



UNIVERSITA DEGLI STUDI DI FOGGIA

Dipartimento di Agraria

Cdl in Ingegneria dei Sistemi Logistici per l'Agroalimentare

Corso integrato di Sistemi di Elaborazione

Modulo I

Prof. Crescenzo Gallo

crescenzo.gallo@unifg.it

Basi di dati: introduzione

Introduzione

- Gestione delle informazioni
- Basi di dati / DBMS
- Modello dei dati
- Indipendenza dei dati
- Accesso ai dati
- Vantaggi e svantaggi dei DBMS

Gestione delle informazioni

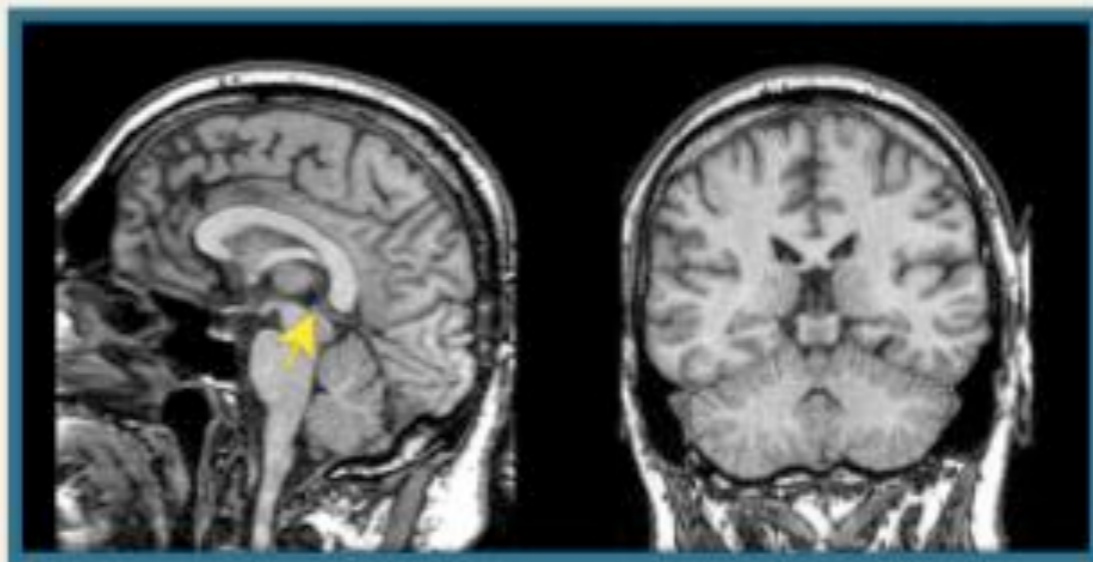
Gestione delle informazioni

Le informazioni sono registrate e scambiate in forme diverse:



Gestione delle informazioni

Nel tempo, sono state introdotte forme di organizzazione e codifica delle informazioni:



Gestione delle informazioni

- Nei sistemi informatici, le informazioni sono rappresentate per mezzo di dati.
- I dati sono simboli grezzi che devono essere interpretati e correlati per fornire informazioni.

Esempio

- *dati*: "Mario Rossi", 424242;
- *informazione*: la persona registrata sulla propria agenda telefonica.



Gestione delle informazioni

- I dati sono molto più stabili nel tempo dei processi che li gestiscono.
 - *Esempio:*
 - *i dati relativi alle applicazioni bancarie hanno una struttura invariata da decenni;*
 - *le procedure che li gestiscono variano di anno in anno.*
- I dati sono una risorsa importante della organizzazione che li gestisce.

Basi di dati / DBMS

Basi di dati

- Definizione generale:
una base di dati è una collezione di dati che rappresenta le informazioni di interesse per un sistema informativo.
- Definizione "tecnica":
una base di dati è una collezione di dati gestita da un DBMS.

DBMS

Un sistema per la gestione di basi di dati o DBMS (Data Base Management System) è un sistema software in grado di gestire collezioni di dati che siano:

- grandi
- condivise
- persistenti

assicurando la loro affidabilità e riservatezza.

