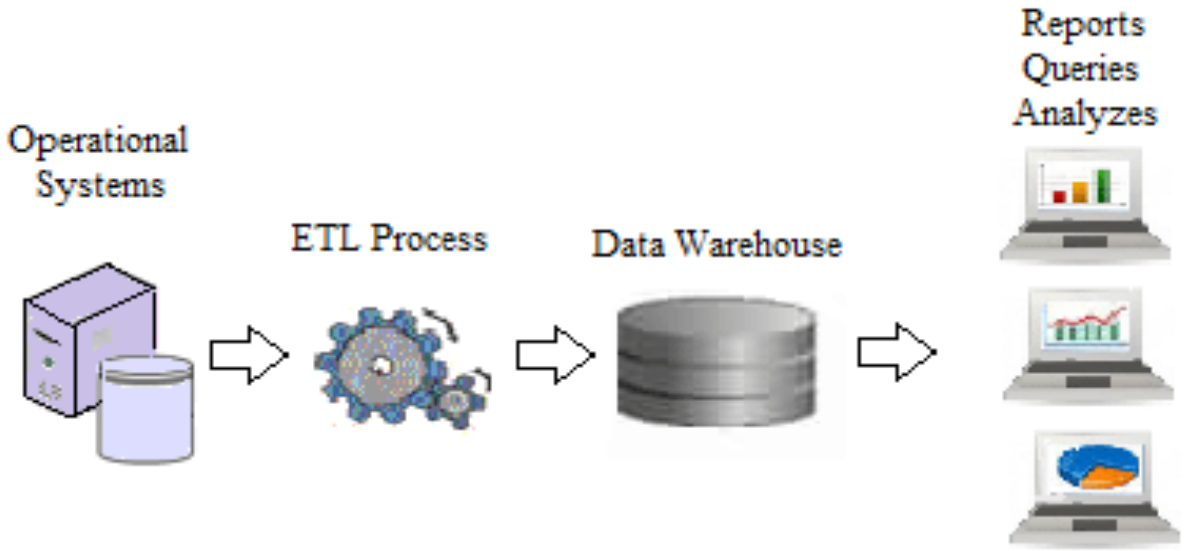
BI Architecture

- The technical structure for implementing BI systems.

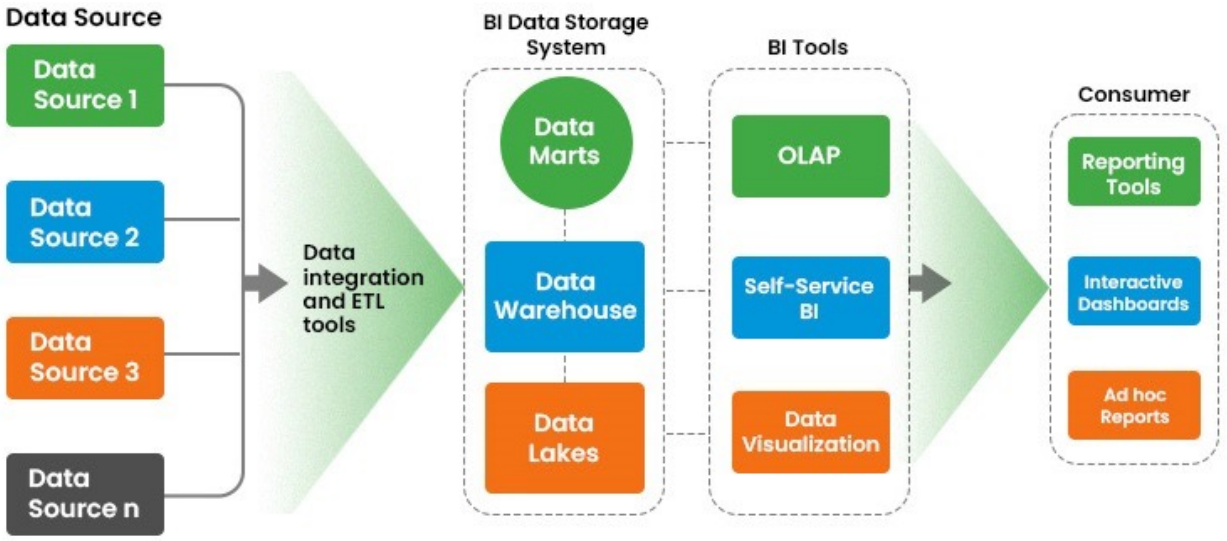
Traditional BI Architecture

- Source Layer → ETL Layer → Data Warehouse Layer → Analysis & Reporting Layer → User Interface



Modern BI Architecture

- **Real-time data streaming** (IoT, sensors).
- **Big Data technologies** (Hadoop, Spark).
- **Cloud-based storage & analytics**.
- **Self-service BI** (users access and analyze data without IT).
- **AI & ML integration** (advanced predictive models).



