



# Corso di Informatica (Basi di Dati)

Lezione 1 (12 dicembre 2008)

Introduzione alle Basi di Dati

Da:

Atzeni, Ceri, Paraboschi, Torlone - Basi di Dati

Lucidi del Corso di Basi di Dati 1, Prof. Carlo Batini, Laurea in Informatica,  
AA 2007-2008

# Definizione informale

---

Una Base di Dati (o Database) è informalmente un insieme organizzato di dati utilizzato per reperire le informazioni necessarie allo svolgimento delle attività di un ente, di un'azienda, di un ufficio, di una persona, etc.

In Informatica è un archivio strutturato che consente la gestione dei dati (inserimento, ricerca, cancellazione e aggiornamento) da parte di un'applicazione software.

# Dato e Informazione

---

Cos'è un dato?

Ciò che è immediatamente presente alla conoscenza prima di ogni elaborazione.

In Informatica è un elemento di informazione costituito da simboli che devono essere elaborati.

Ad esempio le due stringhe "Mario Rossi" e "25775" sono due dati che prima di essere elaborati non significano nulla.

# Dato e Informazione

---

Cos'è un'informazione?

Notizia o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni o modi di essere.

I dati precedenti "Mario Rossi" e "25775" possono essere interpretati per arricchire la nostra conoscenza con il fatto che l'interno telefonico del signor Mario Rossi è 25775.

# Dagli archivi tradizionali... ...alle Basi di Dati

Un semplice esempio di organizzazione di dati per ottenere informazioni è l'agenda telefonica...

A-B	Telefono	C-D	Telefono
Berti	02 34556743	Ciri	02 34554322
Aldini	335 4554679	Cometti	02 34221
Bombi	323 4565430	Dini	02 3443232
		Cirielli	06 3432321

...in cui vengono specificati cognomi e numeri telefonici con prefisso

# Dagli archivi tradizionali... ...alle Basi di Dati

Quali possono essere gli "inconvenienti" legati all'agendina? Ad esempio l'agendina della slide precedente si trasforma in...

A-B	Telefono	C-D	Telefono
Berti	02 34556743	Casa	02 34554322
Aldo	335 4554679	Comune	02 34221
B.	323 4565430	Dini	02 3443232
		Ciro (a Roma)	3432321

...e si vede che non sempre i campi cognome e numero telefonico con prefisso sono sempre rispettati...

# Dagli archivi tradizionali... ...alle Basi di Dati

---

...Al posto del cognome c'è un nome (Aldo, Ciro), o solo l'iniziale con un punto (B.), o un luogo (Casa) o un riferimento sociale (Comune). Nel numero telefonico può mancare il prefisso come nel caso di Ciro in cui la città (Roma) viene specificata tra parentesi accanto al nome. Quindi si può dedurre che il prefisso è 06.

# Dagli archivi tradizionali... ...alle Basi di Dati

---

Un secondo semplice esempio è fornito da due archivi separati gestiti da due uffici distinti di una stessa università:

- **Archivio 1:** orario delle lezioni (Ufficio orario lezioni)
- **Archivio 2:** orario di ricevimento dei docenti (Ufficio orari ricevimento)



# Dagli archivi tradizionali... ...alle Basi di Dati

## Archivio 1



CCS Ingegneria Informatica Orario - Netscape

File Edit View Go Communicator Help

UNIVERSITA' DEGLI STUDI DI CHISSADOVE

**Corso di Studi in Ingegneria Informatica**


**ORARIO DELLE LEZIONI PER L'ANNO  
ACCADEMICO 1999-2000**

INSEGNAMENTO	Docente	Aula	Orario
Analisi matematica I	Luigi Neri	N1	8:00-9:30
Basi di dati	Piero Rossi	N2	9:45-11:15
Chimica	Nicola Mori	N1	9:45-11:30
Fisica I	Mario Bruni	N1	11:45-13:00
Fisica II	Mario Bruni	N3	9:45-11:15
Sistemi informativi	Piero Rossi	N3	8:00-9:30

