

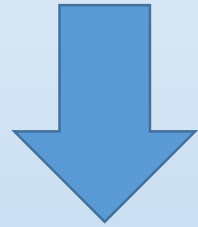
# Corso di Laurea Magistrale in Economia

## Data Science A.A. 2018/2019

### Lez. 2 – Business Intelligence

# Dati

Tecnologie di memorizzazione a basso costo + connettività



## **Accesso a grandi moli di dati**

- Transazioni finanziarie, commerciali e amministrative
- Navigazione su web
- Email, testi e ipertesti
- ...

# Dati, informazioni e conoscenza

È possibile trasformare i dati in informazioni e in conoscenze utilizzabili dai *decision maker*?

## *Business Intelligence*

Insieme di modelli matematici e metodologie di analisi che esplorano i dati per ricavare informazioni e conoscenze per i processi decisionali

# Decisioni

I *decision maker* elaborano le proprie decisioni utilizzando «metodologie» semplici e intuitive, basate su esperienze passate, conoscenza del contesto, informazioni disponibili.

- Indole statica, incapace ad adattarsi a condizioni mutevoli
- Necessità di utilizzare metodologie analitiche e modelli matematici per definire decisioni **efficaci** e **tempestive**

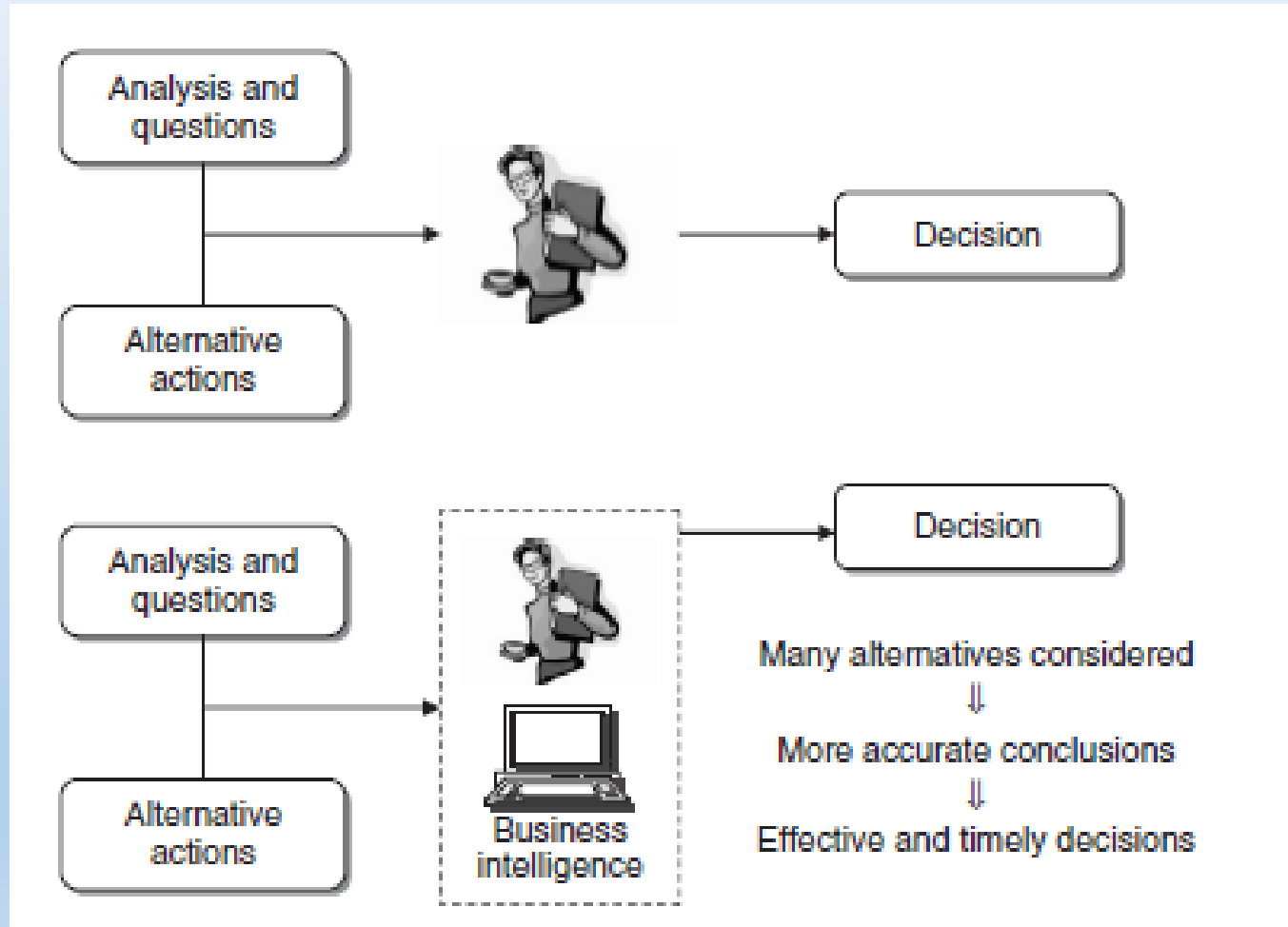
# Esempio

**Example 1.1 – Retention in the mobile phone industry.** The marketing manager of a mobile phone company realizes that a large number of customers are discontinuing their service, leaving her company in favor of some competing provider. As can be imagined, low customer loyalty, also known as customer *attrition* or *churn*, is a critical factor for many companies operating in service industries. Suppose that the marketing manager can rely on a budget adequate to pursue a customer retention campaign aimed at 2000 individuals out of a total customer base of 2 million people. Hence, the question naturally arises of how she should go about choosing those customers to be contacted so as to optimize the effectiveness of the campaign. In other words, how can the probability that each single customer will discontinue the service be estimated so as to target the best group of customers and thus reduce churning and maximize customer retention? By knowing these probabilities, the target group can be chosen as the 2000 people having