USIU AFRICA

United States International University-Africa

DATABASE SYSTEMS

Presented by Prof. Elisha T. O. Omulo

USIU AFRICA

WEEK 10 AGENDA

- **Advanced Database Applications.**
- **Weaknesses of the Relational DBMSs.**
- **Object-Oriented Concepts.**

Course Textbook: **Carlos Coronel, Steven Morris, Peter Rob and Keeley Crockett Database Principles: Fundamentals of Design, Implementation, and Management, 14th Edition, 2022, ISBN-13978-0357673034.**

USIU AFRICA

Advanced Database Applications.

Database Performance Tuning and Query Optimization

Database performance tuning

- This is a set of activities and procedures designed to reduce the response time of the database system—that is, to ensure that an end-user query is processed by the DBMS in the minimum amount of time.
- The goal of database performance is to execute queries as fast as possible.
- Database performance must be closely monitored and regularly tuned.

Factors affecting the time required by a query to return a result:

- CPU processing power,
- Available primary memory (RAM),
- Input/output (hard disk and network) throughput.
- Algorithms being used

Task: Discuss database performance tuning.

Advanced Database Applications.

Database Performance Tuning and Query Optimization

Database performance tuning- factors

	SYSTEM RESOURCES	CLIENT	SERVER
Hardware	CPU	The fastest possible Dual-core CPU or higher	The fastest possible Multiple processors (quad-core technology)
	RAM	The maximum possible	The maximum possible
	Hard disk	Fast SATA/EIDE hard disk with sufficient free hard disk space	Multiple high-speed, high-capacity hard disks (SCSI/SATA/Firewire/Fibre Channel) in RAID configuration
	Network	High-speed connection	High-speed connection
Software	Operating system	Fine-tuned for best client application performance	Fine-tuned for best server application performance
	Network	Fine-tuned for best throughput	Fine-tuned for best throughput
	Application	Optimize SQL in client application	Optimize DBMS server for best performance

USIU AFRICA

Task: **Discuss database performance tuning.**

Advanced Database Applications.

Database Performance Tuning and Query Optimization

Query optimization

This the selection of a strategy for querying so that the performance is optimal.

Automatic query optimization

The DBMS finds the most cost-effective access path without user intervention.

Static query optimization

- **Takes place at compilation time.**
- **The best optimization strategy is selected when the query is compiled by the DBMS.**
- **Occurs mostly when SQL statements are embedded in procedural programming languages such as C# or Visual Basic .NET.**

Dynamic query optimization

- **Takes place at execution time.**
- **Therefore, access strategy is dynamically determined by the DBMS at run time, using the most up-to-date information about the database.**
- **Dynamic query optimization is efficient, but it has its costs in processing overhead.**

USIU AFRICA

Task: **Discuss query optimization.**

Advanced Database Applications: Distributed Database Management Systems

Desktop database- A single-user database that runs on a personal computer.

Multiuser database- supports many users at the same time.

Workgroup database- is a multiuser database for fewer users <=50

Enterprise database- is a multiuser database that supports >50, usually hundreds across many departments in an entire organization.

Centralized database- supports data located at a single site.

Distributed database - supports data distributed across several different sites.

General-purpose databases - contain a wide variety of data used in multiple disciplines eg. e, a census database that contains general demographic data.

