

Basi di Dati

Federico Matteoni

A.A. 2019/20

Indice

1	Introduzione	5
2	Base di Dati	7
2.1	Costruzione di una Base di Dati	7
2.1.1	Figure Coinvolte	7
2.1.2	Sistemi Informativi	7
2.1.3	Sistemi Informatici	8
2.1.4	Classificazione dei sistemi informatici	9
2.1.5	Requisiti per l'Analisi dei Dati	10

Capitolo 1

Introduzione

Obiettivi del corso Modelli dei dati, linguaggi e sistemi per lo sviluppo di applicazioni che prevedono l'uso di grandi quantità di dati permanenti organizzati in **basi di dati**.

Testo di Riferimento *Fondamenti di Basi di Dati*, A. Albano, G. Ghelli e R. Orsini, Zanichelli. Scaricabile liberamente da fondamentidibasididati.it

Esame Scritto + orale

Terminologia

Base di dati: tecnologia di base, gestione delle attività quotidiane dell'organizzazione e **tema di questo corso**

Data Warehouse, Data Lake, Big Data, Data Science: termini che hanno a che vedere con l'**analisi dei dati** e che non rientrano nei temi trattati nel corso.

Capitolo 2

Base di Dati

Cos'è una base di dati? Una **base di dati** è un **insieme organizzato di dati** usati per il supporto allo svolgimento di una attività (di un ente, azienda, ufficio, persona...)

Qualche esempio

Titolo	Codice	Materie	Syllabus
Basi di Dati	AA024	Progettazione e interrogazione...	
Reti di Calcolatori	AA019	Realizzazione e uso di reti, protocollo TCP...	

Materia	AA	Corsi	Titolare
		Semestre	
AA024	2007	1	Albano
AA024	2007	1	Ghelli
AA019	2007	1	Brogi

2.1 Costruzione di una Base di Dati

2.1.1 Figure Coinvolte

Committente

Dirigente
Operatore

Fornitore

Direttore del progetto
Analista
Progettista del DB
Programmatore di applicazioni che usano il DB

Manutenzione e messa a punto del DB – Gestione del DBMS

Amministratore del DBMS

2.1.2 Sistemi Informativi

Definizione Un **sistema informativo** di un'organizzazione è una **combinazione di risorse, umane e materiali, e di procedure** organizzate per raccolta, archiviazione, elaborazione e scambio **delle informazioni** necessarie alle attività:

Operative (informazioni di servizio)

Programmazione e controllo (informazioni di gestione)

Pianificazione strategica (informazioni di governo)

Esempi di sistemi informativi Un comune

Gestione servizi demografici (anagrafe, stato civile, servizio elettorale e vaccinale) e della rete viaria

Gestione attività finanziaria secondo la normativa vigente

Gestione del personale per il calcolo della retribuzione in base al tipo di normativa contrattuale

Gestione dei servizi amministrativi e sanitari delle USL

Gestione della cartografia generale e tematica del territorio

Sistema informativo nelle organizzazioni



2.1.3 Sistemi Informatici

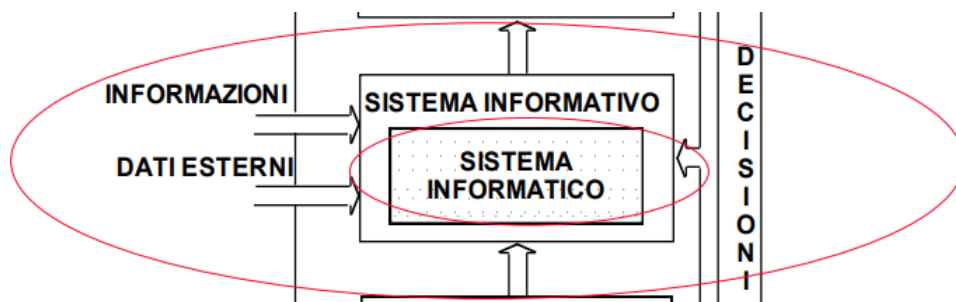
Sistema Informativo Automatizzato Quella parte del sistema informativo in cui le informazioni sono raccolte, elaborate, archiviate e scambiate usando un **sistema informatico**.

Sistema Informatico Insieme delle tecnologie informatiche e della comunicazione (ICT, Information and Communication Technologies) a supporto delle attività di un'organizzazione.