the highest churn likelihood among the customers of high business value. Without the support of advanced mathematical models and data mining techniques, described in Chapter 5, it would be arduous to derive a reliable estimate of the churn probability and to determine the best recipients of a specific marketing campaign.

# Esempio

**Example 1.2 – Logistics planning.** The logistics manager of a manufacturing company wishes to develop a medium-term logistic-production plan. This is a decision-making process of high complexity which includes, among other choices, the allocation of the demand originating from different market areas to the production sites, the procurement of raw materials and purchased parts from suppliers, the production planning of the plants and the distribution of end products to market areas. In a typical manufacturing company this could well entail tens of facilities, hundreds of suppliers, and thousands of finished goods and components, over a time span of one year divided into weeks. The magnitude and complexity of the problem suggest that advanced optimization models are required to devise the best logistic plan. As we will see in Chapter 14, optimization models allow highly complex and large-scale problems to be tackled successfully within a business intelligence framework.

# Decisioni e *Business Intelligence*



# Dati, informazioni e conoscenza

**Dati:** rappresentano una codifica strutturata delle singole entità primarie e delle interazioni tra queste. Ad es., clienti, prodotti, vendite.

**Informazioni:** sono il risultato di operazioni di estrazione ed elaborazione sui dati, hanno un significato per chi le utilizza in uno specifico contesto. Ad es., numero di scontrini di importo  $> \text{€ } 100$  in una settimana.

**Conoscenza:** le informazioni diventano conoscenza se, inserite in contesto e arricchite dall'esperienza e dal *know-how*, vengono utilizzate per prendere delle decisioni e per implementare le azioni conseguenti. Ad es., l'analisi delle vendite in una zona con riduzione del fatturato  $\rightarrow$  nuovo concorrente. Decisione: consegna a domicilio gratuita.