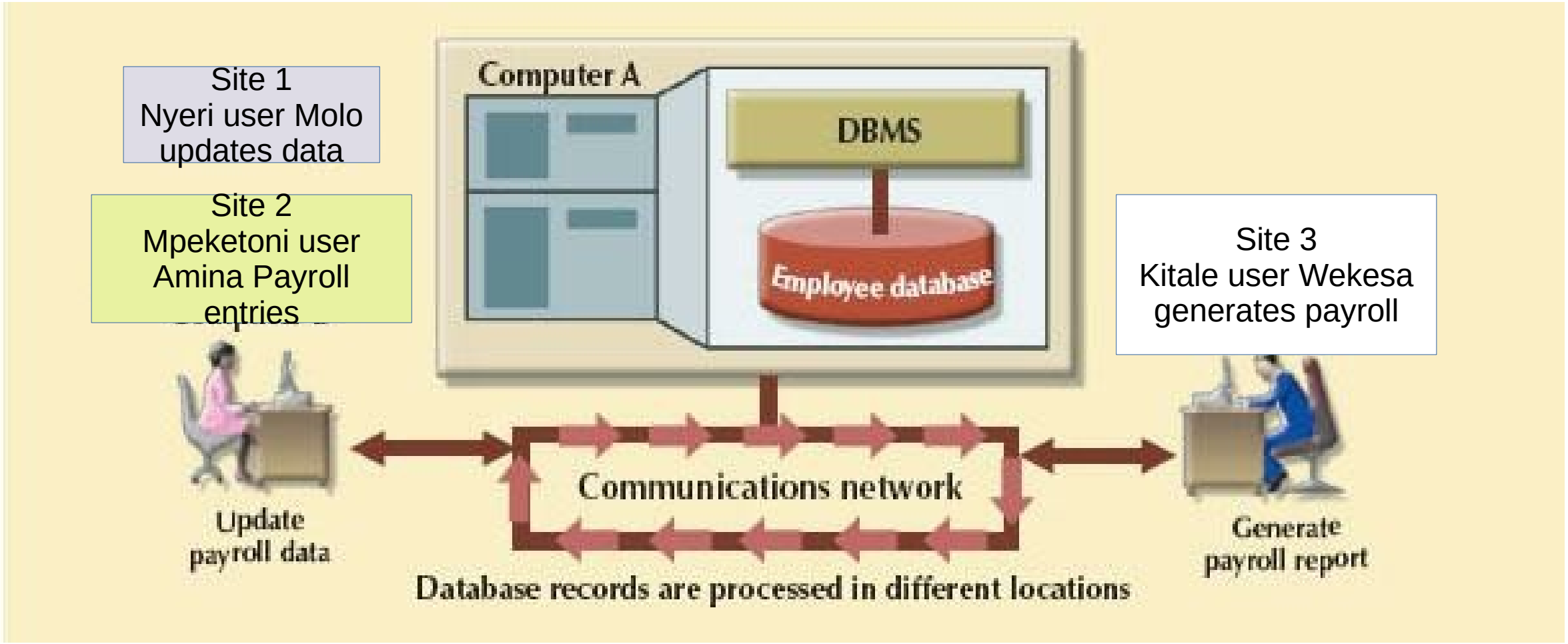
Discipline-specific databases- contain data focused on specific subject areas eg. CRSP for financial data, GIS databases that store geospatial and other related data, and medical databases.

Operational database - also known as *online transaction processing* (OLTP) , transactional, or production database. These are designed primarily to support a company's day-to-day operations.

