Definitions

Definitions

- Data
  - Context-free facts about a concept, object, or event.
- Information
  - Data in context. Multiple data items arranged in such a way as to allow an analytical interpretation of the facts.
- Data management
  - A systematic process for the storage and retrieval of data.
- Metadata (schema)
  - Data about the data. The definition of the tables and other objects in which data are stored.
- Query
  - The technical implementation of a “data question” asked of the database.

The Database Approach

Learning Objectives

- Formalize the concepts of data, information, data management, and metadata (schema)
- Explain what a database is and why databases are important
- Describe a database management system
- Differentiate between the DBMS and a database
- Describe the different data models and abstraction layers

What Are Databases?

Databases are collections of data, usually describing events, objects, and concepts.

Common database models

- Flat
- **Relational***
- Dimensional
- Object oriented
- Document model

** We will be working exclusively with the relational model in this class.*

Traditional File Processing Systems

Early electronic storage was inefficient

- Indexed Sequential Access Method (ISAM)
- Program-data dependency
- Duplication of data
- Lengthy development times
- Heavy maintenance load

The Database Approach

Simple tabular databases work well for very small data sets.

TO DO LIST

Wednesday, June 1, 2016

Due Today: 1

Overdue: 2

Done	Description	Due Date	Priority	Assigned to
	Mow grass	6/2/2016	High	Mom
✔	~~Clean Room~~	~~5/31/2016~~	~~Medium~~	~~Child 1~~
	Clean Room	5/31/2016	High	Child 2
✔	~~Clean Room~~	~~5/31/2016~~	~~Medium~~	~~Child 3~~
	Clean Room	5/31/2016	Medium	Child 4
	Organize Pantry	6/1/2016	Low	Dad

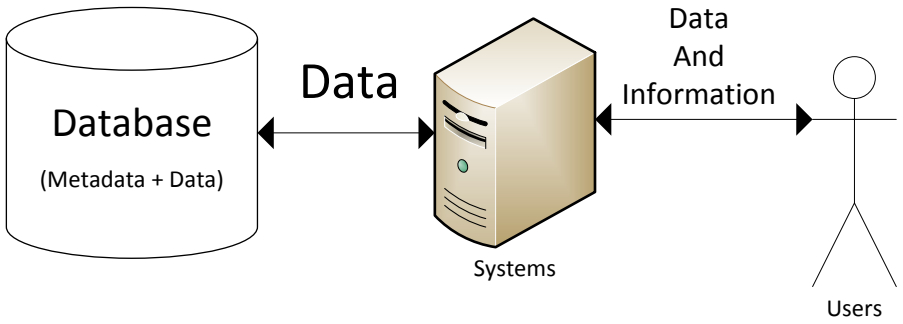
Elements of a DBMS

- Data definition mechanism
  - Provides a means for structuring and describing the stored data
- Storage and retrieval mechanism
  - Provides a means for data entry and recall
- Data administration mechanism
  - Provides the means to limit access to data, backup and restore database elements, manage database performance, and other tasks

Fundamental Axioms of DBMS

- Users communicate with computer applications (websites, etc.)
- Computer applications communicate with DBMS.
- Users do **not** communicate with DBMS directly.
- As a result, DBMS, although they can be used interactively, are **not** used interactively.
- **A DBMS is not a replacement for sound database design principles.**

Systems use DBMS.



