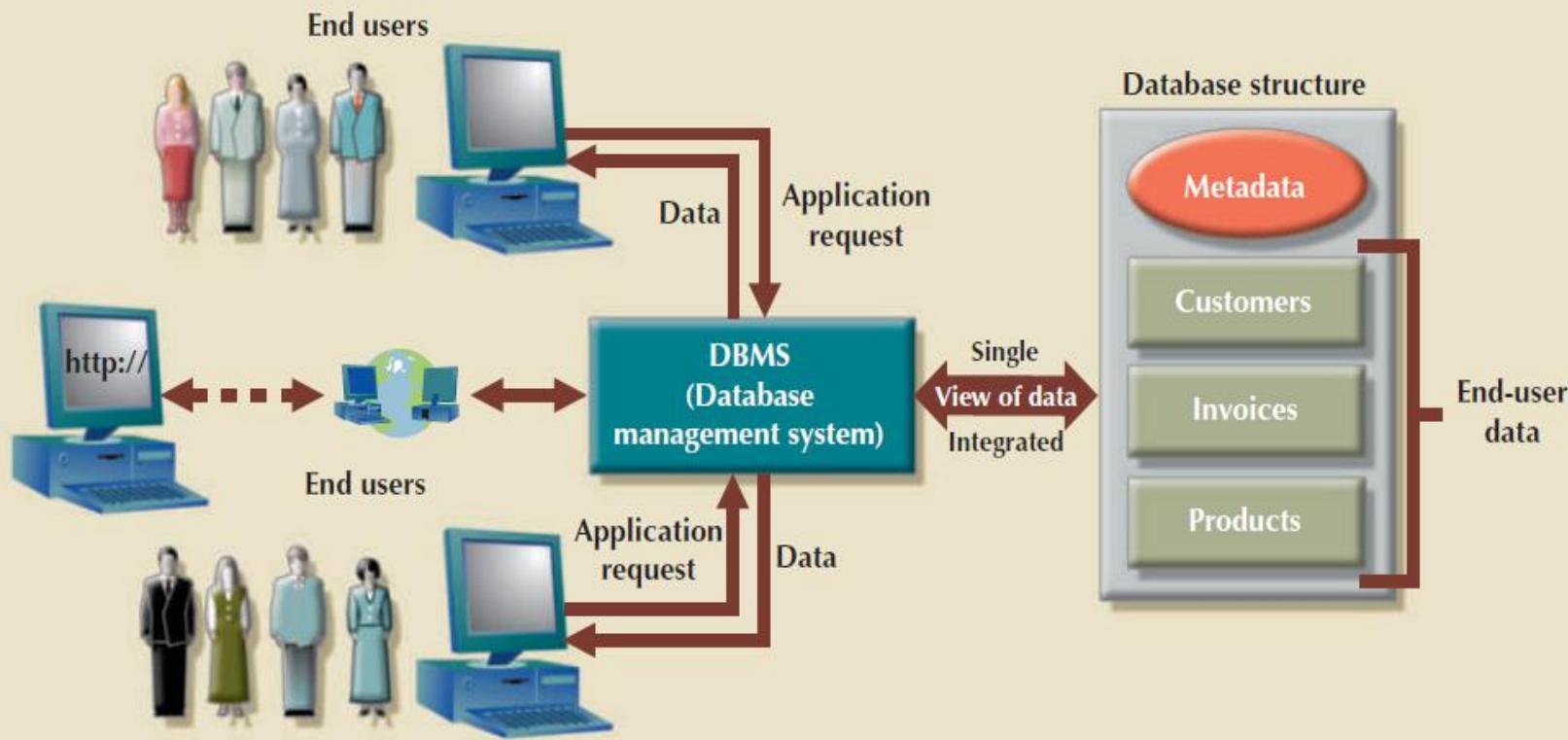
- End-user data—that is, raw facts of interest to the end user
- Metadata, or data about data, through which the end-user data is integrated and Managed
- The metadata describes the data characteristics and the set of relationships that links the data found within the database

Database management system (DBMS)



A DBMS: is a collection of programs that manages the database structure and controls access to the data stored in the database.

