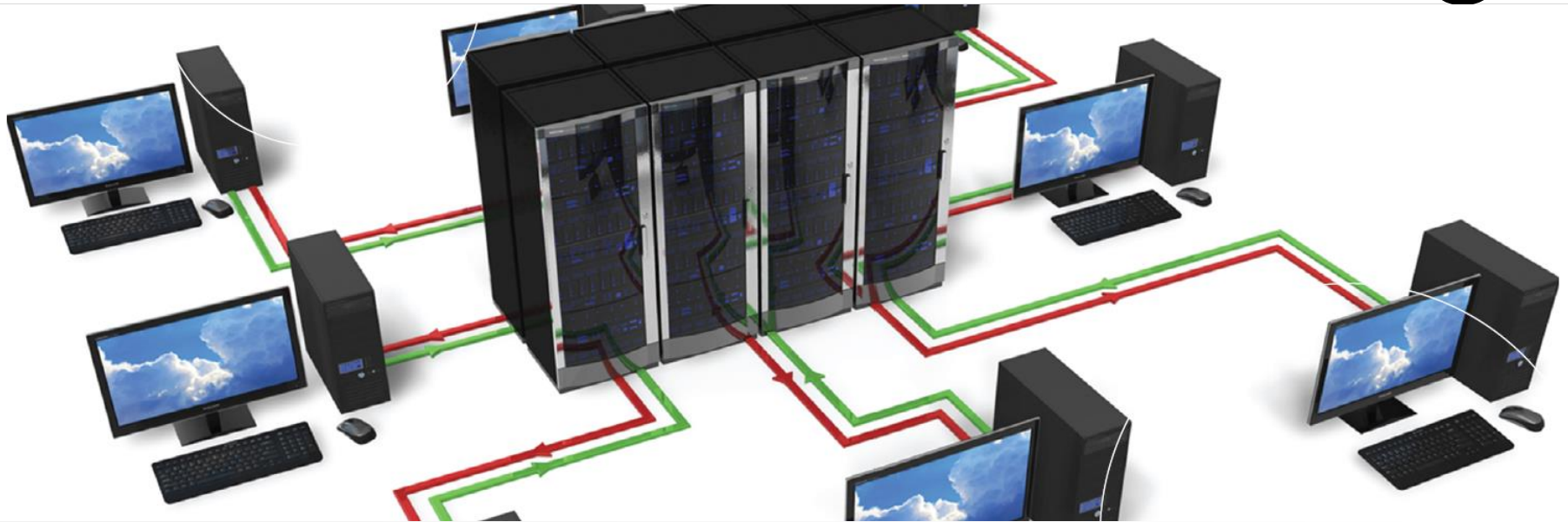



CHAPTER 3



Data and Knowledge Management

-
1. Managing Data
 2. The Database Approach
 3. Big Data
 4. Data Warehouses and Data Marts
 5. Knowledge Management



-
1. Discuss ways that common challenges in managing data can be addressed using data governance.
 2. Discuss the advantages and disadvantages of relational databases.
 3. Define Big Data, and discuss its basic characteristics.
-
- 

-
4. Explain the elements necessary to successfully implement and maintain data warehouses.
 5. Describe the benefits and challenges of implementing knowledge management systems in organizations.



Opening Case

- **Database Saves the State of Washington Medicaid Dollars**
 1. Describe additional benefits (beyond those discussed in the case) of the State of Washington's EDIE database.
 2. Describe potential disadvantages of the State of Washington's EDIE database.

About Business 3.1

- Giving Data Away for Free
 1. At first glance, giving away data for free seems to be a questionable business strategy. Describe how Jana's business model makes this practice a success.
 2. Why are Facebook and Google so interested in bringing Internet access to the developing world?
 3. Discuss the pros and cons of the zero-rating system.
 4. Is Jana correct in claiming that its system is not really a zero-rating system? Why or why not?