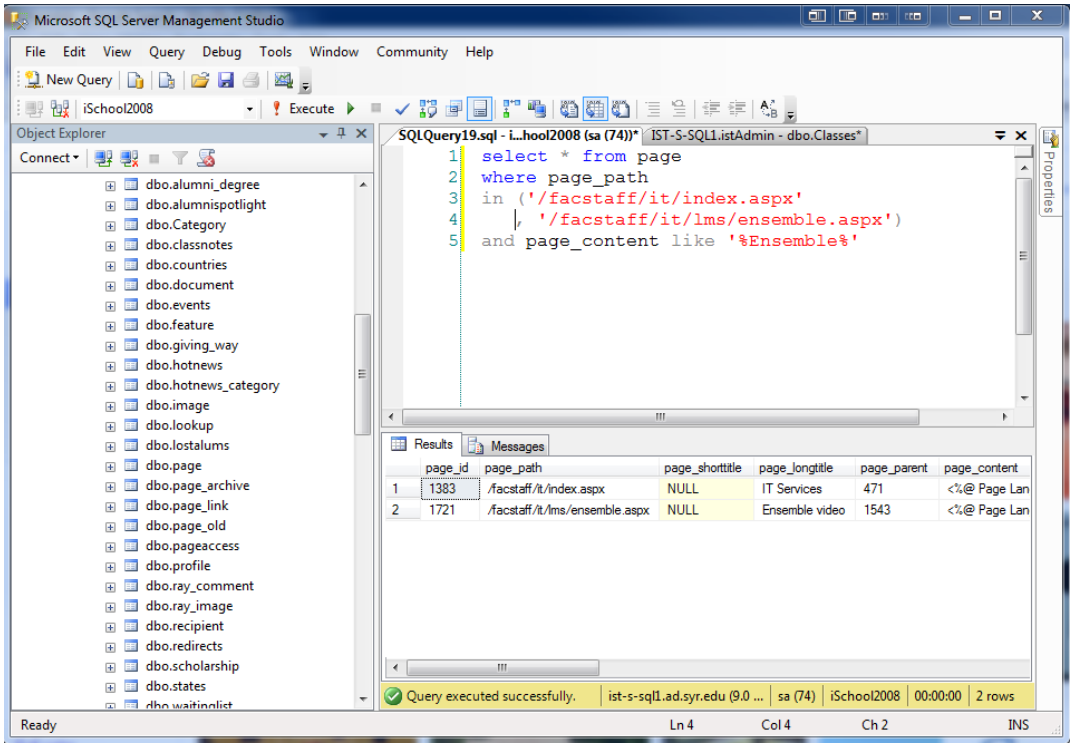
Users use Systems

Popular Database Management Systems

Learning Objectives

- Explain the similarities and differences among DBMS products (see additional reading with class notes)
- Explain DBMS history and modern uses

SQL Server: Database or DBMS?

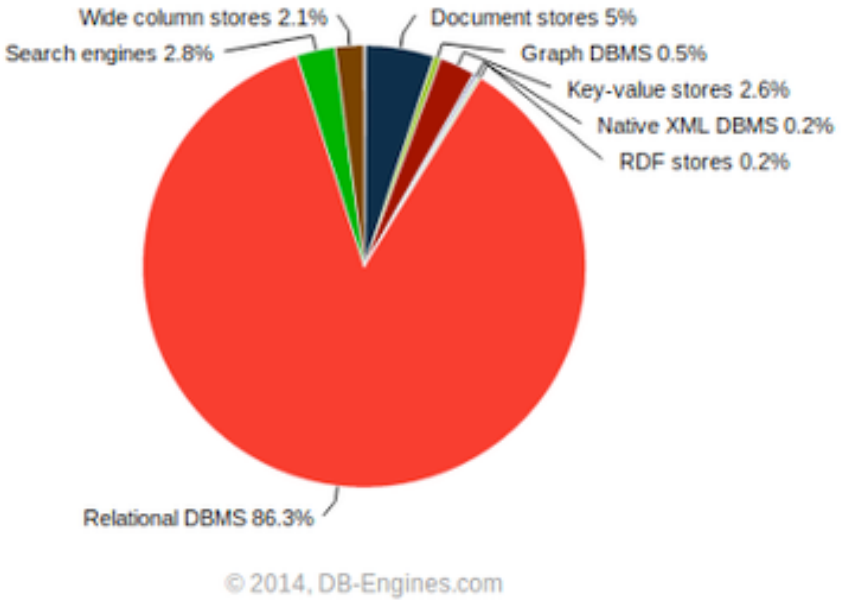


2015 DBMS Ranking

Rank Aug 2015	Rank Jul 2015	Rank Aug 2014	DBMS	Database Model	Score Aug 2015	Score Jul 2015	Score Aug 2014
1.	1.	1.	Oracle	Relational DBMS	1453.02	-3.70	-17.83
2.	2.	2.	MySQL	Relational DBMS	1292.03	+8.69	+10.81
3.	3.	3.	Microsoft SQL Server	Relational DBMS	1108.66	+5.60	-133.84
4.	4.	↑ 5.	MongoDB	Document store	294.65	+7.26	+57.30
5.	5.	↓ 4.	PostgreSQL	Relational DBMS	281.86	+9.04	+32.01
6.	6.	6.	DB2	Relational DBMS	201.23	+3.12	-5.19
7.	7.	7.	Microsoft Access	Relational DBMS	144.20	-0.10	+4.58
8.	8.	↑ 10.	Cassandra	Wide column store	113.99	+1.28	+32.09
9.	9.	↓ 8.	SQLite	Relational DBMS	105.82	-0.05	+16.95
10.	10.	↑ 11.	Redis	Key-value store	98.81	+3.73	+28.01

Source: ***http://db-engines.com/en/ranking*** ***(08/23/2015)***

DBMS Deployment Rankings



Source:Db-engines.com

An Employable Skill in the Future

Projections data from the National Employment Matrix

Occupational Title	SOC Code	Employment, 2008	Projected Employment, 2018	Change, 2008-18	
				Number	Percent
Computer network, systems, and database administrators	—	961,200	1,247,800	286,600	30
Database administrators	15-1061	120,400	144,700	24,400	20
Network and computer systems administrators	15-1071	339,500	418,400	78,900	23
Network systems and data communications analysts	15-1081	292,000	447,800	155,800	53
All other computer specialists	15-1099	209,300	236,800	27,500	13

NOTE: Data in this table are rounded. See the discussion of the employment projections table in the *Handbook* introd on *Occupational Information Included in the Handbook*.