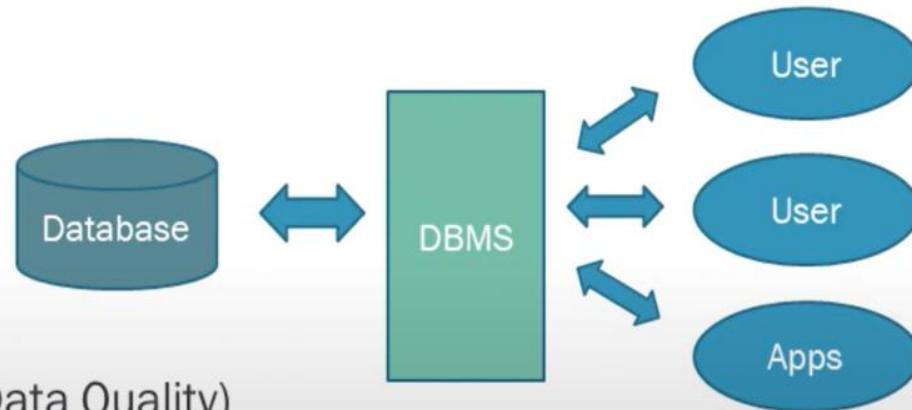
FIGURE 1.4 THE DBMS MANAGES THE INTERACTION BETWEEN THE END USER AND THE DATABASE



Database Management System (DBMS)

DBMS – An interface between the user and the data

- Data Sharing
- Data Security
- Data Integration
- Minimize Inconsistency
- Data Access (Query)
- Improved Decision Making (Data Quality)
- Increased Productivity



Types of Databases

Multi-user types:

Workgroup database - < 50 users
Enterprise database - > 50 users

- Single User (Desktop) vs. Multi-User
- Centralized vs. Distributed (multi-location)
- Cloud database (MS Azure, Amazon AWS, etc)
- General Purpose vs. Discipline-Specific
- Operational vs. Analytical
- Structured vs. Unstructured (vs Semi-structured)
- XML (Extensible Markup Language)
- NO-SQL (Not Only SQL)

Cloud DB:

Microsoft Azure
Amazon AWS
Oracle Cloud Database

Operational DB:

- Online Transaction Processing (OLTP)
- Transactional
- Production

03

Why Database Design Is Important?

Database design



- Personal productivity tools such as spreadsheets and desktop database programs suffer from lack proper data-modelling and database design skills
- **Database design:** refers to the activities that focus on the design of the database structure that will be used to store and manage end-user data.

Poor Database design (Example)



- Consider the data shown in the following figure.
- What is the problems you notes in this design.

Poor Database design (Example)

FIGURE 1.5 EMPLOYEE SKILLS CERTIFICATION IN A POOR DESIGN

Why are there
blanks in rows
9 and 10?

How to produce
an alphabetical
listing of
employees?

How to count how
many employees are
certified in Basic
Database Manipulation?

Is Basic Database
Manipulation the
same as Basic DB
Manipulation?

What if an employee
acquires a fourth
certification?

Do we add
another column?