Task: Discuss the various types of databases

Advanced Database Applications: Distributed Database Management Systems

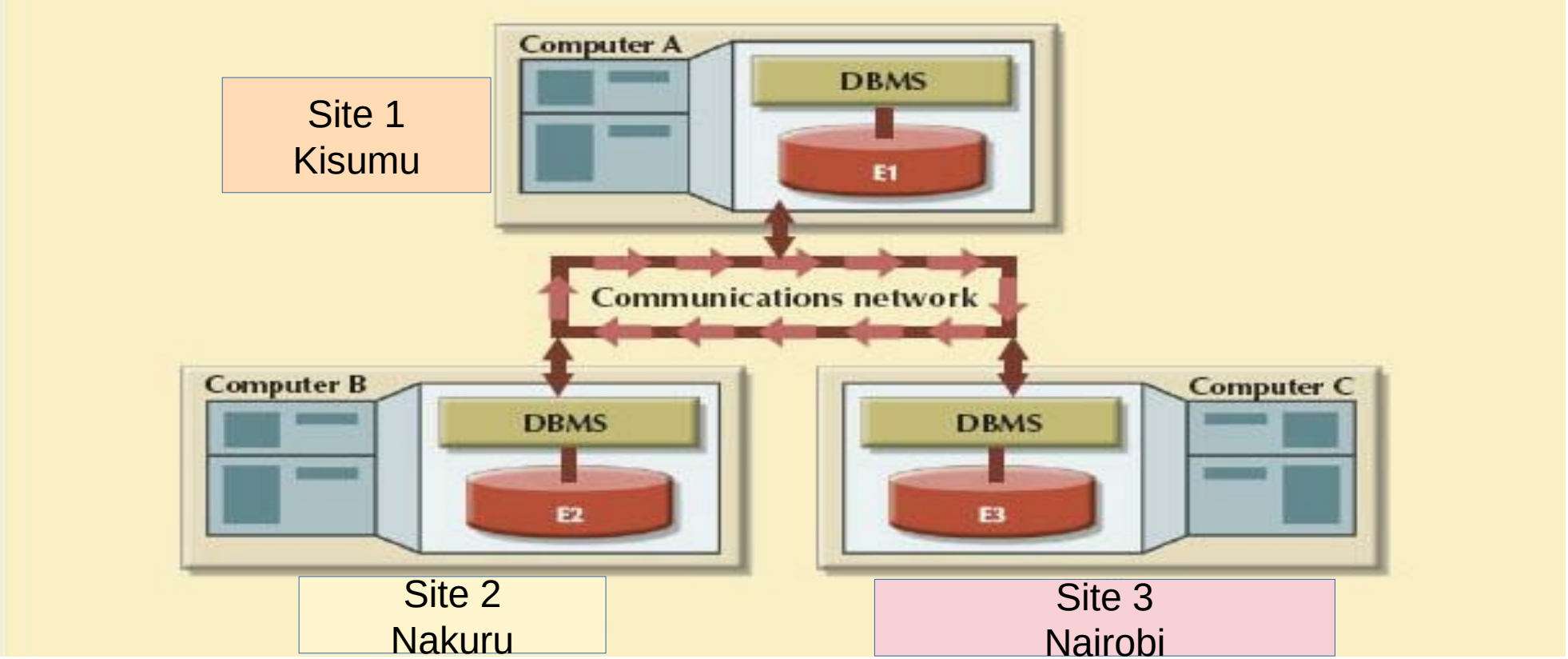
Distributed processing- can be done without a DDBMS



Task: **Discuss distributed database management systems**

Advanced Database Applications: Distributed Database Management Systems

Distributed database environment



Task: **Discuss distributed database management systems**

Advanced Database Applications: Distributed Database Management Systems

Functions of a DDBMS

- **Application interface- to interact with the end user, application programs, and other DBMSs within the distributed database**
- **Validation- to analyze data requests for syntax correctness**
- **Transformation- to decompose complex requests into atomic data request components**
- **Query optimization- to find the best access strategy (which database fragments must be accessed by the query, and how must data updates, if any, be synchronized?)**
- **Mapping- to determine the data location of local and remote fragments**
- **I/O interface- to read or write data from or to permanent local storage**
- **Formatting- to prepare the data for presentation to the end user or to an application program**

Task: Discuss distributed database management systems

Advanced Database Applications: Distributed Database Management Systems

Functions of a DDBMS

- **Security- to provide data privacy at both local and remote databases**
- **Backup and recovery- to ensure the availability and recoverability of the database in case of a failure**
- **DB administration features for the database administrator provided**
- **Concurrency control- to manage simultaneous data access and to ensure data consistency across database fragments in the DDBMS**
- **Transaction management- to ensure that the data move from one consistent state to another; this activity includes the synchronization of local and remote transactions as well as transactions across multiple distributed segments**
- **Support all functions of the centralized database management**

Task: Discuss distributed database management systems

Advanced Database Applications: Business Intelligence and Data Warehouse

Business intelligence (BI)

- Is using a comprehensive, cohesive, and integrated set of tools and processes used to capture, collect, integrate, store, and analyze data with the purpose of generating and presenting information to support business decision making.
- Depends on learning and understanding the facts about the business environment.
- Also seen as a framework that allows a business to transform data into information, information into knowledge, and knowledge in to wisdom.
- Has the potential to positively affect a company's culture by creating continuous business performance improvement through active decision support at all levels in an organization.