La conoscenza può essere estratta dai dati:

- in modo passivo (criteri suggeriti dal *decision maker*)
- in modo attivo (modelli matematici di apprendimento automatico e ottimizzazione)



# Ruolo dei modelli matematici

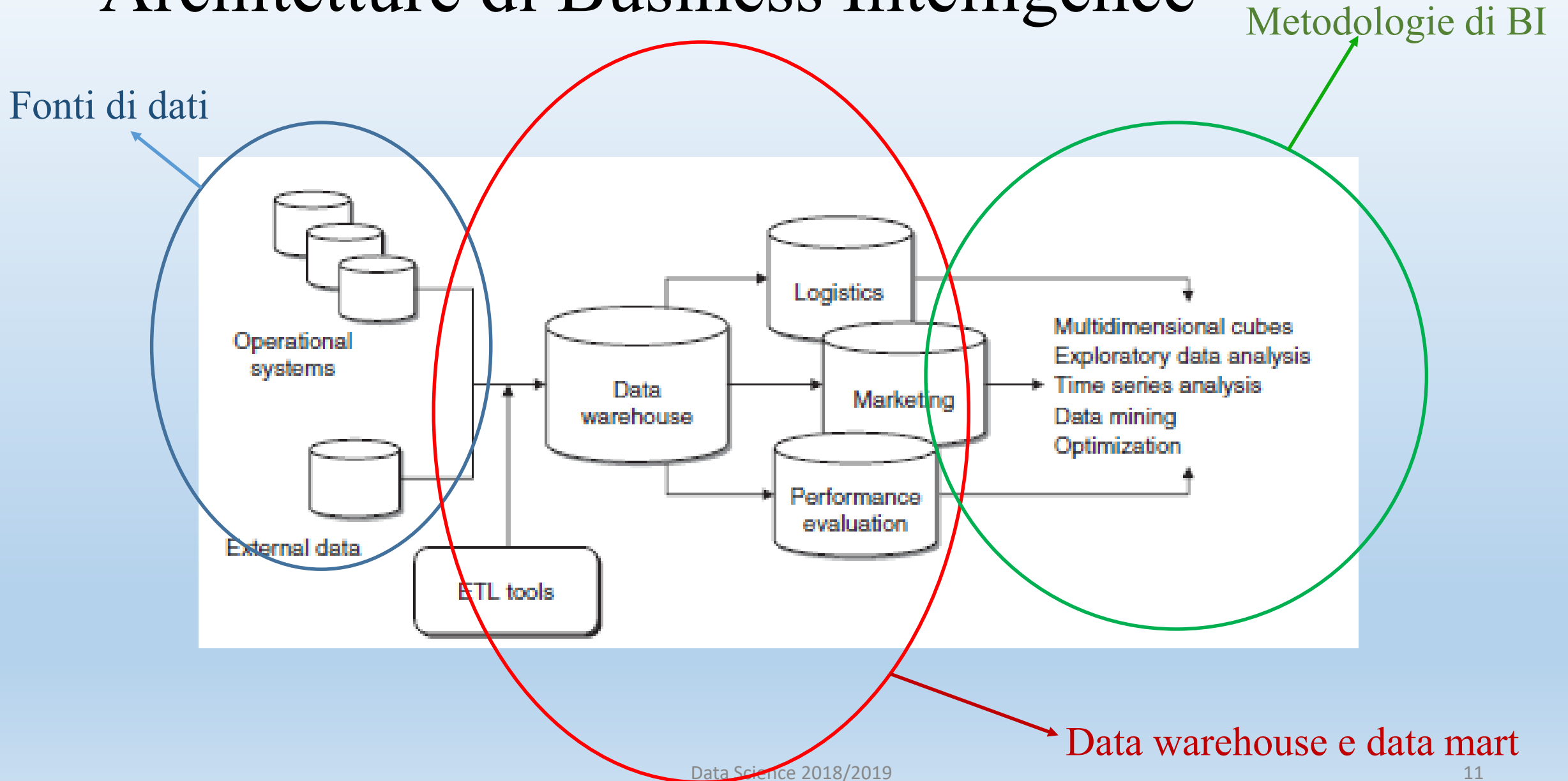
Approccio delle analisi di *Business Intelligence*

- Individuazione obiettivi ed identificazione indicatori di prestazione per la valutazione delle alternative
- Sviluppo di modelli matematici (rappresentazione delle relazioni tra le variabili ed i parametri e le metriche di valutazione)
- Analisi *what-if* per la valutazione degli effetti sulle prestazioni di variazioni delle variabili e/o dei parametri