Terminologia

Sistema informativo → Sistema informativo automatizzato

Sistema informativo automatizzato → Sistema informatico



2.1.4 Classificazione dei sistemi informatici

Sistemi Informatici Operativi → Sistemi Informatici Direzionali

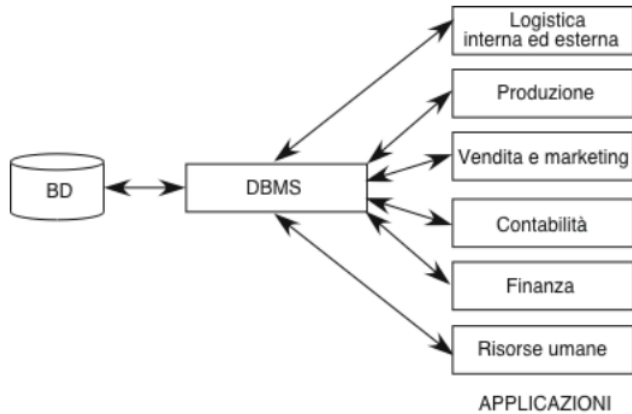
Sistemi Informatici Operativi I dati sono organizzati in DB. Le applicazioni si usano per svolgere le classiche attività strutturate e ripetitive dell'azione nelle aree amministrativa e finanziaria: vendite, risorse umane, produzione. . .

Alcune sigle:

DP Data Processing

EDP Electronic Data Processing

TPS Transaction Processing Systems



DBMS Le caratteristiche del DB sono **garantite da un sistema per la gestione della base di dati** (DBMS, Data Base Management System) che ha il controllo dei dati e li rende accessibili agli utenti autorizzati.

OLTP On-Line Transaction Processing, modo d'uso principale dei DBMS. Tradizionale elaborazione di transazioni, che realizzano processi operativi per il funzionamento di organizzazioni:

Operazioni predefinite e relativamente semplici

Ogni operazione coinvolge *pochi* dati

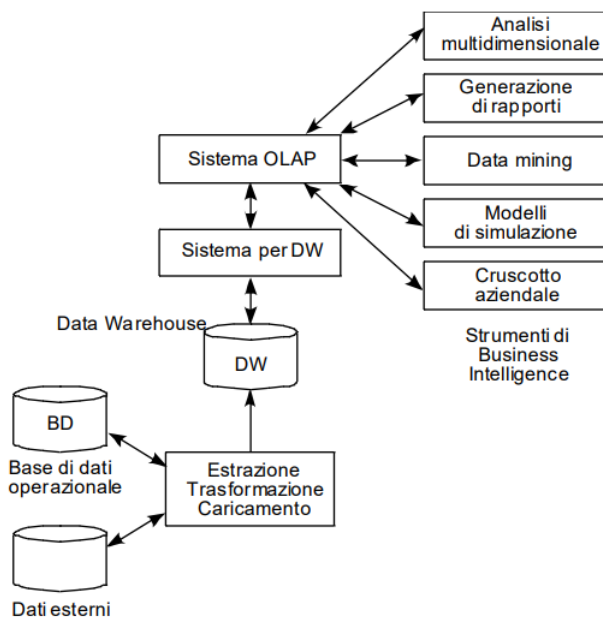
Dati di dettaglio, aggiornati

Sistemi Informatici Direzionali I dati sono organizzati in data warehouse (DW) e gestiti ad un opportuno sistema. Le applicazioni, dette di **business intelligence**, sono strumenti di supporto ai processi di controllo delle prestazioni aziendali e di decisione manageriale. Terminologia:

MIS Management Information Systems

DSS Decision Support Systems, data-based o model-based

EIS Executive Information System



OLAP On-Line Analytical Processing modo d'uso principale dei DW. Analisi dei dati di supporto alle decisioni:

Operazioni complesse e casuali

Ogni operazione può coinvolgere *molte* dati

Dati aggregati, storici, anche non attualissimi

Differenze tra OLTP e OLAP

	OLTP	OLAP
Scopi	Supporto operatività	Supporto decisioni
Utenti	Molti, esecutivi	Pochi, dirigenti e analisti
Dati	Analitici, relazionali	Sintetici, multidimensionali
Usi	Noti a priori	Poco prevedibili
Quantità di dati per attività	Bassa (decine)	Alta (milioni)
Orientamento	Applicazione	Soggetto
Aggiornamenti	Frequenti	Rari
Visione dei dati	Corrente	Storica
Ottimizzati per	Transazioni	Analisi

2.1.5 Requisiti per l'Analisi dei Dati

Aggregati Non interessa **un** dato, ma la **somma**, la **media**, il **minimo**/**massimo** di una misura. . .

Multidimensionale Interessa **incrociare le informazioni**, per analizzarle da punti di vista diversi e valutare i risultati del business per intervenire sui problemi critici o per cogliere nuove opportunità

Diversi livelli di dettaglio Per esempio, una volta scoperto un calo delle vendite in un determinato periodo in una specifica regione, si passa ad un'analisi dettagliata nell'area di interesse per cercare di scoprirne le cause (dimensioni con **gerarchie**)

2.1.6 Big Data

Ampio Big data è un termine ampio riferito a situazioni in cui l'approccio "schema-first" tipico di DB e DW risulta troppo restrittivo o troppo lento.

3 V Volume, Varietà, Velocità

I Big Data sono in genere associati a sistemi NoSQL, machine learning e approcci Data Lake.