3.1 Managing Data

- Difficulties of Managing Data
- Data Governance

The Difficulties of Managing Data

- The amount of data increases exponentially over time
 - Data are scattered throughout organizations
 - Data are generated from multiple sources (internal, personal, external)
 - Clickstream Data
 - New sources of data
-

The Difficulties of Managing Data (continued)

- Data Degradation
 - Data Rot
 - Data security, quality, and integrity are critical
 - Legal requirements change frequently and differ among countries & industries
-

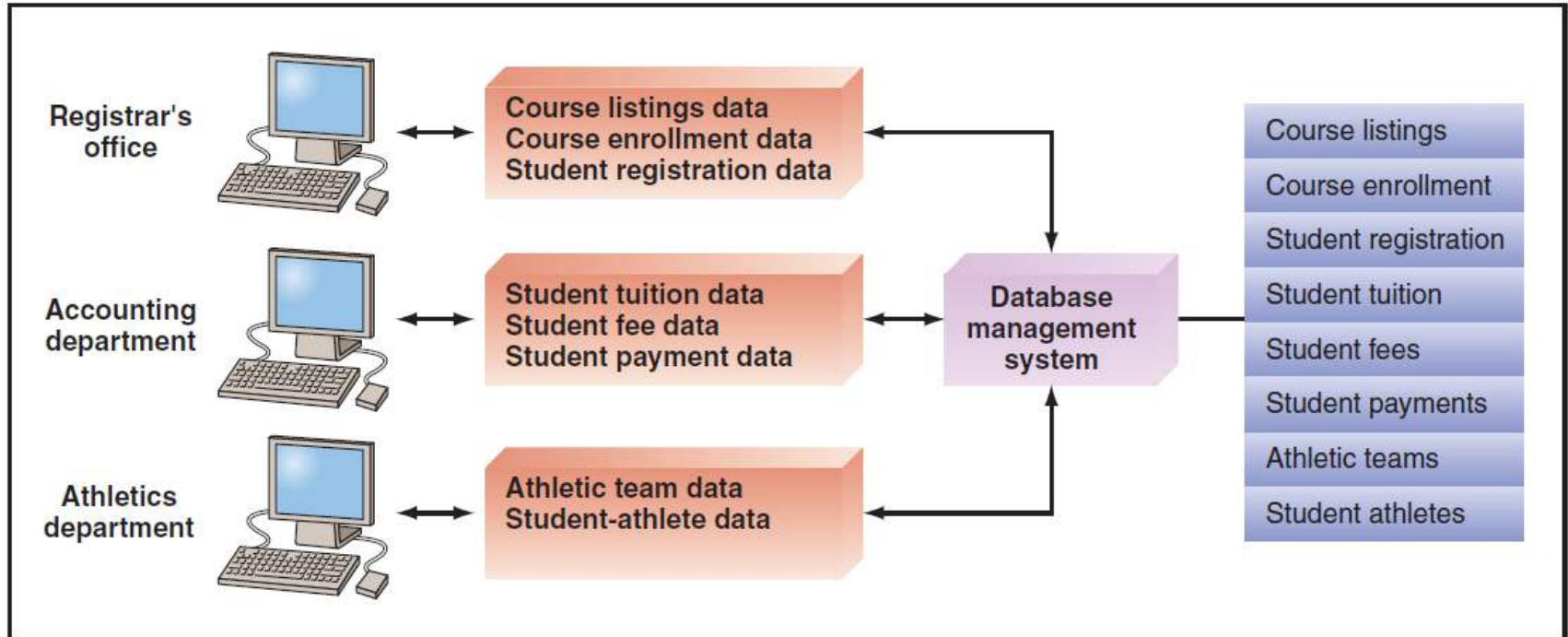
Data Governance

- Master Data Management
- Master Data

3.2 The Database Approach

- Data File
 - Database Systems Minimize & Maximize Three Things
 - The Data Hierarchy
 - The Relational Database Model
-

Figure 3.1: Database Management System



Database Management Systems (DBMS) Minimize:

- Data Redundancy
- Data Isolation
- Data Inconsistency

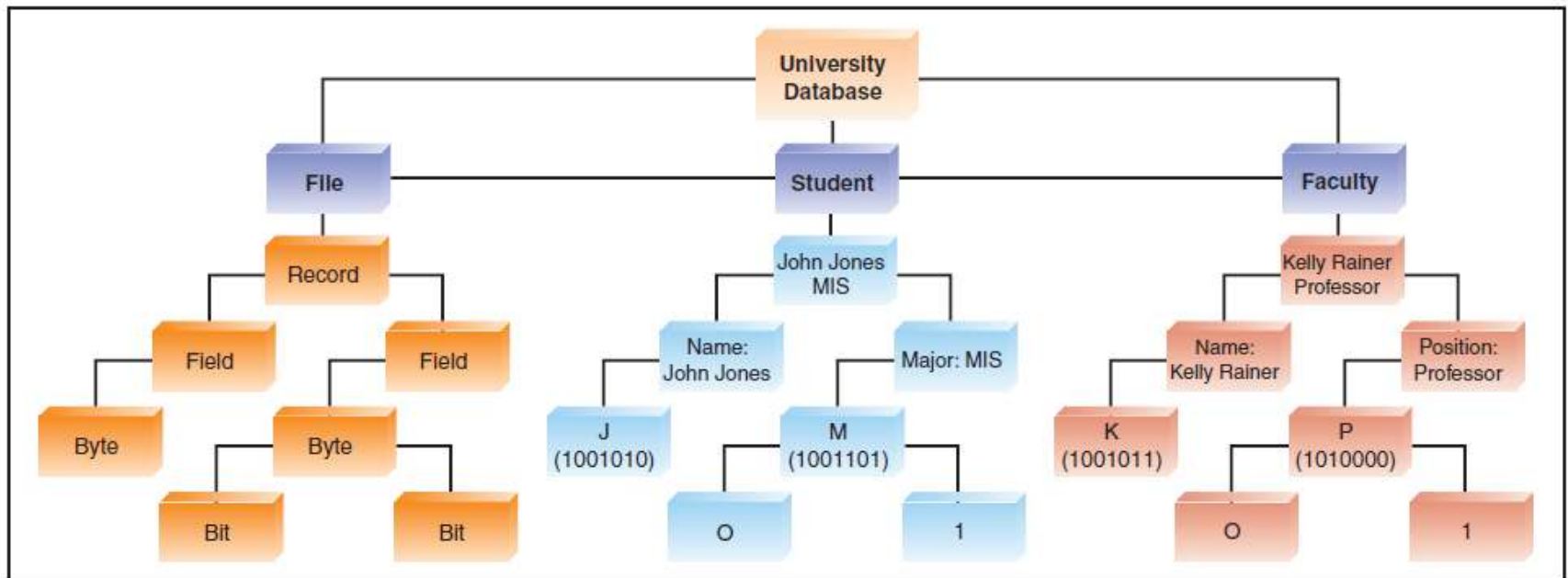
Database Management Systems (DBMS) Maximize:

- Data Security
- Data Integrity
- Data Independence

Data Hierarchy

- Bit
 - Byte
 - Field
 - Record
 - Data File (Table)
 - Database
-

Figure 3.2: Hierarchy of Data for a Computer-based File



The Relational Database Model

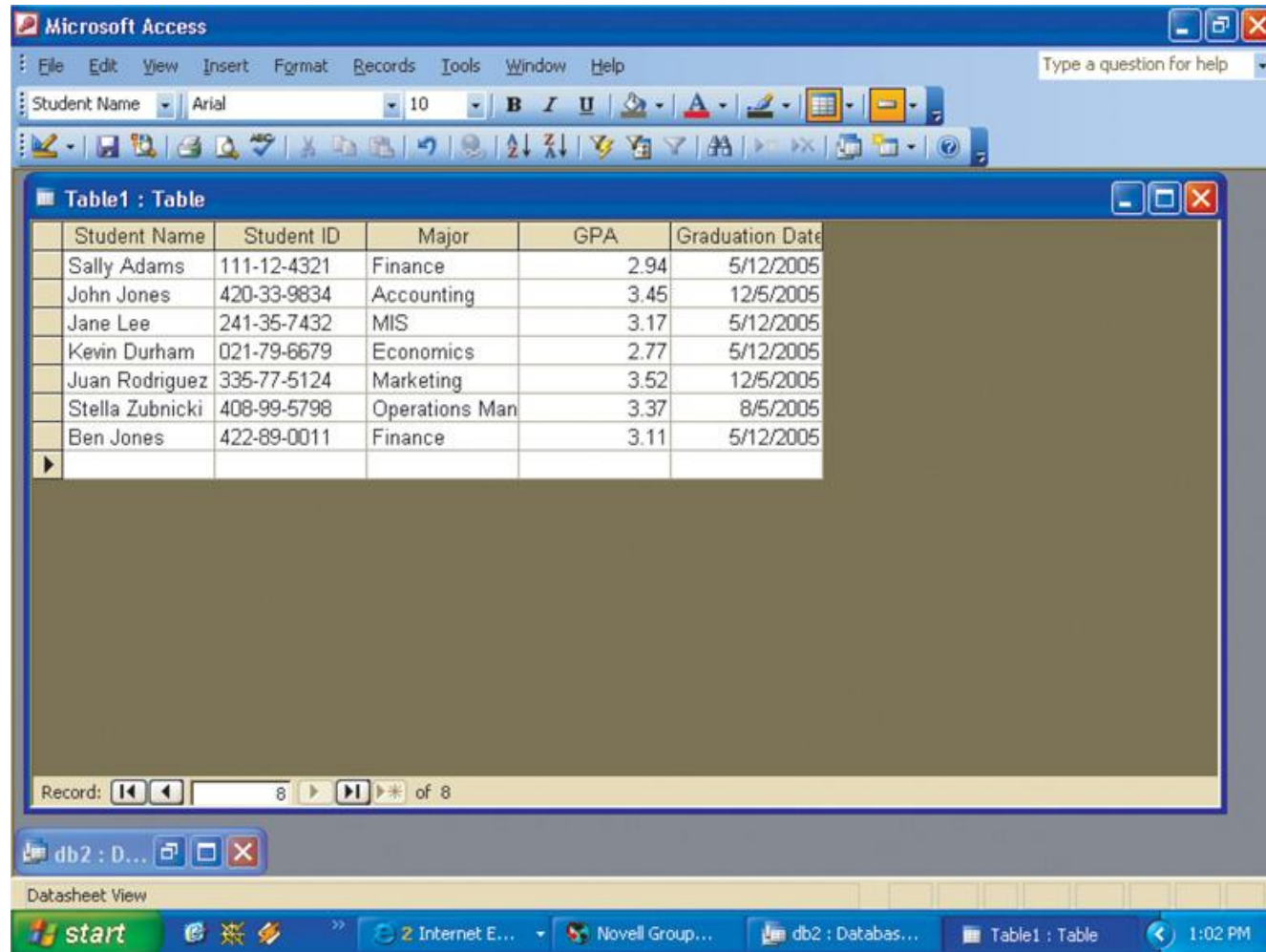
- Database Management System (DBMS)
 - Relational Database Model
 - Data Model
 - Entity
 - Instance
 - Attribute
-

The Relational Database Model (continued)

- Primary Key
- Secondary Key
- Foreign Key
- Unstructured Data



Figure 3.3: Student Database Example



3.3 Big Data

- Defining Big Data
 - Characteristics of Big Data
 - Issues with Big Data
 - Managing Big Data
 - Putting Big Data to Use
 - Big Data Used in the Functional Areas of the Organization
-


Defining Big Data

- **Gartner (www.gartner.com)**
- **Big Data Institute**


Defining Big Data: Gartner

- Diverse, high volume, high-velocity information assets that require new forms of processing to enable enhanced decision making, insight discovery, and process optimization.

Defining Big Data: The Big Data Institute (TBDI)

- Vast Datasets that:
 - Exhibit variety
 - Include structured, unstructured, and semi-structured data
 - Generated at high velocity with an uncertain pattern
 - Do not fit neatly into traditional, structured, relational databases
 - Can be captured, processed, transformed, and analyzed in a reasonable amount of time only by sophisticated information systems.
-
- 

Big Data Generally Consist of the Following:

- Traditional Enterprise Data
 - Machine-Generated/Sensor Data
 - Social Data
 - Images Captured by Billions of Devices Located Throughout the World
-
- 


Characteristics of Big Data

- Volume
- Velocity
- Variety

Issues with Big Data

- Untrusted data sources
- Big Data is dirty
- Big Data changes, especially in data streams

Managing Big Data

- Big Data makes it possible to do many things that were previously impossible:
 - Spot business trends more rapidly and accurately
 - tracking the spread of disease
 - tracking crime
 - detecting fraud
-
- 

Managing Big Data (continued)

- **First Step:**
 - Integrate information silos into a database environment and develop data warehouses for decision making.
- **Second Step:**
 - making sense of their proliferating data.



Managing Big Data (continued)

- Many organizations are turning to NoSQL databases to process Big Data



About Business 3.2

- TrueCar Uses Hadoop
 1. Describe how Hadoop manages Big Data in its data lake.
 2. Discuss why relational databases experienced problems with the variety of data that TrueCar has to manage and analyze.
 3. What are the benefits of Big Data to TrueCar?