Task: Discuss business intelligence

Advanced Database Applications: Business Intelligence and Data Warehouse

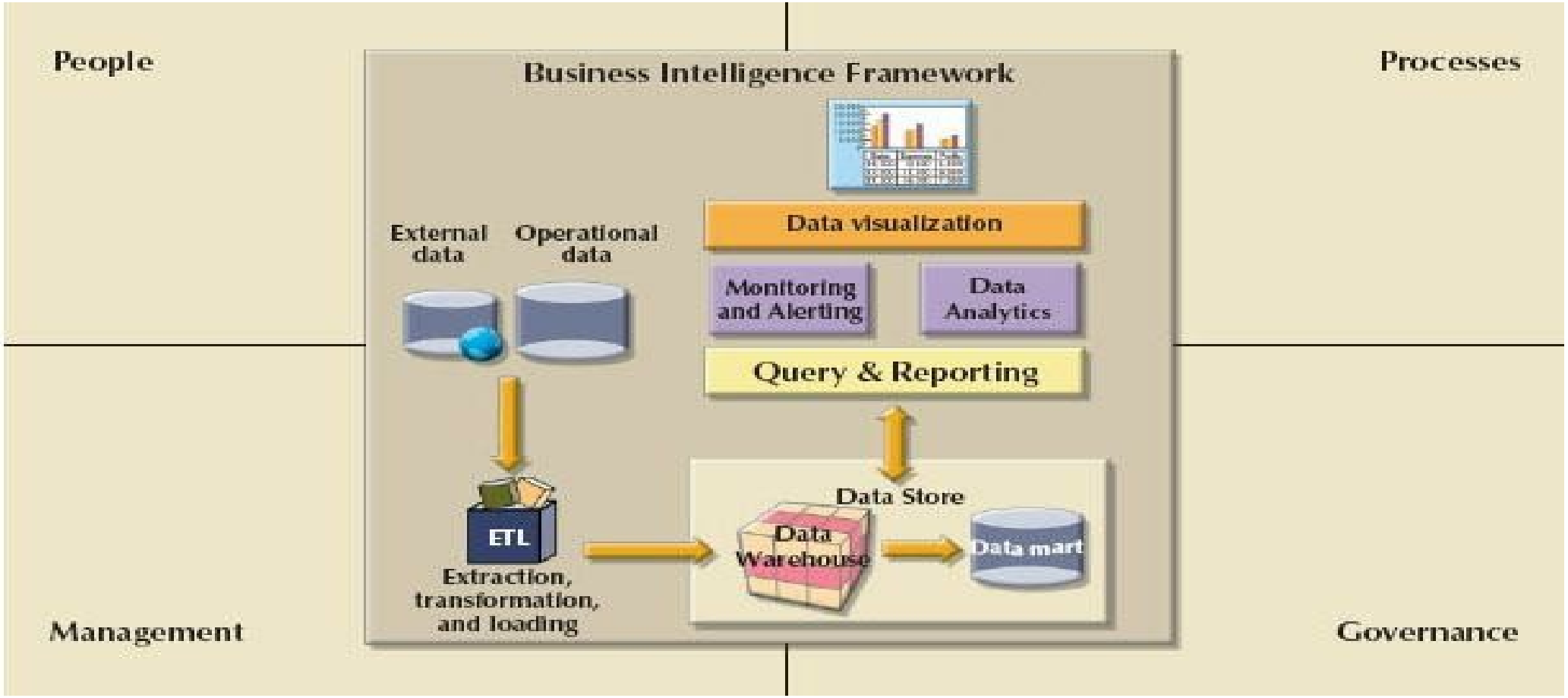
Business intelligence (BI)

- Provides business insight that empowers users to make sound decisions based on the accumulated knowledge of the business.
- Its initial adopters were high-volume industries such as financial services, insurance, and healthcare companies; spread to other industries such as telecommunications, retail/merchandising, manufacturing, media, government, education, etc.
- BI can be seen as a collection of best practices and software tools developed to support business decision making in this age of globalization, emerging markets, rapid change, and increasing regulation.
- The main activities involve gathering, generating, and presenting information to business decision makers, with data warehouses playing important roles.