ID	Enum	Name	Title	HireDate	Skill	Skill1Date	Skill2	Skill2Date	Skill3	Skill3Date
1 02345	Brian Oates	DBA		2/14/1997	Basic Database Management	2/14/2004	Advanced Database Management	2/14/2007	Basic Web Design	8/9/2005
2 08273	Marco Biercz	Analyst		7/28/2008	Basic Web Design	3/8/2011	Advance Process Modeling	8/19/2014		
3 06234	Jasmine Patel	Programmer		8/10/2007	Basic Web Design	8/10/2009	Advanced C# programming	8/10/2009	Basic DB manipulation	1/29/2014
4 03373	Franklin Johnson, Jr.	Purchasing Agent		3/15/2004	Advanced Spreadsheets	6/20/2013				
5 13567	Almond, Robert	Analyst		9/30/2014	Basic Process Modeling	9/30/2016	Basic Database Design	5/23/2017		
6 10282	Richardson, Amanda	Clerk		4/11/2013						
7 09382	Susan Mathis	Database Programmer		8/2/2012	Basic DB Design	8/2/2014	Basic Database Manipulation	8/2/2014	Advanced DB Manipulation	5/1/2015
8 14311	Duong, Lee	Programmer		9/1/2016	Basic Web Design	9/1/2018				
9					Master Database Programming					
10					Basic Spreadsheets					
11 09002	Wade Gaither	Clerk		5/20/2012	Advanced Spreadsheets	5/18/2015	Basic Web Design	5/18/2015		
12 13383	Raymond F. Matthews	Programmer		3/12/2014	Basic C# Programming	3/12/2016				
13 09263	Chavez, Juan	Clerk		7/4/2012						
14 04883	Patricia Richards	DBA		6/11/2006	Advanced Database Management	6/11/2008	Advanced Database Manipulation	9/20/2014		
15 13932	Lee, Meagan	Programmer		9/29/2015						

Poor Database design (Example)



1. With the improved structure in Figure 1.6, you can use simple commands in a standard data manipulation language to do the following:
2. With the improved structure in Figure 1.6, you can use simple commands in a standard data manipulation language to do the following:

- Produce an alphabetical listing of employees by last name:
 - `SELECT * FROM EMPLOYEE ORDER BY EMPLOYEE_LNAME;`
- Determine how many employees are certified in Basic Database

Manipulation:

- `SELECT Count (*)`
- `FROM SKILL JOIN CERTIFIED ON SKILL.SKILL_ID = CERTIFIED.SKILL_ID`
- `WHERE SKILL_NAME = 'Basic Database Manipulation';`

FIGURE 1.6 EMPLOYEE SKILLS CERTIFICATION IN A GOOD DESIGN

Table name: EMPLOYEE

Employee_ID	Employee_FName	Employee_Lname	Employee_HireDate	Employee_Title
02345	Brian	Oates	2/14/1999	DBA
03373	Franklin	Johnson	3/15/2006	Purchasing Agent
04883	Patricia	Richards	6/11/2008	DBA
06234	Jasmine	Patel	8/10/2009	Programmer
08273	Marco	Blenz	7/28/2010	Analyst
09002	Wade	Gaither	5/20/2014	Clerk
09283	Juan	Chavez	7/4/2014	Clerk
09362	Susan	Mathis	8/2/2014	Database Programmer
10282	Amanda	Richardson	4/11/2015	Clerk
13383	Raymond	Matthews	3/12/2016	Programmer
13667	Robert	Almond	9/30/2016	Analyst
13932	Megan	Lee	9/29/2017	Programmer
14311	Lee	Duong	9/1/2018	Programmer

Table name: SKILL

Skill_ID	Skill_Name	Skill_Description
100	Basic Database Management	Create and manage database user accounts.
110	Basic Web Design	Create and maintain HTML and CSS documents.
120	Advanced Spreadsheets	Use of advanced functions, user-defined functions, and macroing.
130	Basic Process Modeling	Create core business process models using standard libraries.
140	Basic Database Design	Create simple data models.
150	Master Database Programming	Create integrated trigger and procedure packages for a distributed environment.
160	Basic Spreadsheets	Create single tab worksheets with basic formulas.
170	Basic C# Programming	Create single-tier data aware modules.
180	Advanced Database Management	Manage Database Server Clusters.
190	Advance Process Modeling	Evaluate and Redesign cross-functional internal and external business processes.
200	Advanced C# Programming	Create multi-tier applications using multi-threading.
210	Basic Database Manipulation	Create simple data retrieval and manipulation statements in SQL.
220	Advanced Database Manipulation	Use of advanced data manipulation methods for multi-table inserts, set operations, and correlated subqueries.