Putting Big Data to Use

- Making Big Data Available
 - Enabling Organizations to Conduct Experiments
 - Micro-Segmentation of Customers
 - Creating New Business Models
 - Organizations Can Analyze Far More Data
-

About Business 3.3

- Combining Big Data and Open Data to Fight Ebola
 1. Provide examples of open data mentioned in this case.
 2. Provide examples of Big Data mentioned in this case.
 3. Why was the integration of open data and Big Data essential to help lessen the impact of the Ebola virus?

Big Data Used in the Functional Areas of the Organization

- Human Resources
 - Product Development
 - Operations
 - Marketing
 - Government Operations
-
- 

3.4 Data Warehouses and Data Marts

- Describing Data Warehouses and Data Marts
- A Generic Data Warehouse Environment

Describing Data Warehouses and Data Marts

- Organized by business dimension or Use online analytical processing (OLAP)
 - Integrated
 - Time variant
 - Nonvolatile
 - Multidimensional
-

A Generic Data Warehouse Environment

- Source Systems
 - Data Integration
 - Storing the Data
 - Metadata
 - Data Quality
 - Governance
 - Users
-

Figure 3.4: Data Warehouse Framework

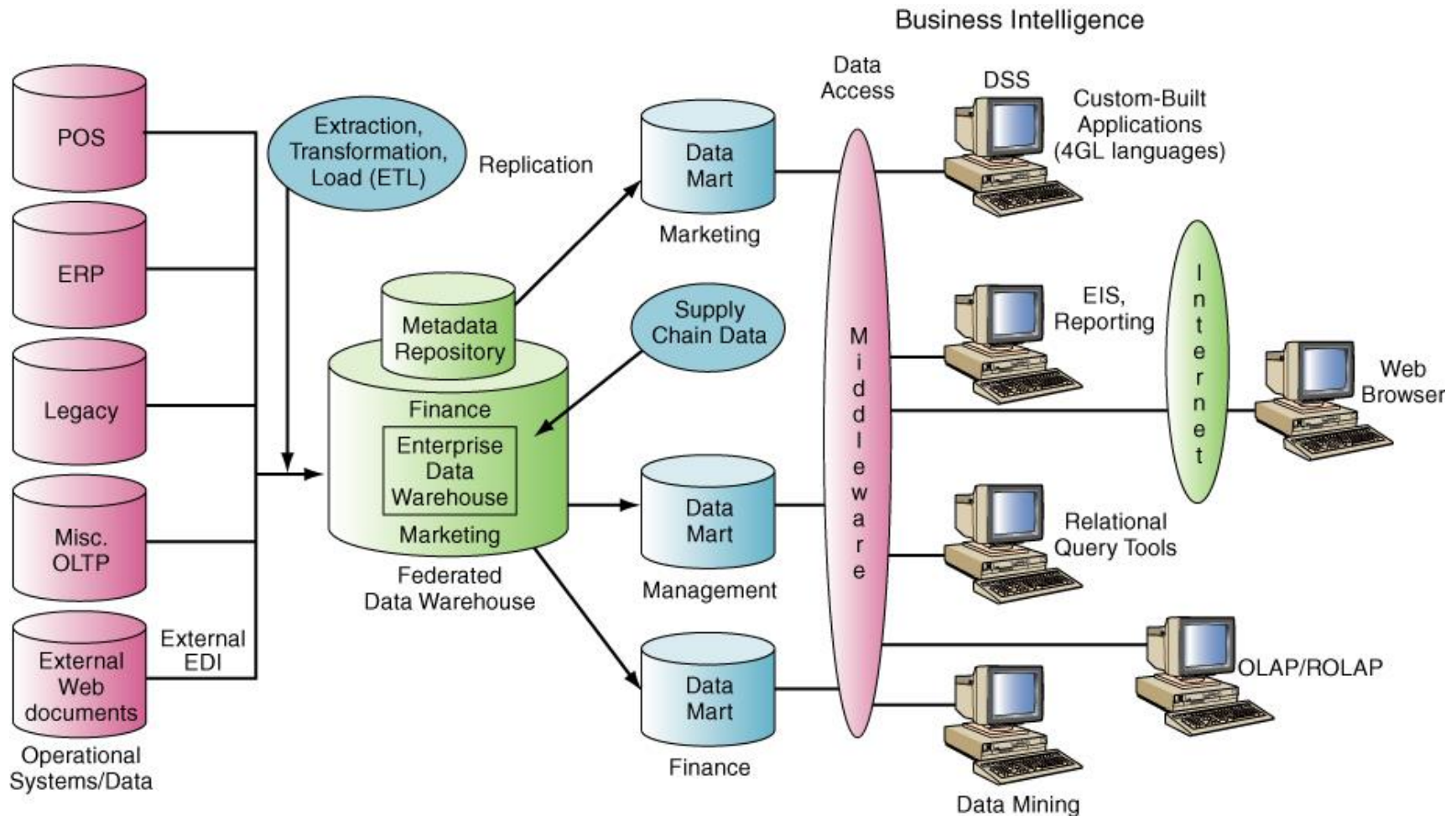


Figure 3.5: Relational Databases

(a) 2012

Product	Region	Sales
Nuts	East	50
Nuts	West	60
Nuts	Central	100
Screws	East	40
Screws	West	70
Screws	Central	80
Bolts	East	90
Bolts	West	120
Bolts	Central	140
Washers	East	20
Washers	West	10
Washers	Central	30

(b) 2013

Product	Region	Sales
Nuts	East	60
Nuts	West	70
Nuts	Central	110
Screws	East	50
Screws	West	80
Screws	Central	90
Bolts	East	100
Bolts	West	130
Bolts	Central	150
Washers	East	30
Washers	West	20
Washers	Central	40

(c) 2014

Product	Region	Sales
Nuts	East	70
Nuts	West	80
Nuts	Central	120
Screws	East	60
Screws	West	90
Screws	Central	100
Bolts	East	110
Bolts	West	140
Bolts	Central	160
Washers	East	40
Washers	West	30
Washers	Central	50

Figure 3.6: Data Cube

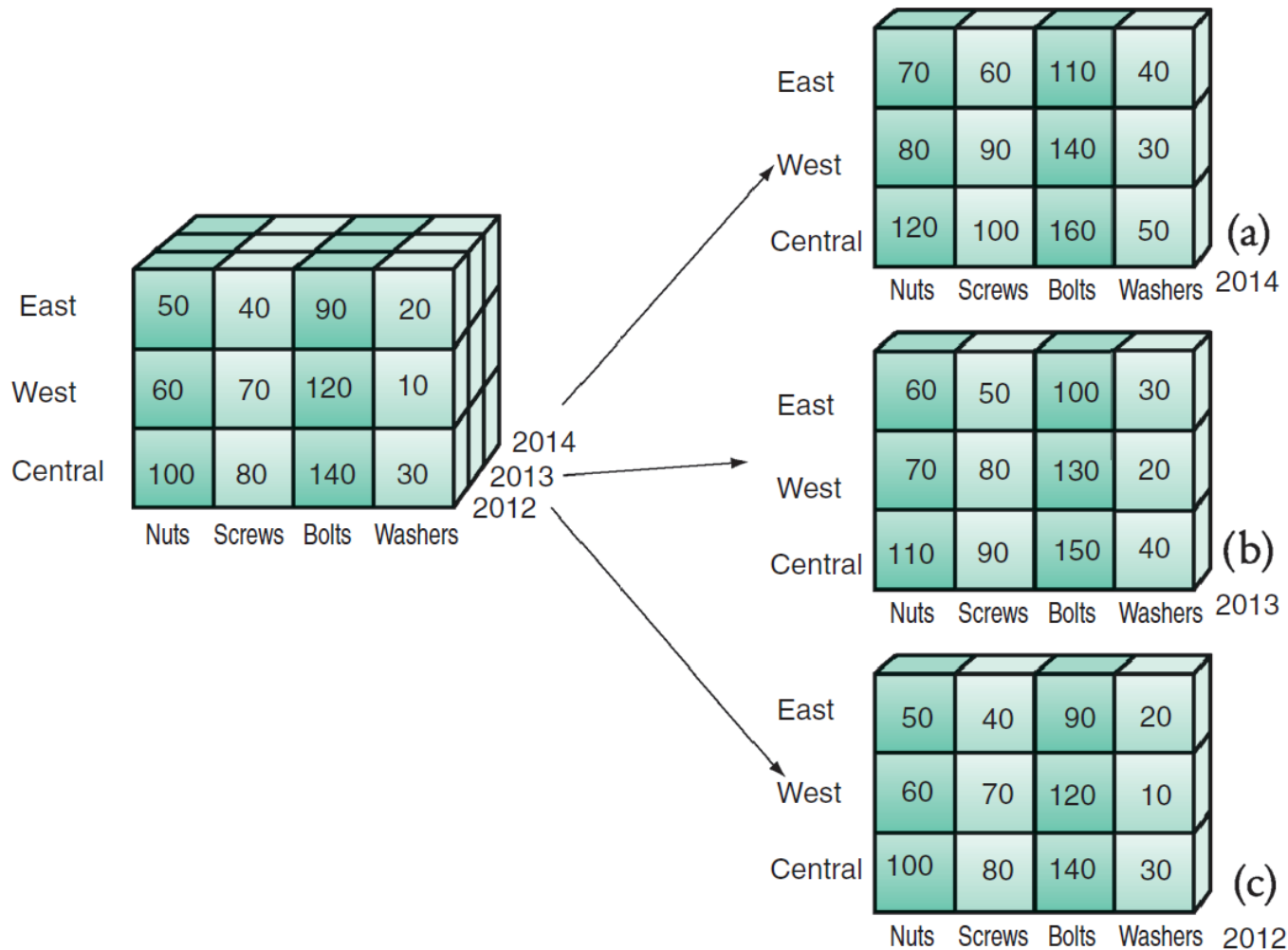
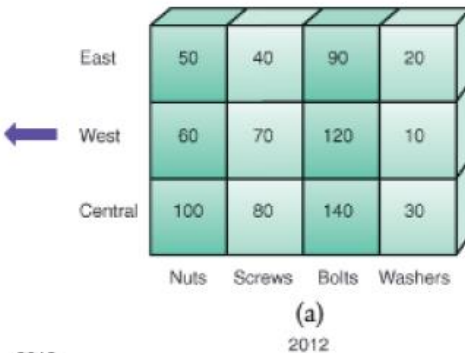


