

**Corso di Laurea Magistrale in Economia**

**Corso di Laurea in Ingegneria Industriale**

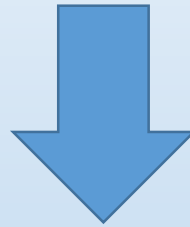
# Data Science and Decision Support Systems

A.A. 2019/2020

Lez. 2 – Sistemi di 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**

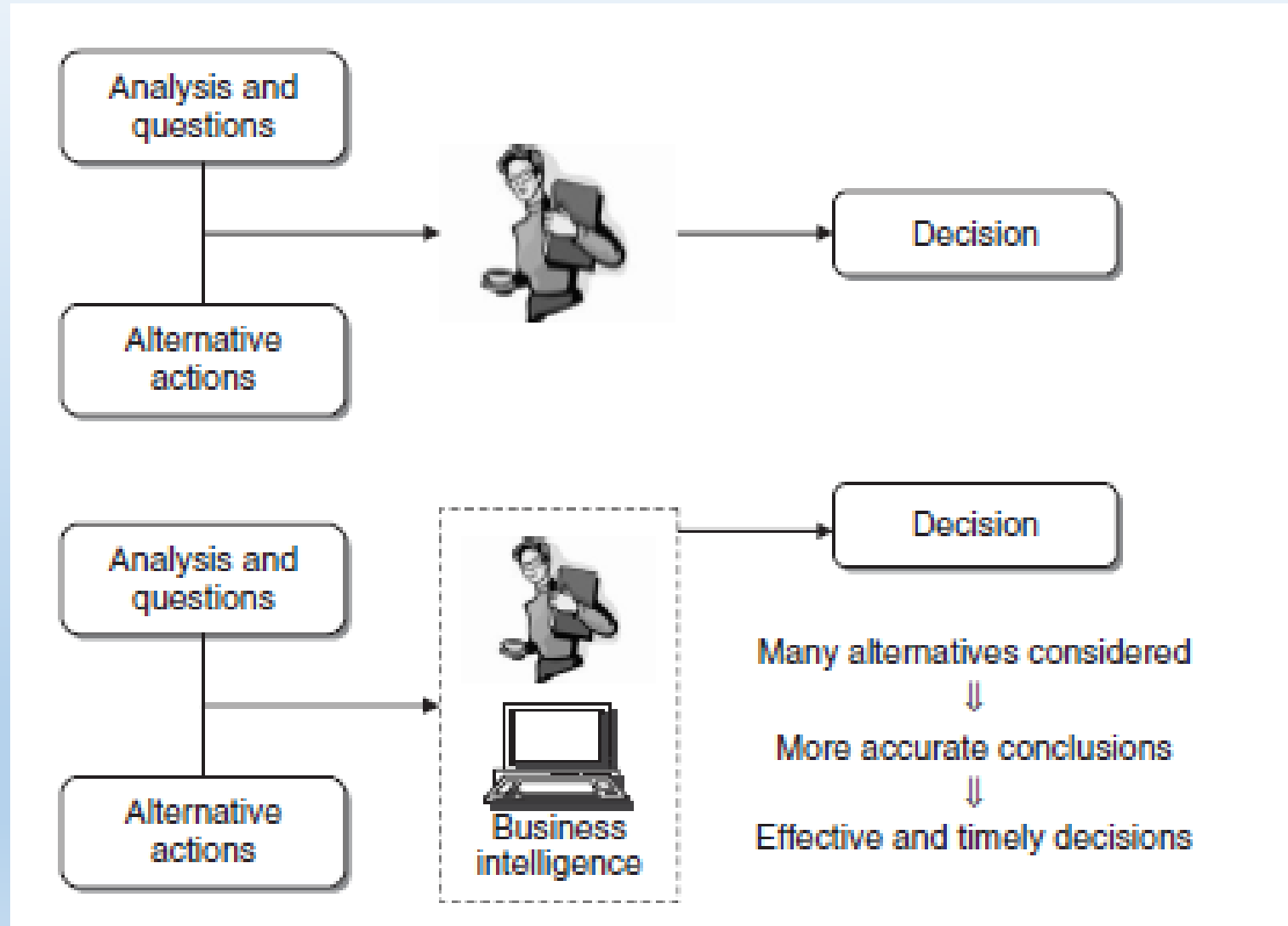
# Esempio

- Il marketing manager di un'azienda telefonica si rende conto che un gran numero di clienti sta lasciando la compagnia verso un altro operatore. Come noto, la scarsa «fedeltà» dei clienti (*churn*) è un fattore critico per diverse aziende del settore servizi. Il manager riesce ad ottenere risorse per svolgere una campagna mirata nei confronti di 2000 utenti (su una base di circa 2 milioni). La questione è: come scegliere gli utenti cui destinare le promozioni? In sostanza, come si può stimare la probabilità di abbandono, in modo da indirizzare le azioni verso gli utenti più a rischio?
- Senza l'ausilio di modelli matematici e tecniche di data mining l'operazione è praticamente impossibile