Database name: Ch01_Text

Table name: CERTIFIED

Employee_ID	Skill_ID	Certified_Date
02345	100	2/14/2004
02345	110	8/9/2005
02345	180	2/14/2007
03373	120	6/20/2013
04883	180	6/11/2008
04883	220	9/20/2014
06234	110	8/10/2009
06234	200	8/10/2009
06234	210	1/29/2014
08273	110	3/8/2011
08273	190	8/19/2014
09002	110	5/16/2015
09002	120	5/16/2015
09362	140	8/2/2014
09362	210	8/2/2014
09362	220	5/1/2015
13383	170	3/12/2016
13667	130	9/30/2016
13667	140	5/23/2017
14311	110	9/1/2018

04

Evolution of File System Data Processing

Evolution of File System Data Processing



- Pencil and Paper
- File folders and cabinets
- If data was small, worked fine
- Moved to computers
- Required Data Processing (DP) Specialist
- Bad design caused issues

Example of simple Computerized File

FIGURE 1.7 CONTENTS OF THE CUSTOMER FILE

Database name: Ch01_Text

C_NAME	C_PHONE	C_ADDRESS	C_ZIP	A_NAME	A_PHONE	TP	AMT	REN
Alfred A. Ramas	615-844-2573	218 Fork Rd., Babs, TN	36123	Leah F. Hahn	615-882-1244	T1	100.00	05-Apr-2018
Leona K. Dunne	713-894-1238	Box 12A, Fox, KY	25246	Alex B. Alby	713-228-1249	T1	250.00	16-Jun-2018
Kathy W. Smith	615-894-2285	125 Oak Ln, Babs, TN	36123	Leah F. Hahn	615-882-2144	S2	150.00	29-Jan-2019
Paul F. Olowksi	615-894-2180	217 Lee Ln., Babs, TN	36123	Leah F. Hahn	615-882-1244	S1	300.00	14-Oct-2018
Myron Orlando	615-222-1672	Box 111, New, TN	36155	Alex B. Alby	713-228-1249	T1	100.00	28-Dec-2018
Amy B. O'Brian	713-442-3381	387 Troll Dr., Fox, KY	25246	John T. Okon	615-123-5589	T2	850.00	22-Sep-2018
James G. Brown	615-297-1228	21 Tye Rd., Nash, TN	37118	Leah F. Hahn	615-882-1244	S1	120.00	25-Mar-2019
George Williams	615-290-2556	155 Maple, Nash, TN	37119	John T. Okon	615-123-5589	S1	250.00	17-Jul-2018
Anne G. Farriss	713-382-7185	2119 Elm, Crew, KY	25432	Alex B. Alby	713-228-1249	T2	100.00	03-Dec-2018
Olette K. Smith	615-297-3809	2782 Main, Nash, TN	37118	John T. Okon	615-123-5589	S2	500.00	14-Mar-2019