DBMS: caratteristiche

- Dimensioni molto maggiori della memoria centrale disponibile
 - *gestione dei dati in memoria secondaria*
- Condivisione dei dati tra applicazioni ed utenti: una base di dati è una risorsa *integrata*, condivisa da più settori aziendali
 - *riduzione della ridondanza dei dati*
 - *riduzione delle inconsistenze tra i dati*
 - *meccanismo di controllo dell'accesso concorrente*

DBMS: caratteristiche

- Persistenza dei dati
 - *tempo di vita non limitato a quello dell'esecuzione dei programmi che li utilizzano*
- Affidabilità dei dati in caso di malfunzionamenti hardware e software
 - *funzionalità di salvataggio (backup) e ripristino (recovery)*
- Riservatezza dei dati
 - *meccanismi di autorizzazione per abilitare gli utenti*



DBMS: caratteristiche

- Efficienza
 - *capacità di svolgere le operazioni utilizzando un insieme di risorse (tempo e spazio) accettabile per gli utenti \Rightarrow sistema informatico adeguatamente dimensionato*
- Efficacia
 - *capacità di rendere produttive le attività degli utenti*

DBMS: caratteristiche

- Approccio "semplificato" alla gestione dei dati: dati memorizzati in modo persistente nella memoria di massa all'interno di file
 - *è possibile memorizzare e ricercare dati*
 - *semplici meccanismi di accesso (lettura sequenziale)*
 - *semplici meccanismi di condivisione (condivisione in lettura e blocco del file in scrittura)*
- I DBMS estendono le funzionalità dei file system, fornendo più servizi in modo integrato

Modello dei dati

Modello dei dati

- Un modello dei dati è un insieme di concetti utilizzati per organizzare i dati di interesse e descriverne la struttura in modo da essere comprensibile ad un elaboratore
 - *tipi di dato elementari (intero, carattere, ...)*
 - *meccanismo di strutturazione per definire strutture più complesse (costruttore di record, array, ...)*



Modello dei dati

- Modello concettuale
 - *permette di rappresentare i dati in modo indipendente dal modello logico*
 - *descrive concetti del mondo reale*
 - *utilizzato nella fase di progettazione*
 - *esempio: modello Entità-Relazione*
- Modello logico
 - *descrive la struttura dei dati nel DBMS*
 - *utilizzato dai programmi che accedono ai dati*
 - *indipendente dalle strutture fisiche*
 - *esempio: modello relazionale*

Modello relazionale

- Modello dei dati più diffuso
- Definisce il costruttore di relazione (sottoinsieme del prodotto cartesiano degli attributi-dominio), che organizza i dati in insiemi di record omogenei (ennuple a struttura fissa)
 - *le relazioni sono rappresentate sotto forma di tabelle*

Modello relazionale

Corsi

Codice	Nome	Docente
M2170	Fondamenti di informatica	Verdi
M4880	Sistemi di elaborazione	Bianchi
F0410	Basi di dati	Neri

Docenti

Nome	Dipartimento	Telefono
Verdi	Informatica	123456
Bianchi	Elettronica	636363
Neri	Informatica	414243

Altri modelli dei dati

- Prima del modello relazionale, erano utilizzati modelli più vicini alle strutture fisiche di memorizzazione (poco astratti)
 - *modello gerarchico*
 - *modello reticolare*
- Dopo il modello relazionale
 - *modello ad oggetti*
 - *XML*

Schema e istanza

- In una base di dati sono definiti:
 - ➔ lo **schema**, che descrive la struttura dei dati
 - *praticamente invariante nel tempo*
 - *è rappresentato dall'intestazione di ogni tabella (nome tabella e nomi colonne)*

Esempio:

Corsi	Codice	Nome	Docente
Docenti	Nome	Dipartimento	Telefono

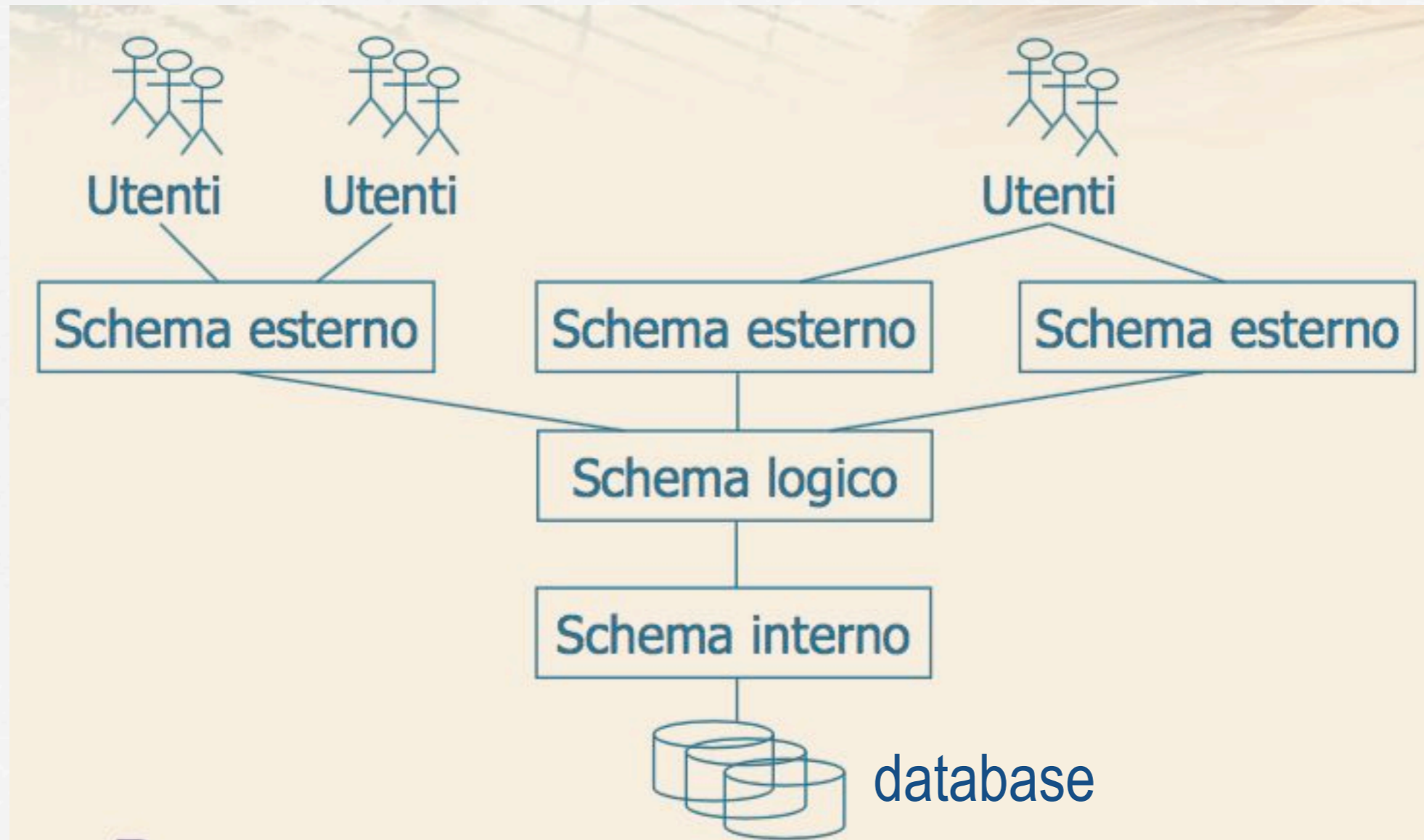
Schema e istanza

- In una base di dati sono definiti:
 - ➔ l'**istanza**, costituita dal contenuto di ogni tabella
 - *variabile nel tempo, anche molto rapidamente*
 - *è rappresentata dalle righe delle tabelle*

Esempio (istanza della tabella Docenti):

Verdi	Informatica	123456
Bianchi	Elettronica	636363
Neri	Informatica	414243

Livelli di astrazione in un DBMS



Architettura a tre livelli ANSI/SPARC

- Schema logico
 - *descrizione della base di dati mediante il modello logico del DBMS*
- Schema interno
 - *rappresentazione dello schema logico mediante strutture fisiche di memorizzazione*
- Schema esterno
 - *descrizione di parti della base di dati, denominate “viste”, che riflette il punto di vista di particolari utenti*
 - *è definito sul modello logico*