Types of BI Architectures

- **Centralized DW Architecture**
  - One central repository.
    - Pros: Consistency, easy maintenance.
    - Cons: Expensive, slower for diverse needs.
- **Decentralized (Data Mart) Architecture**
  - Multiple independent data marts.
    - Pros: Tailored for departments, quick implementation.
    - Cons: Data silos, redundancy.
- **Federated Architecture**
  - Combines centralized and decentralized approaches.
    - Pros: Flexibility, reduced redundancy.
    - Cons: Complex integration.

Types of BI Architectures

- **Cloud-based BI Architecture**
  - Data and tools hosted on cloud platforms.
    - Pros: Scalability, cost-effective, remote access.
    - Cons: Security and compliance concerns.

Types of BI Architectures

Centralized DW

Single central repository
+ Consistency, easy maintenance
- Expensive, less flexible

Decentralized Data Marts

Independent departmental marts
+ Tailored, quick setup
- Data silos, redundancy

Federated

Hybrid of central + marts
+ Flexible, less redundancy
- Complex integration

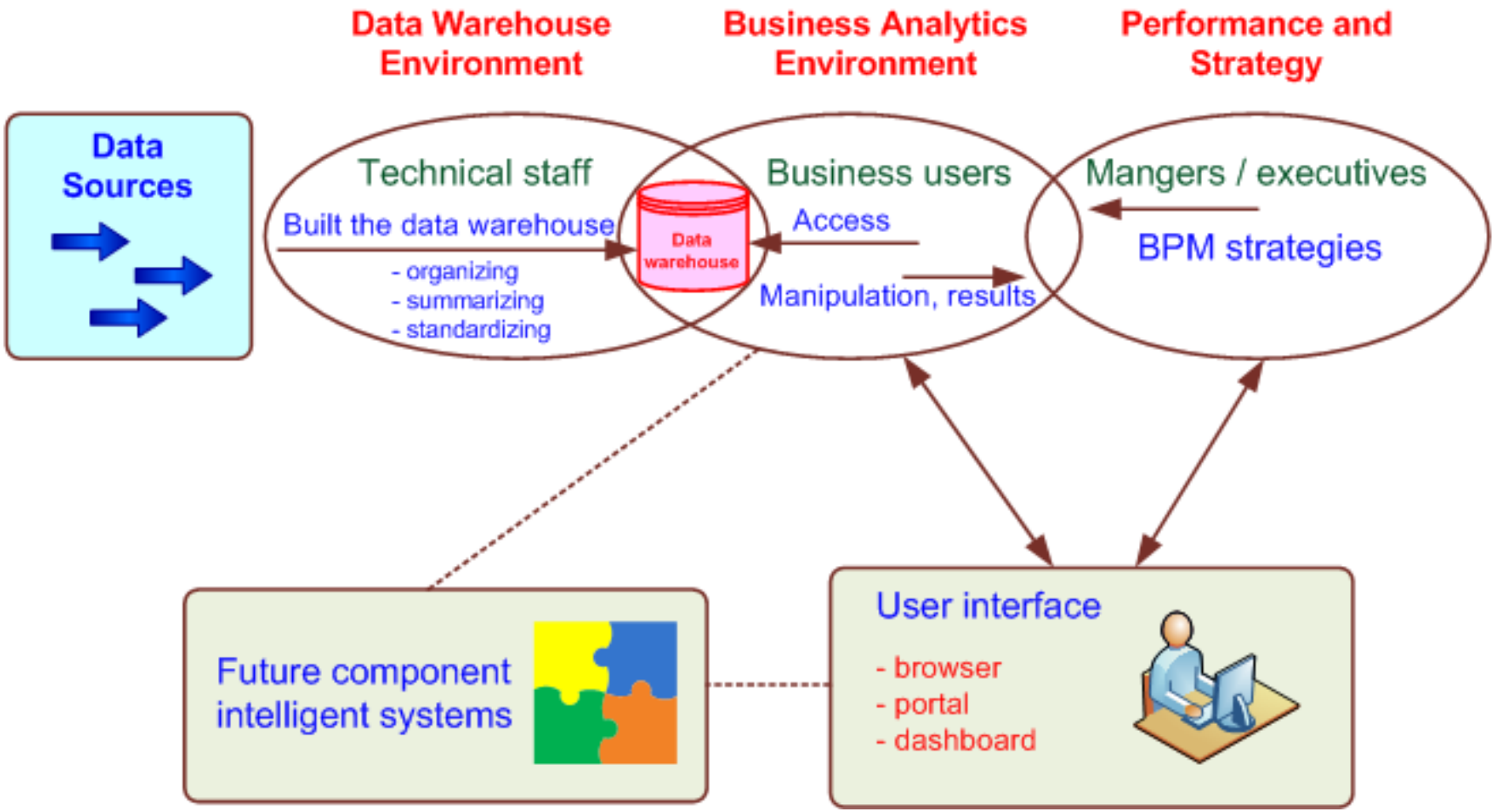
Cloud-based

Hosted on cloud platforms
+ Scalable, cost-effective
- Security concerns

BI Tools & Technologies

- **ETL Tools:** Talend, Informatica, SSIS.
- **Data Warehousing:** Amazon Redshift, Snowflake, Oracle DW.
- **OLAP Tools:** Microsoft SSAS, IBM Cognos.
- **Visualization Tools:** Power BI, Tableau, QlikView.
- **Big Data Tools:** Hadoop, Spark, Google BigQuery.

A High-level Architecture of BI



Components in a BI Architecture

- The data warehouse is the cornerstone of any medium-to-large BI system.
  - Originally, the data warehouse included only historical data that was organized and summarized, so end users could easily view or manipulate it.
  - Today, some data warehouses include access to current data as well, so they can provide real-time decision support (for details see Chapter 2).

Components in a BI Architecture

- The data warehouse is the cornerstone of any medium-to-large BI system.
  - Originally, the data warehouse included only historical data that was organized and summarized, so end users could easily view or manipulate it.
  - Today, some data warehouses include access to current data as well, so they can provide real-time decision support
- Business analytics are the tools that help users transform data into knowledge (e.g., queries, data/text mining tools, etc.).

- A data lake is a platform that stores data from various pipelines. The data doesn't undergo any processing and is stored as it is, structured or unstructured.
- It is simply a storage platform to aggregate all the data in its original form.
- Not designed for a specific purpose but acts as a large, single-point repository store data from which other platforms can collect and analyze data.

Characteristics of a data lake

- It collects and stores data in the raw format.
- It is a massive data storage unit for structured, unstructured, and semi-structured data.
- It supports multiple data source types with full and incremental retrieval capabilities.
- Provides complete metadata such as source, format, and all stored data types permissions.