C_NAME = Customer name

C_PHONE = Customer phone

C_ADDRESS = Customer address

C_ZIP = Customer zip code

A_NAME = Agent name

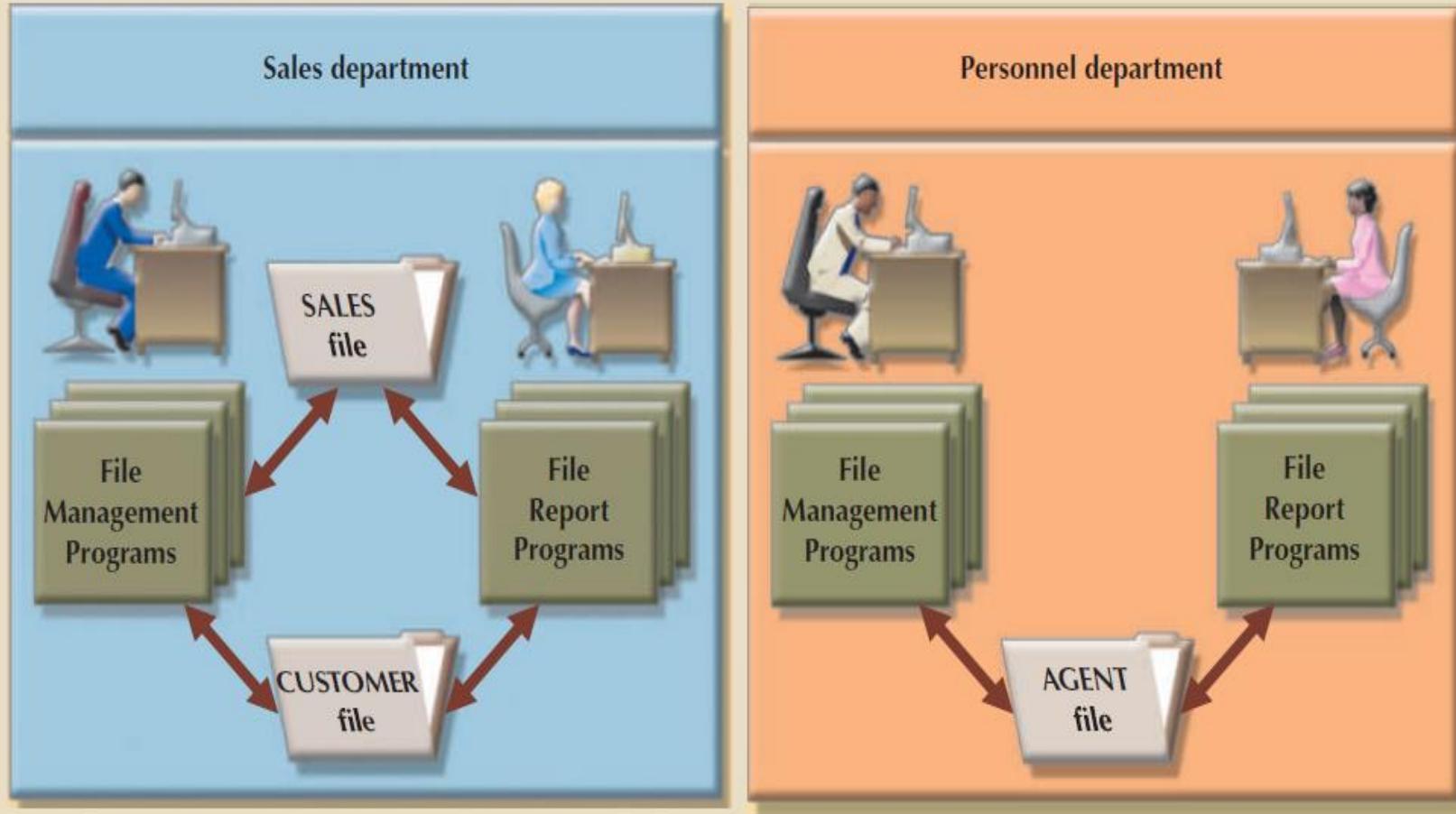
A_PHONE = Agent phone

TP = Insurance type

AMT = Insurance policy amount, in thousands of \$

REN = Insurance renewal date

FIGURE 1.9 A SIMPLE FILE SYSTEM



Issues with Computer File Data



- Lengthy development time
- Difficulty getting answers
- Complex system administration
- Lack of security
- Limited data sharing
- Extensive programming

More Data Issues

STRUCTURE/ DATA DEPENDENCE/INDEPENDENCE

Can you change the **structure** without breaking the system?

Can you change the **data** without breaking the system?