Task: Discuss business intelligence

Advanced Database Applications: Business Intelligence and Data Warehouse

Business intelligence (BI)-framework



Task: Discuss business intelligence

Advanced Database Applications: Business Intelligence and Data Warehouse

COMPONENT	DESCRIPTION
ETL tools	Data **extraction, transformation, and loading (ETL)** tools collect, filter, integrate, and aggregate internal and external data to be saved into a data store optimized for decision support. Internal data are generated by the company during its day-to-day operations, such as product sales history, invoicing, and payments. The external data sources provide data that cannot be found within the company but are relevant to the business, such as stock prices, market indicators, marketing information (such as demographics), and competitors' data. Such data are generally located in external databases provided by industry groups or companies that market the data.
Data store	The data store is optimized for decision support and is generally represented by a *data warehouse* or a *data mart*. The data are stored in structures that are optimized for data analysis and query speed.
Query and reporting	This component performs data selection and retrieval, and is used by the data analyst to create queries that access the database and create the required reports. Depending on the implementation, the query and reporting tool accesses the operational database, or more commonly, the data store.
Data visualization	This component presents data to the end user in a variety of meaningful and innovative ways. This tool helps the end user select the most appropriate presentation format, such as summary reports, maps, pie or bar graphs, mixed graphs, or dashboards.
Data monitoring and alerting	This component allows real-time monitoring of business activities. The BI system will present the concise information in a single integrated view for the data analyst. This integrated view could include specific metrics about the system performance or activities, such as number of orders placed in the last four hours, number of customer complaints by product by month, and total revenue by region. Alerts can be placed on a given metric; once the value of a metric goes below or above a certain baseline, the system will perform a given action, such as e-mailing shop floor managers, presenting visual alerts, or starting an application.
Data analytics	This component performs data analysis and data-mining tasks using the data in the data store. This tool advises the user as to which data analysis tool to select and how to build a reliable business data model. Business models are generated by special algorithms that identify and enhance the understanding of business situations and problems. Data analysis can be either explanatory or predictive. Explanatory analysis uses the existing data in the data store to discover relationships and their types, and predictive analysis creates statistical models of the data that allow predictions of future values and events.

Task: **Discuss the components of business intelligence**

Advanced Database Applications: Business Intelligence and Data Warehouse

TOOL	DESCRIPTION	SAMPLE VENDORS
Dashboards and business activity monitoring	**Dashboards** use Web-based technologies to present key business performance indicators or information in a single integrated view, generally using graphics that are clear, concise, and easy to understand.	Salesforce IBM/Cognos BusinessObjects Information Builders iDashboards
Portals	**Portals** provide a unified, single point of entry for information distribution. Portals are a Web-based technology that use a Web browser to integrate data from multiple sources into a single Web page. Many different types of BI functionality can be accessed through a portal.	Oracle Portal Actuate Microsoft SAP
Data analysis and reporting tools	These advanced tools are used to query multiple and diverse data sources to create integrated reports.	Microsoft Reporting Services MicroStrategy SAS WebReportStudio
Data-mining tools	These tools provide advanced statistical analysis to uncover problems and opportunities hidden within business data.	SAP Teradata MicroStrategy MS Analytics Services
Data warehouses (DW)	The data warehouse is the foundation of a BI infrastructure. Data are captured from the production system and placed in the DW on a near real-time basis. BI provides company-wide integration of data and the capability to respond to business issues in a timely manner.	Microsoft Oracle IBM/Cognos Teradata
OLAP tools	Online analytical processing provides multidimensional data analysis.	IBM/Cognos BusinessObjects Oracle Microsoft
Data visualization	These tools provide advanced visual analysis and techniques to enhance understanding and create additional insight of business data and its true meaning.	Dundas Tableau QlikView Actuate

Task: **Discuss the tools used in business intelligence**

Advanced Database Applications: Business Intelligence and Data Warehouse

Data Warehouse

Is an integrated, subject-oriented, time-variant, nonvolatile collection of data that provides support for decision making.

Integrated – it is a centralized, consolidated database that integrates data derived from the entire organization and from multiple sources with diverse formats. All business entities, data elements, data characteristics, and business metrics are described in the same way throughout the enterprise.

Subject-oriented - its data are arranged and optimized to provide answers to questions from diverse functional areas within a company. Data are organized and summarized by topic, such as sales, marketing, finance, distribution, and transportation. For each topic, the data warehouse contains specific subjects of interest—products, customers, departments, regions or promotions.

Task: Discuss data warehousing notion.