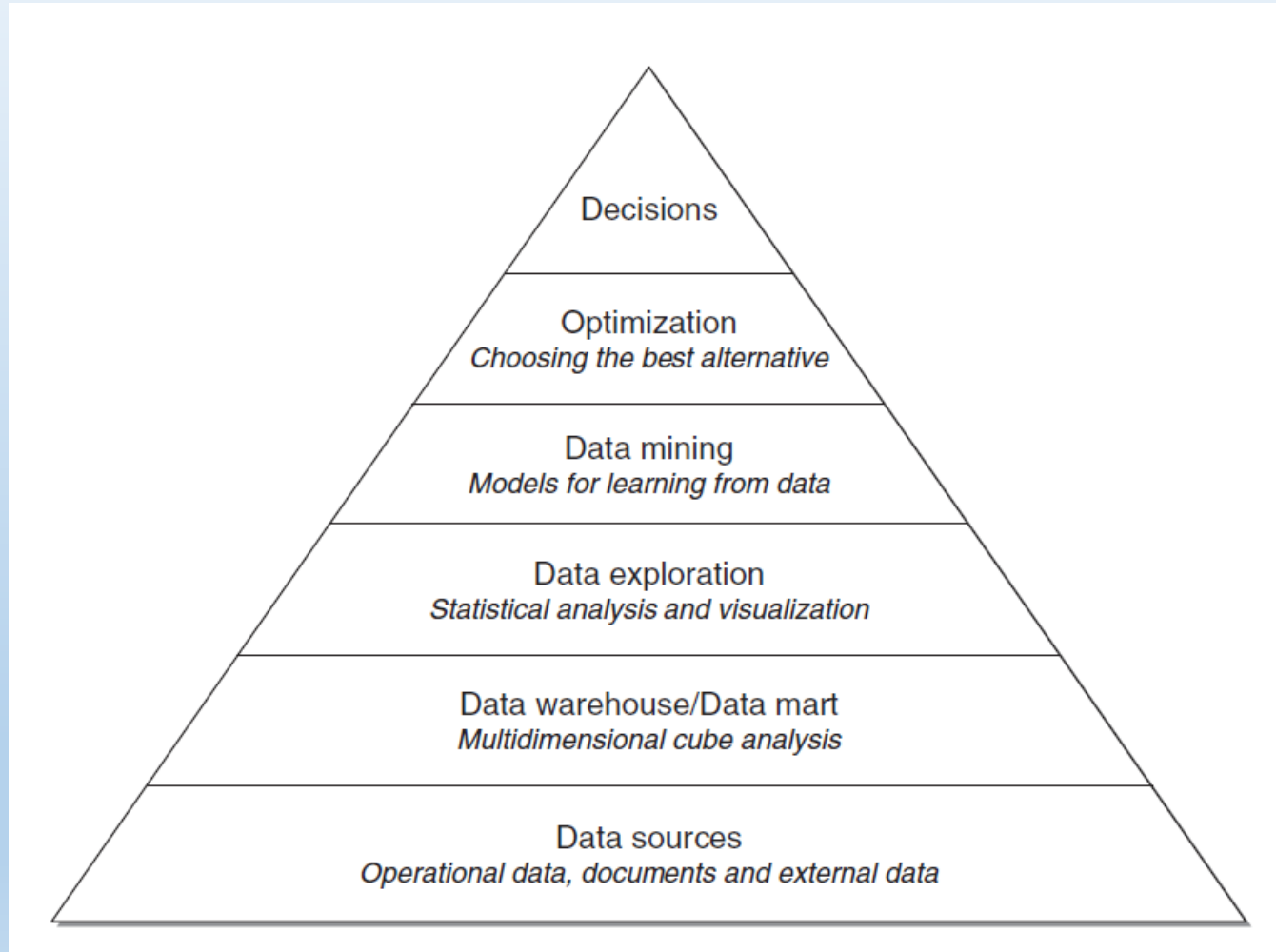
# Vantaggi

- Aumento efficacia del processo decisionale
- Maggiore attenzione agli aspetti rilevanti
- Conservazione e trasferimento conoscenza all'interno dell'organizzazione
- Possibilità di riutilizzare la metodologia definita per uno specifico problema anche in altre situazioni simili