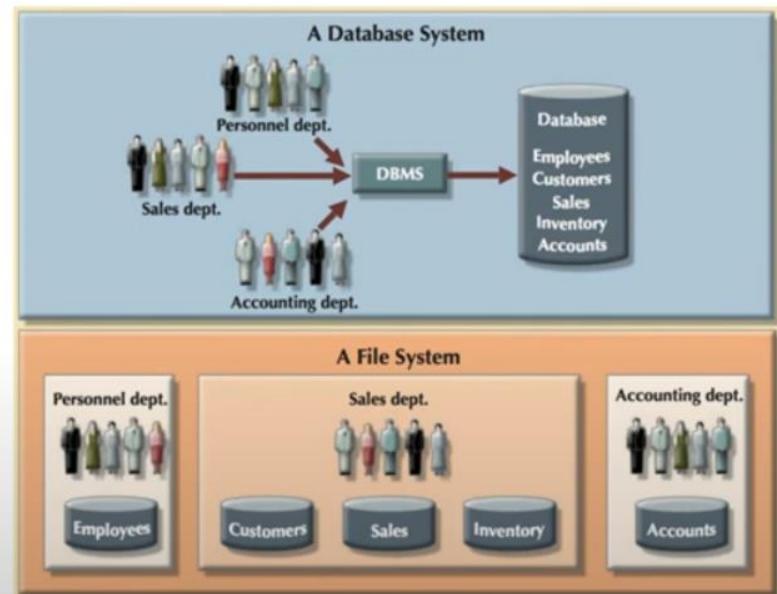
DATA REDUNDANCY

- Poor data security
- Data inconsistency
- Data Entry errors
- Data integrity
- Data Anomalies

Logical data format (how you visualize) vs. physical data format (actual database)

Database System Environment

- Hardware – Physical Devices
- Data / Information
- Software
 - Operating System Software
 - DBMS
 - Application / Utility
- People
 - System Admin
 - Database Admin
 - Database Designer
 - System Analyst / Programmer
 - End Users
- Procedures / Workflow



DBMS Functions

- ❖ Data Dictionary Management
- ❖ Data Storage Management
- ❖ Data Transformation and Presentation
- ❖ Security Management
- ❖ Multi-user Access Control
- ❖ Backup and Recovery Management
- ❖ Data Integrity Management
- ❖ Database Access Languages (SQL)
- ❖ Application Programming Interface
- ❖ Database Communication Interface

Some issues with DBMS

Increase Costs

Management Complexity

Maintaining Currency (keeping current, not money)

Vendor Dependence

Frequent Update/Replacement Cycles

Want a Job with Databases?

JOB TITLE	DESCRIPTION	SAMPLE SKILLS REQUIRED
Database Developer	Create and maintain database-based applications	Programming, database fundamentals, SQL
Database Designer	Design and maintain databases	Systems design, database design, SQL
Database Administrator	Manage and maintain DBMS and databases	Database fundamentals, SQL, vendor courses
Database Analyst	Develop databases for decision support reporting	SQL, query optimization, data warehouses
Database Architect	Design and implementation of database environments (conceptual, logical, and physical)	DBMS fundamentals, data modeling, SQL, hardware knowledge, etc.
Database Consultant	Help companies leverage database technologies to improve business processes and achieve specific goals	Database fundamentals, data modeling, database design, SQL, DBMS, hardware, vendor-specific technologies, etc.
Database Security Officer	Implement security policies for data administration	DBMS fundamentals, database administration, SQL, data security technologies, etc.
Cloud Computing Data Architect	Design and implement the infrastructure for next-generation cloud database systems	Internet technologies, cloud storage technologies, data security, performance tuning, large databases, etc.
Data Scientist	Analyze large amounts of varied data to generate insights, relationships, and predictable behaviors	Data analysis, statistics, advanced mathematics, SQL, programming, data mining, machine learning, data visualization

Remember...

Data is raw facts, Information is processed data

Data is stored in a database, that is managed by a DBMS

Database design is very, very important

Limitations include programming, system admin and money

TERM	DEFINITION
Data	Raw facts, such as a telephone number, a birth date, a customer name, and a year-to-date (YTD) sales value. Data has little meaning unless it has been organized in some logical manner.
Field	A character or group of characters (alphabetic or numeric) that has a specific meaning. A field is used to define and store data.
Record	A logically connected set of one or more fields that describes a person, place, or thing. For example, the fields that constitute a record for a customer might consist of the customer's name, address, phone number, date of birth, credit limit, and unpaid balance.
File	A collection of related records. For example, a file might contain data about the students currently enrolled at Gigantic University.