Advanced Database Applications: Business Intelligence and Data Warehouse

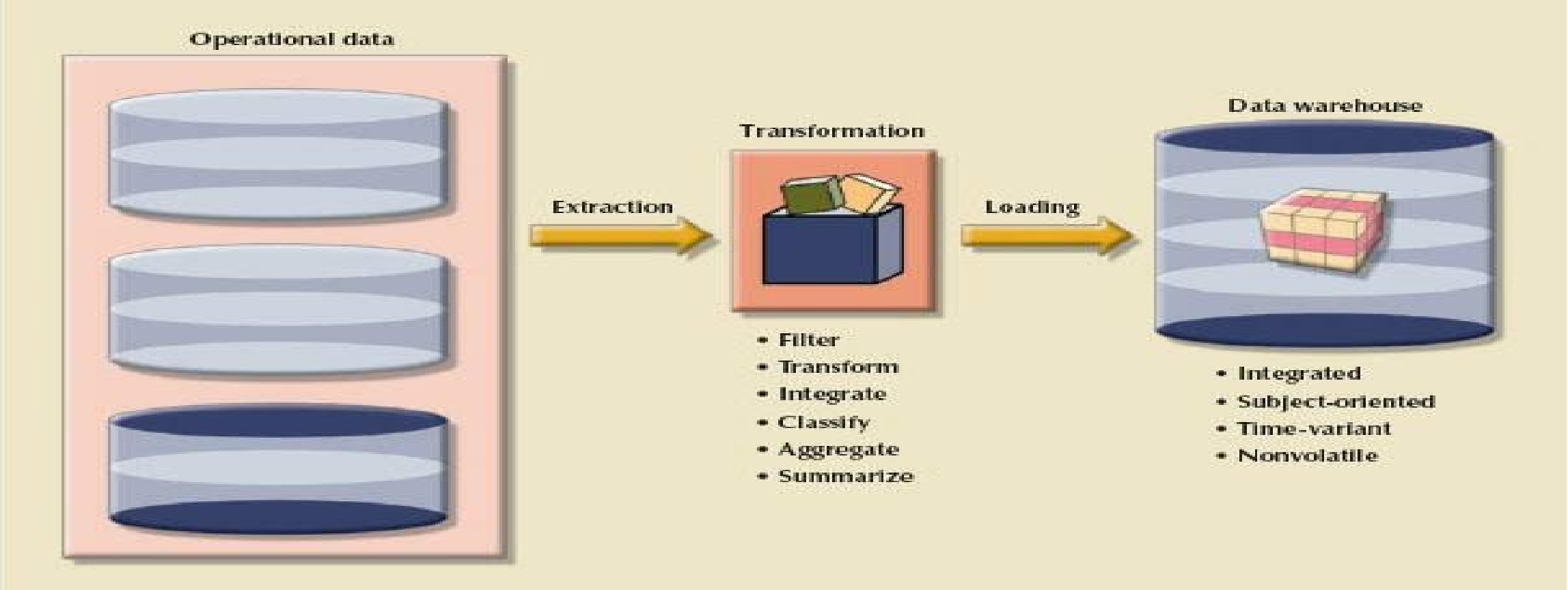
Data Warehouse: Is an integrated, subject-oriented, time-variant, nonvolatile collection of data that provides support for decision making.

CHARACTERISTIC	OPERATIONAL DATABASE DATA	DATA WAREHOUSE DATA
Integrated	Similar data can have different representations or meanings. For example, Social Security numbers may be stored as ###-##-#### or as #########, and a given condition may be labeled as T/F or 0/1 or Y/N. A sales value may be shown in thousands or in millions.	Provide a unified view of all data elements with a common definition and representation for all business units.
Subject-oriented	Data are stored with a functional, or process, orientation. For example, data may be stored for invoices, payments, and credit amounts.	Data are stored with a subject orientation that facilitates multiple views of the data and decision making. For example, sales may be recorded by product, division, manager, or region.
Time-variant	Data are recorded as current transactions. For example, the sales data may be the sale of a product on a given date, such as $342.78 on 12-MAY-2012.	Data are recorded with a historical perspective in mind. Therefore, a time dimension is added to facilitate data analysis and various time comparisons.
Nonvolatile	Data updates are frequent and common. For example, an inventory amount changes with each sale. Therefore, the data environment is fluid.	Data cannot be changed. Data are added only periodically from historical systems. Once the data are properly stored, no changes are allowed. Therefore, the data environment is relatively static.

Task: Discuss the tools used in business intelligence

Advanced Database Applications: Business Intelligence and Data Warehouse

Data Warehouse: Is an integrated, subject-oriented, time-variant, nonvolatile collection of data that provides support for decision making.