# Esempio

- Il *logistic manager* di una azienda manifatturiera vuole predisporre un piano di sviluppo a medio-lungo termine.
- Tale processo decisionale è particolarmente complesso in quanto include:
  - L'allocazione della domanda delle diverse aree ai siti di produzione
  - L'approvvigionamento di materie prime dai fornitori
  - La pianificazione della produzione di ciascun impianto
  - La distribuzione dei prodotti finiti nelle diverse aree
  - ...
- In una azienda di medie dimensioni possono esserci decine di siti logistici, centinaia di fornitori e migliaia di prodotti finiti e componenti, e il processo decisionale può abbracciare periodi anche di svariate settimane.
- Per avere un supporto decisionale efficace è indispensabile avvalersi di modelli di ottimizzazione avanzati per problemi a grandi dimensioni

# 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

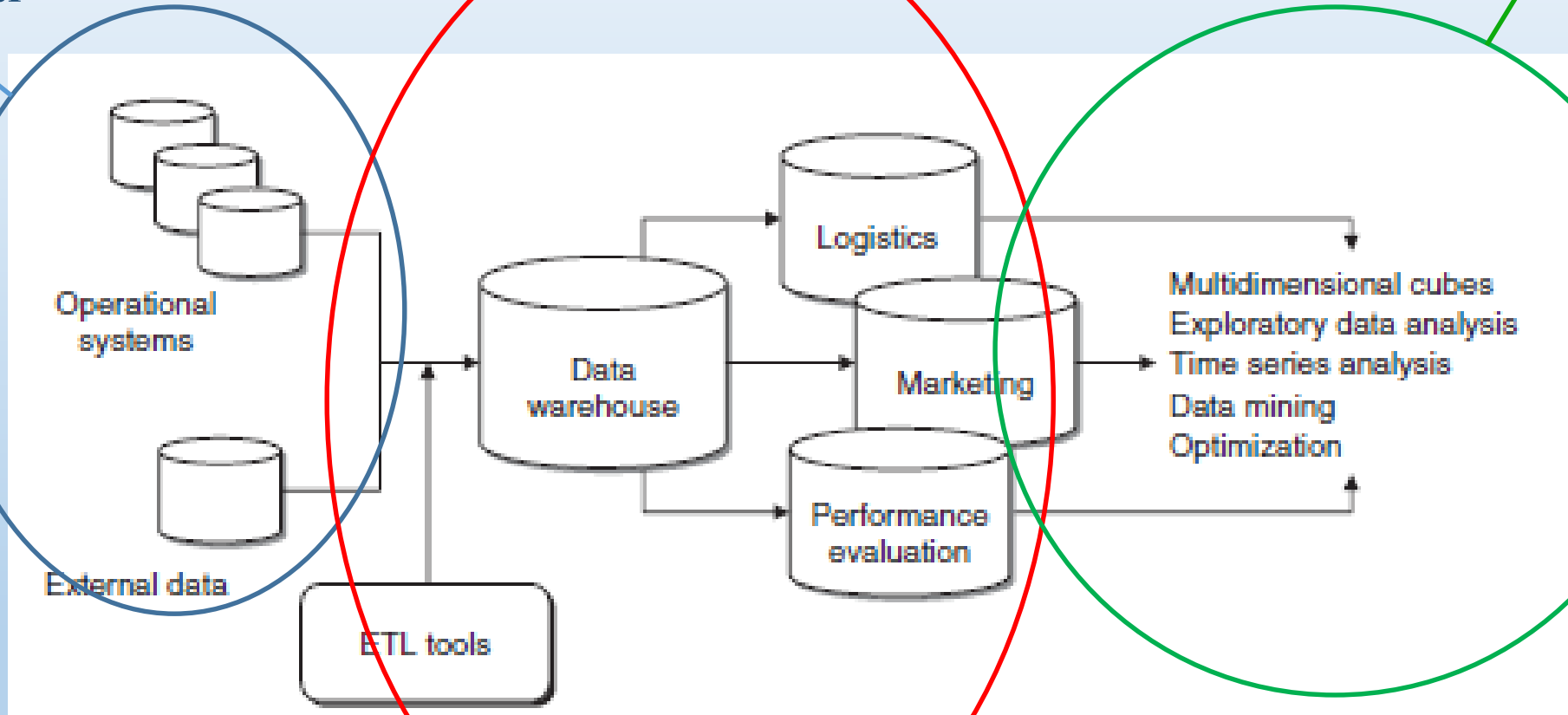