# Architetture di Business Intelligence



# Ambiente di Business Intelligence



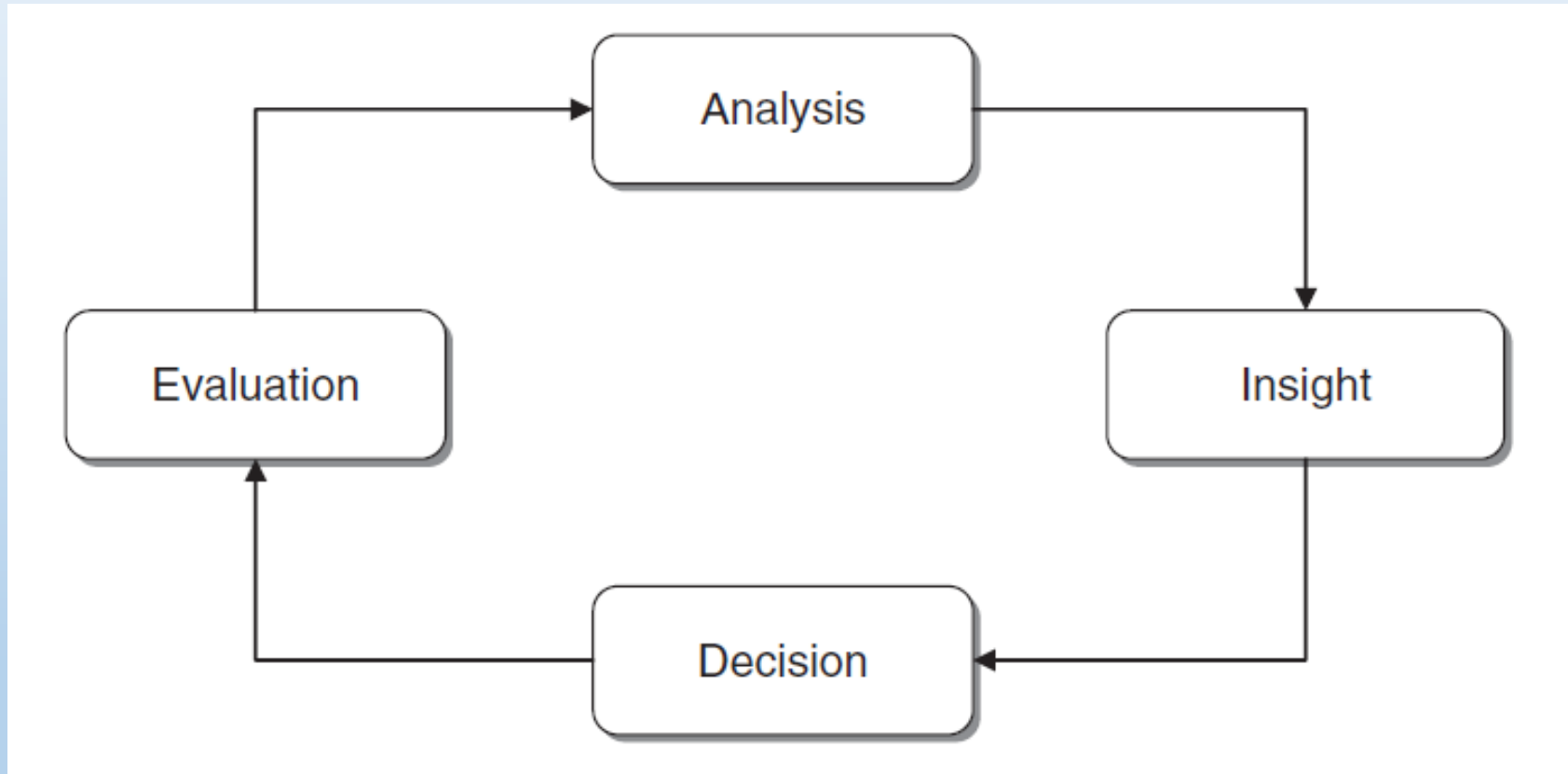
# Ambiente di Business Intelligence

- **Fonti di dati:** raccolta e integrazioni di dati da diverse sorgenti eterogenee
- **Data warehouse e data mart:** mediante strumenti di ETL (*extract, transform, load*) i dati vengono immagazzinati nei DB preposti alle analisi di BI
- **Esplorazione dei dati:** strumenti per le analisi «passive» dei dati (metodi di analisi statistica, sistemi di reporting). Il decisore deve formulare preventivamente un'ipotesi che potrà essere verificata grazie a tali strumenti.

# Ambiente di Business Intelligence

- **Modelli di apprendimento:** metodologie «attive» per l'estrazione di informazioni e conoscenza a partire dai dati (*data mining*). Estendono le conoscenze del decisore senza la necessità di ipotesi preliminari.
- **Modelli di ottimizzazione:** consentono di individuare la decisione migliore tra un insieme molto ampio o infinito di alternative.
- **Decisioni:** l'identificazione della decisione (azione) definitiva rimane di pertinenza del decisore, che può utilizzare notizie informali e non strutturate per «adattare» le soluzioni suggerite.

# Ciclo delle analisi di Business Intelligence



# Ciclo delle analisi di Business Intelligence

- **Analisi**
  - Identificazione e formulazione del problema
  - Elaborazione modello e selezione dei fattori più rilevanti
  - Rapido ciclo di domande/risposte
- **Comprensione**
  - Maggiore conoscenza delle dinamiche del problema
  - Informazioni → conoscenze
  - Utilizzo sia di valutazioni soggettive che di modelli di apprendimento
- **Decisione**
  - Conoscenze → decisioni → azioni
  - Valutazione e aggiustamento delle soluzioni suggerite da strumenti di BI
- **Valutazione**
  - Misura delle prestazioni attraverso indicatori oggettivi e pertinenti



# Fattori abilitanti

- **Tecnologie**

- Potenza di calcolo → algoritmi più sofisticati, visualizzazione grafica
- Capacità dispositivi di archiviazione → moli di dati sempre maggiori
- Connettività → accesso a fonti di dati, diffusione di informazioni e conoscenze
- Facile integrazione tra hw e sw → portabilità degli strumenti di BI