Source BLS: 2008
http://www.bls.gov/oes/current/oes151061.htm

Advantages of the Database Approach

Program-Data Independence and Improved Data Sharing

- Separates the data from the applications that use it

- Allows for many different applications to interact with the data
  - PCs/Macs
  - Mobile platforms
  - Web applications
  - Service frameworks

Planned Data Redundancy and Improved Data Consistency

- Does not completely eliminate redundancy but enables control over the type and amount of redundancy
- Minimizing redundancy improves data consistency.
- “Single source of the truth”

Employees' Skills

Employee ID	Employee Address	Skill
426	87 Sycamore Grove	Typing
426	87 Sycamore Grove	Shorthand
519	94 Chestnut Street	Public Speaking
519	96 Walnut Avenue	Carpentry

Productive Application Development

- Reduces the cost and time of developing data-centric applications
- Standardized database formats allow for many purpose built tools
- Common design activities strengthen communication between stakeholders and developers

Improved Data Quality and Enforcement of Standards

- By establishing rules about how our data are stored, we can improve the quality of the data being stored.

- We can use business rules to inform standards that improve data quality.

StudentID	Name	GPA
123	Jeff	3.08
124	Kim	3.53
125	Julie	3.51
126	Josh	5.00

Other Advantages

- Improved accessibility

- Reduced maintenance

- Improved decision support

```
1  select Play.*, Act.* from Act
2  inner join Play on Play.PlayID = Act.ActID
```

100 %

Results | Messages

	PlayID	Title	Synopsis	GenreID	ActID	Act	PlayID
1	1	Henry IV, Part I	Henry IV, Part 1, culminates in the battle of Shrew...	3	1	1	1
2	2	Henry VI, Part 1	With an underage boy now king of England, Henr...	3	2	2	1
3	3	Henry IV, Part 2	Henry IV, Part 2, continues the story of Henry IV, ...	3	3	3	1
4	4	Henry VI, Part 2	With a weak, unworldly king on the throne, the E...	3	4	4	1

Click to select the

Some Cautions

- DBMSs do many things but still require good design principles.
- DBMS implementations require certain sets of specialized skills.
- Shared data creates a need for organizational concordance (data governance).

A Data Modeling Exercise

A Data Modeling Exercise

Trip Name	Associated Course	Dates	Student Cost	Coordinator Name	Satisfies Global Req	Notes
EuroTech	GET487/687	Aug 1 – Aug 19 (2015) July 19 – Aug 10 (2014)	4500.00 4250	Carrie Allen	Yes Yes	Limit: 30 Other: Faculty is vegan.
Retail Road Trip - NYC		10/25/12 thru 10/27	$110	Julie Salas	No	Alumni: Reception, Site Hosts Limit: 42
Buffalo Road Trip - Buffalo	NA	Jan 18 2013	45	Salas, Julie	NA	Alumni: Site Hosts Limit: 35
AsiaTech	GET487, GET687	June 1, 2013 thru 6/15	$4,500.00	Carrie Allyn	True	Limit: 30
SBinSV	IST500	3/9/15 – 3/15/15	$500.00	Jane Libby	No	Limit: 15
Data by the Lake-Chicago	IST 687	March 9 – 13 (2015)	$500	Jen Pepper		Limit: 16
EntreTech-NYC	IST 500	5/10/15 thru 5/15/15	$500	Jane Libby	No	Max Cap: 15 Alumni: Reception Other: Don't use Billy Bus Service again!
Great Libraries - Florence	YES	5/25/14- 6/1/14	$1500	Julie Salas	Yes	Limit: 6 Cost: Doesn't include flight or meals Other: Allow Alums (4)
Sports Road Trip - NYC		2/21/13 – 2/22/13	$125.00	Julie Salas	No	Limit: 44 Alumni: Reception, Site Hosts
Media/ Entertainment Road Trip - NYC	No	April 14 2014 – April 18 2014	$40	Julie Salas	No	Limit: 53
AsiaTech	GET487/678	5/30/14 – 6/14/14	4,250	Carrie Allyn	yes	Limit: 30

** All data is purely fictitious – it in no way reflects actual iSchool related road trips! Facts and figures have been changed to protect the innocent.*


SYRACUSE UNIVERSITY
SUOS
CULTORES
SCIENTIA
CORONAT