Indipendenza dei dati

Indipendenza dei dati

- L'indipendenza dei dati garantisce che gli utenti e i programmi applicativi che utilizzano una base di dati possano ignorare i dettagli realizzativi utilizzati nella costruzione della base di dati
- È una conseguenza della suddivisione in livelli di astrazione



Indipendenza fisica

- consente di interagire con il DBMS in modo indipendente dalla struttura fisica dei dati;
- l'accesso a una relazione (livello logico o esterno) avviene sempre nello stesso modo, indipendentemente dalla modalità di memorizzazione;
- è possibile modificare il modo in cui i dati sono memorizzati fisicamente senza influire sui programmi che utilizzano i dati.

Indipendenza logica

- consente di interagire con il livello esterno in modo indipendente dal livello logico;
- è possibile modificare il livello logico mantenendo inalterate le strutture esterne (purché le corrispondenze siano inalterate);
- è possibile aggiungere nuove viste o modificare viste esistenti senza modificare lo schema logico.

Accesso ai dati

Linguaggi di accesso ai dati

- Interfacce "*user friendly*" che permettono di specificare interrogazioni senza utilizzare un linguaggio testuale
- Linguaggi testuali interattivi (SQL, dichiarativo)
- Comandi simili a quelli introdotti in linguaggi di programmazione tradizionale 3G (C, C++, COBOL, Java, ...), detti linguaggi "ospite"
- Comandi simili a quelli interattivi introdotti in linguaggi di sviluppo ad hoc 4G, spesso con funzionalità specifiche (generazione di grafici, stampe complesse, videate)

Linguaggi di accesso ai dati

Sono divisi in due categorie:

- **linguaggi di definizione dei dati** (Data Definition Language o DDL) utilizzati per definire gli schemi logici, esterni e fisici e le autorizzazioni per l'accesso;
- **linguaggi di manipolazione dei dati** (Data Manipulation Language o DML) utilizzati per l'interrogazione e l'aggiornamento delle istanze della base di dati.

Figure professionali

- **Database administrator (DBA):** responsabile del controllo (centralizzato) e della gestione della base di dati
 - *garantisce prestazioni sufficienti;*
 - *assicura l'affidabilità del sistema;*
 - *gestisce le autorizzazioni di accesso ai dati.*

Figure professionali

- **Progettisti e programmatori:** definiscono e realizzano
 - *la struttura della base di dati;*
 - *i programmi che accedono alla base di dati.*
- **Utenti:** utilizzano la base di dati per le proprie attività
 - *utenti finali: utilizzano transazioni, cioè programmi che realizzano attività predefinite;*
 - *utenti casuali: formulano interrogazioni (o aggiornamenti) non predefinite mediante i linguaggi interattivi di accesso alla base di dati.*

Transazioni

- **Transazione:** sequenza di operazioni da svolgere tutte o nessuna (blocco "atomico").
- Sono implementate da programmi che realizzano attività frequenti e predefinite. Esempi:
 - *prenotazione di volo aereo*
 - *bonifico bancario*
- Generalmente realizzate con l'introduzione di SQL all'interno di un linguaggio ospite.

Vantaggi e svantaggi dei DBMS

Vantaggi dei DBMS

- Dati come risorsa comune di tutta l'organizzazione
 - *riduzione di ridondanze e inconsistenze*
- Modello dei dati unificato e preciso della realtà di interesse per l'organizzazione
- Possibile il controllo centralizzato dei dati
 - *standardizzazione, economie di scala*
- Indipendenza dei dati

Svantaggi dei DBMS

- Sono prodotti costosi, complessi, che richiedono
 - ➔ investimenti diretti
 - *acquisto del prodotto*
 - ➔ investimenti indiretti
 - *acquisizione delle risorse hardware e software necessarie*
 - *conversione delle applicazioni formazione del personale*
- Forniscono un insieme di servizi in forma integrata
 - *non è possibile scorporare servizi inutilizzati che possono causare riduzioni di prestazioni*