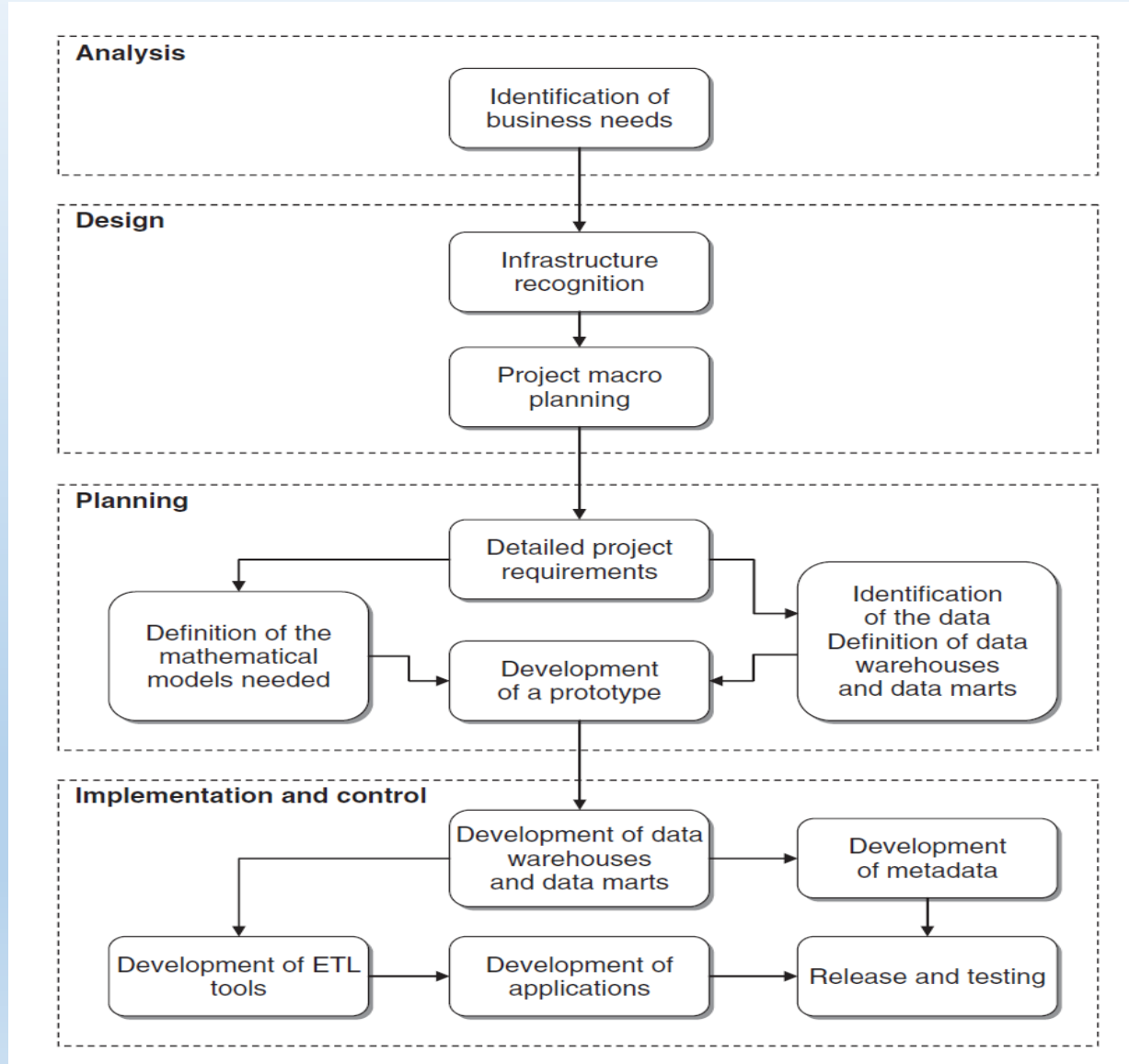
- **Metodologie analitiche**

- Modelli di apprendimento evoluti
- Modelli di ottimizzazione efficaci
- Algoritmi efficienti

- **Risorse umane**

- Capacità di assimilare informazioni e definizione di decisioni/azioni
- Competenze necessarie per l'efficace utilizzo di strumenti di BI

# Cicli di sviluppo ambiente di BI



# Fasi per lo sviluppo di un ambiente di BI

- **Giustificazione**
  - Valutazione delle esigenze (anche mediante interviste/questionari)
  - Identificazione obiettivi generali e priorità
  - Analisi costi/benefici
- **Pianificazione**
  - Progettazione di massima che tiene conto sia degli sviluppi immediati che delle evoluzioni future
  - Ricognizione delle infrastrutture esistenti
  - Esame dei processi decisionali da supportare mediante strumenti di BI
  - Pianificazione del progetto, identificazione fasi di sviluppo, tempi, costi, ruoli e risorse, con tecniche di *Project Management*

# Fasi per lo sviluppo di un ambiente di BI

- **Progettazione**

- Identificazione e descrizione di dettaglio delle funzionalità
- Ricognizione dei dati esistenti e/o reperibili all'esterno
- Definizione architettura (Data Warehouse/Data Mart)
- Definizione modelli matematici e verifica capacità algoritmi
- Realizzazione prototipo
- Verifica della rispondenza alle esigenze

- **Realizzazione e collaudo**

- Sviluppo Data Warehouse/Data Mart
- Sviluppo archivio metadati
- Sviluppo procedure di ETL
- Realizzazione applicazioni di BI (modelli, algoritmi)
- Collaudo del sistema