# Vantaggi

- 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

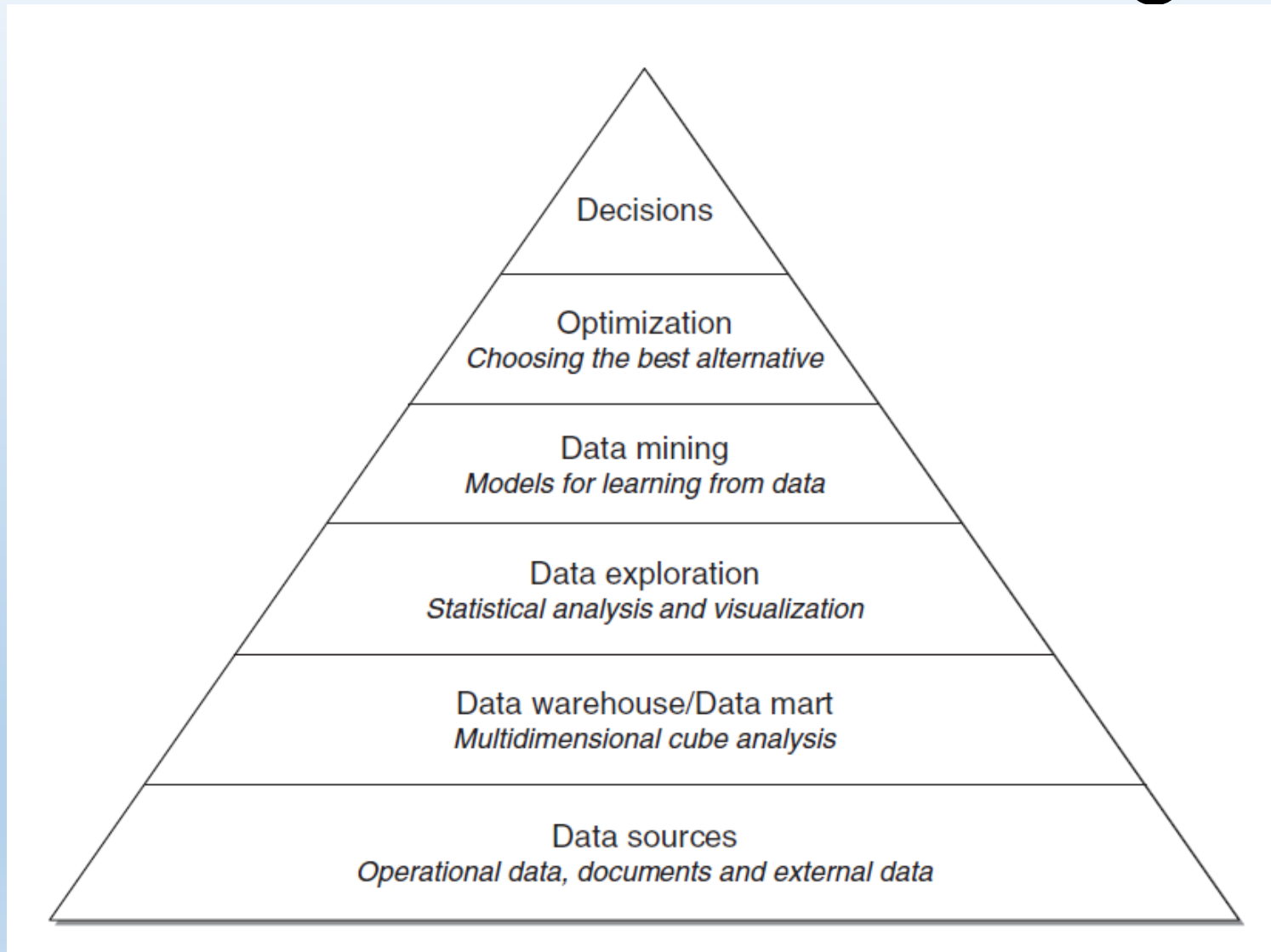
Metodologie di BI

Fonti di dati



Data warehouse e data mart

# 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.

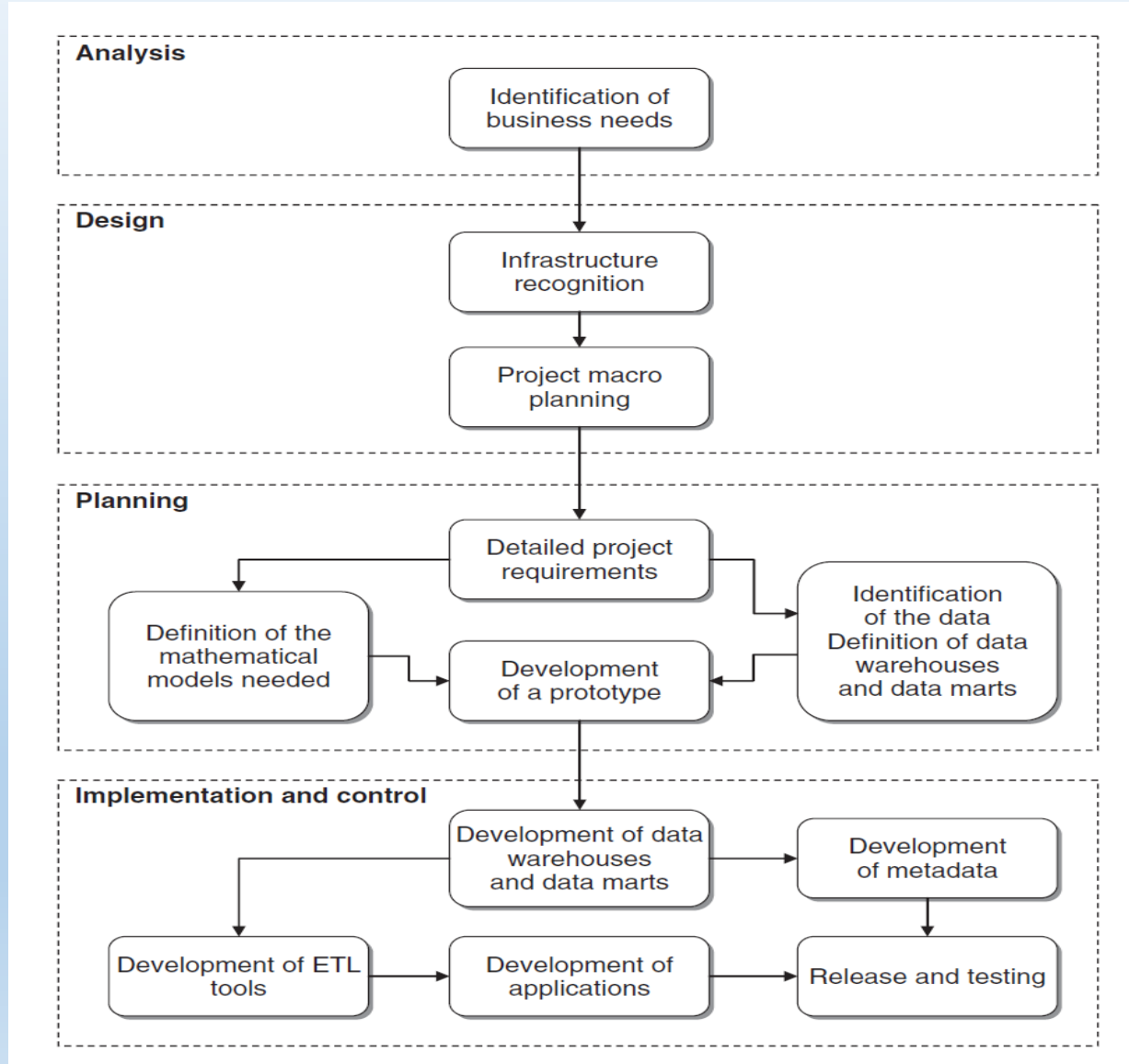
# Ambiente di Business Intelligence

- **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.

# 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