Task: Discuss the extraction, transformation, loading (ETL) process

Weaknesses of the Relational DBMSs.

Poor representation of 'real world' entities – normalization changes the real world model instead it fragments it

Semantic overloading- the relation is the only item for representation, where two entities are involved for many to many relationships we have to create another one which is not an entity.

Poor support for integrity and enterprise constraints - is no support for general constraints in the relational model, which again means they have to be built into the DBMS or the application.

Task: Discuss the weaknesses of RDB Model.

Weaknesses of the Relational DBMSs.

Homogeneous data structure- is too restrictive for many 'real world' objects that are complex. Relations have columns and rows that are fixed, same degree and columns have data from a domain.

Limited operations - has only a fixed set of operations, such as set and tuple-oriented operations, operations that are provided in the SQL specification. This is too restrictive to model the behavior of many 'real world' objects. For example, a GIS application typically uses points, lines, line groups, and polygons, and needs operations for distance, intersection, and containment.

Difficulty handling recursive queries - Atomicity of data means that repeating groups are not allowed in the relational model. As a result, it is extremely difficult to handle recursive queries, that is, queries about relationships that a relation has with itself (directly or indirectly).

Task: discuss the weaknesses of the RD model.

Weaknesses of the Relational DBMSs.

Impedance mismatch- **SQL is a declarative language that handles rows of data, whereas a high-level language such as 'C' is a procedural language that can handle only one row of data at a time. SQL and 3GLs use different models to represent data. For example, SQL provides the built-in data types Date and Interval, which are not available in traditional programming languages.**

Other problems with RDBMSs associated with concurrency, schema changes, and poor navigational access.

Task: discuss the weaknesses of the RD model.

Object-Oriented Concepts.

OODM (logical)- data model that captures the semantics of objects supported in object-oriented programming.

OODB - persistent and sharable collection of objects defined by an OODM. An object-oriented database can be collection of object-oriented programming and relational database. It is organized around objects rather than actions, and data rather than logic, eg, a multimedia record in a relational database can be a definable data object, as opposed to an alphanumeric value.

OODBMS- The manager of an OODB. It has:

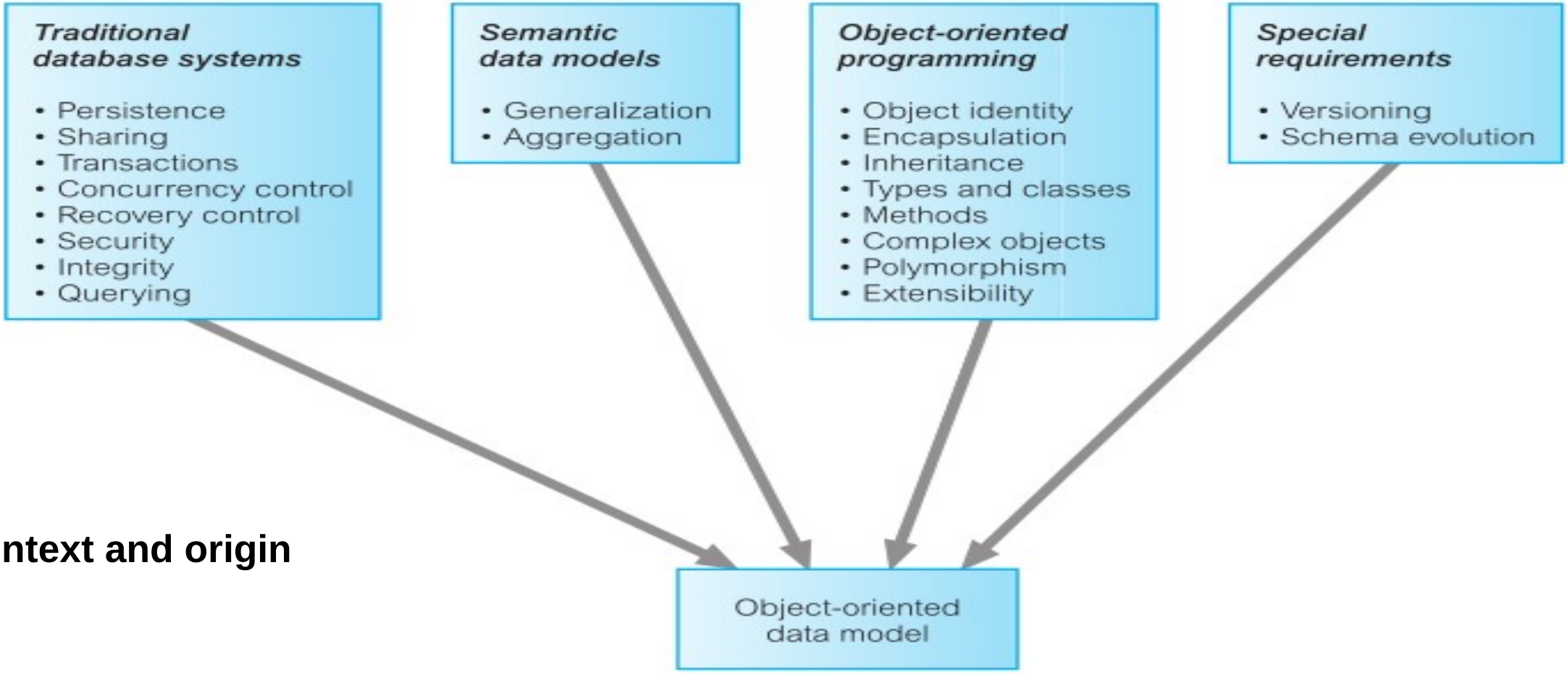
(1) object-orientation = abstract data types + inheritance + object identity;