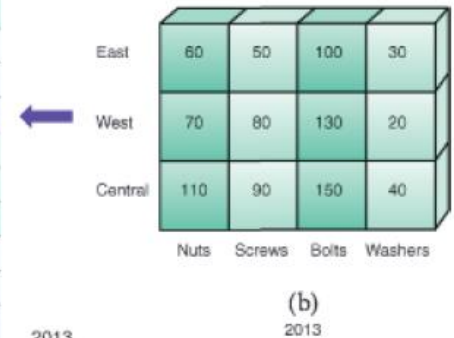
Figure 3.7: Equivalence Between Relational and Multidimensional Databases

Product	Region	Sales
Nuts	East	50
Nuts	West	60
Nuts	Central	100
Screws	East	40
Screws	West	70
Screws	Central	80
Bolts	East	90
Bolts	West	120
Bolts	Central	140
Washers	East	20
Washers	West	10
Washers	Central	30



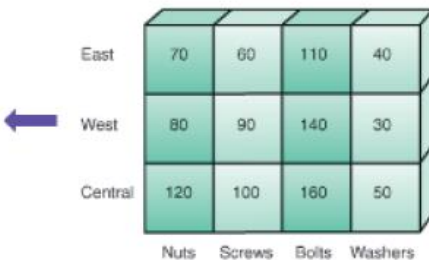
2012

Product	Region	Sales
Nuts	East	60
Nuts	West	70
Nuts	Central	110
Screws	East	50
Screws	West	80
Screws	Central	90
Bolts	East	100
Bolts	West	130
Bolts	Central	150
Washers	East	30
Washers	West	20
Washers	Central	40



2013

Product	Region	Sales
Nuts	East	70
Nuts	West	80
Nuts	Central	120
Screws	East	60
Screws	West	90
Screws	Central	100
Bolts	East	110
Bolts	West	140
Bolts	Central	160
Washers	East	40
Washers	West	30
Washers	Central	50



2014

(c)
2014

3.5 Knowledge Management

- Concepts and Definitions
- Knowledge Management Systems
- The KMS Cycle

Concepts and Definitions

- Knowledge Management
- Knowledge
- Explicit and Tacit Knowledge
- Knowledge Management Systems
- The KMS Cycle

About Business 3.4

- Performance Bicycle Leverages Its Employees' Knowledge
 1. Describe several ways in which Performance Bicycle incorporates employee knowledge in its customer experience.
 2. Is Performance Bicycle capturing and using its employee's tacit knowledge or explicit knowledge? Explain your answer.

Figure 3.8: The Knowledge Management System Cycle