- It acts as the storage platform for the data lifecycle management
- Highly scalable - expands data storage according to business data volume.

- A central storage platform that specifically stores structured data.
- Built to support business intelligence processes such as predictive analytics, forecasting, and machine learning. These require highly organized and structured data.
- The data in a data warehouse is prepared to be analytics-ready.
- Data can be used directly as a source for business intelligence tasks without optimizing data quality.

Characteristics of a data lake

- It stores processed and organized data
- Usually provides information on a single theme.
- The data is integrated from various pipelines to match a particular schema.
- The data must undergo the ETL process before being loaded in the data warehouse.
- Allows data to be analyzed based on time-variant metrics by storing data in specific intervals such as monthly or weekly.

DATA MART

- A data mart is a more focused version of a data warehouse that contains information about a specific business function.
- The data in the data mart is integrated to suit the unique requirements of a business unit. It serves a single team or department and has limited visibility.
- E.g, a data mart may have data generated to contain sales-related information for the business analysts in the audit unit.
- While the unit could access the same data from a warehouse, a data mart optimizes the process by providing only relevant data.

Characteristics of a data lake

- It stores data related to a business function.
- It is a subset of the larger data warehouse.
- Users have read-only access to information relevant to their unit.
- Stores high-quality data that is cleansed and optimized for a business function.
- It integrates and filters data from various sources to aggregate information for specific queries.
- It provides quick access and a faster query response time due to a smaller data pool.
- Highly secure and restricts data access by creating subject-wise segregation.
- No need for advanced technical knowledge to query the data in data marts.

Benefits of BI Framework & Architecture

- Improves decision-making quality.
- Provides real-time insights.
- Enhances operational efficiency.
- Enables performance measurement.
- Supports predictive and prescriptive analytics.

Challenges in BI Implementation

- Data quality issues.
- Integration of heterogeneous data sources.
- High cost and complexity.
- User adoption resistance.
- Security and privacy concerns.

Future Trends in BI Architectures

- Augmented Analytics (AI-driven BI).
- Real-time BI with streaming data.
- Natural Language Processing (NLP) for querying.
- Edge analytics (IoT-driven insights).
- Embedded BI within business applications.

BI frameworks and architectures provide the blueprint for organizing, storing, and analysing data to deliver actionable insights.

Traditional BI focused on centralized data warehouses, while modern BI emphasizes cloud, self-service, real-time analytics, and AI-driven insights.