(2) OODBMS = object-orientation + database capabilities.

Examples of OODBMS: Realm (database); ObjectDB; Wakanda; ObjectDatabase++

Task: Investigate the nature of the object oriented databases.

Object-Oriented Concepts.



The context and origin

Task: Investigate the object oriented database ideas.

Object-Oriented Concepts.

Object

- An object is a real-world entity in an environment and may be physical or imaginary.
- Each object is distinct and identifiable; and has properties with measurable or imaginable values;
- It also has behavior that represents externally visible activities performed by it in terms of changes in its state or values of its properties.
- Examples include: customer, a car, etc.; or an conceptual like a project, a process, etc.

Task: Investigate the object oriented object ideas.

Object-Oriented Concepts.

Class

- A set of objects with same properties and common behavior.
- Is the blueprint or description of the objects that can be created from it.
- A specific member of a class is called an instance.
- It consists of: attributes; and operations (methods) that portray the behavior of the objects of the class.
- Example: Circle
    - Data: x–coord, y–coord, radius
    - Methods: findArea(), findCircumference(), scaleRadius()

Task: Investigate the object oriented class ideas.

Object-Oriented Concepts.

Encapsulation

- Combing both attributes and methods the same class.
- Internal details of a class may be hidden from outside.
- The class components can be accessed from outside only through the interface provided by the class.

Data Hiding

- Class data (attributes) can be accessed only by its class methods and insulated from direct outside access. Sometimes called information hiding.

Task: Investigate the object oriented encapsulation ideas.

Object-Oriented Concepts.

Inheritance

- Allows specific sub-classes to continue with or inherit the properties and behaviour of their parent or super classes.
- Parent or super classes are called the base classes/ parent classes/ super-classes
- New classes are called the derived classes/ child classes/ sub-classes.
- The subclass can inherit or derive the properties and methods of the super-class(es) provided that the super-class allows so.
- the subclass may add its own attributes and methods and may modify any of the super-class methods.
- Inheritance allows an “is – a” relationship.
- Example: Mammal is a superclass of Human, Cat, Dog, Cow, etc., but each has its distinct characteristics.

Task: Investigate the other object oriented ideas.

Week 10 exercises

1)Discuss Advanced Database Applications.

2)Discuss the weaknesses of the Relational DBMSs.

3)Explain Object-Oriented Concepts.

Week 10 Session References

- **[Course Text] Carlos Coronel, Steven Morris, Peter Rob and Keeley Crockett Database Principles: Fundamentals of Design, Implementation, and Management, 14th Edition, 2022, ISBN-13978-0357673034.**

- **Thomas M. Connolly, Carolyn E. Begg (2021). Database Systems: A Practical Approach to Design, Implementation, and Management. Published by Pearson (July 14th 2021). ISBN-13: 9780137517053**

Thank You