Document: Done

# Dagli archivi tradizionali... ...alle Basi di Dati

## Archivio 2



Drari di ricevimento - Netscape

File Edit View Go Communicator Help

**UNIVERSITA' DEGLI STUDI CHISSADOVE**

**Corso di Studi in Ingegneria Informatica**

**Orario di ricevimento dei docenti**

DOCENTE	INSEGNAMENTI	ORARIO
Mario BRUNI	Fisica I Fisica II	Martedì' 10-12
Luigi NERI	Analisi matematica I	Lunedì' 12-13
Piero ROSSI	Basi di dati Sistemi informativi	Giovedì' 11-13
Nicola MORI	Chimica	Martedì' 16-18

Document: Done

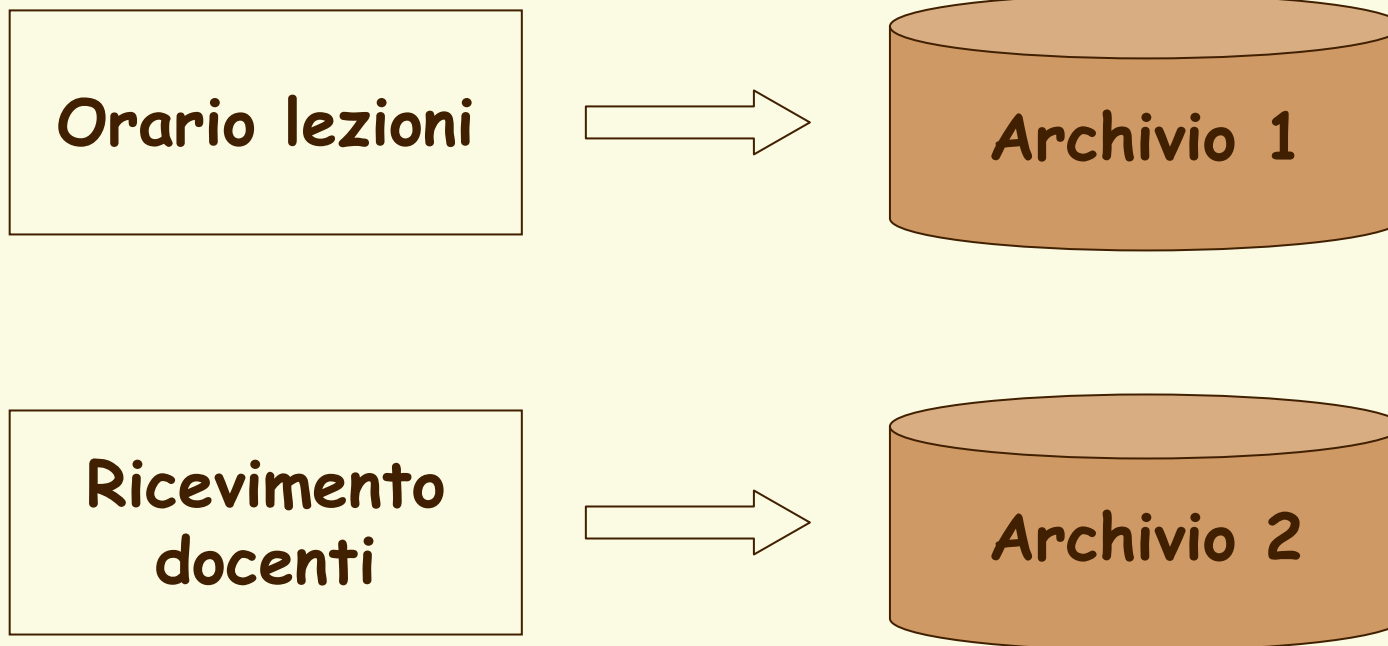
# Dagli archivi tradizionali... ...alle Basi di Dati

---

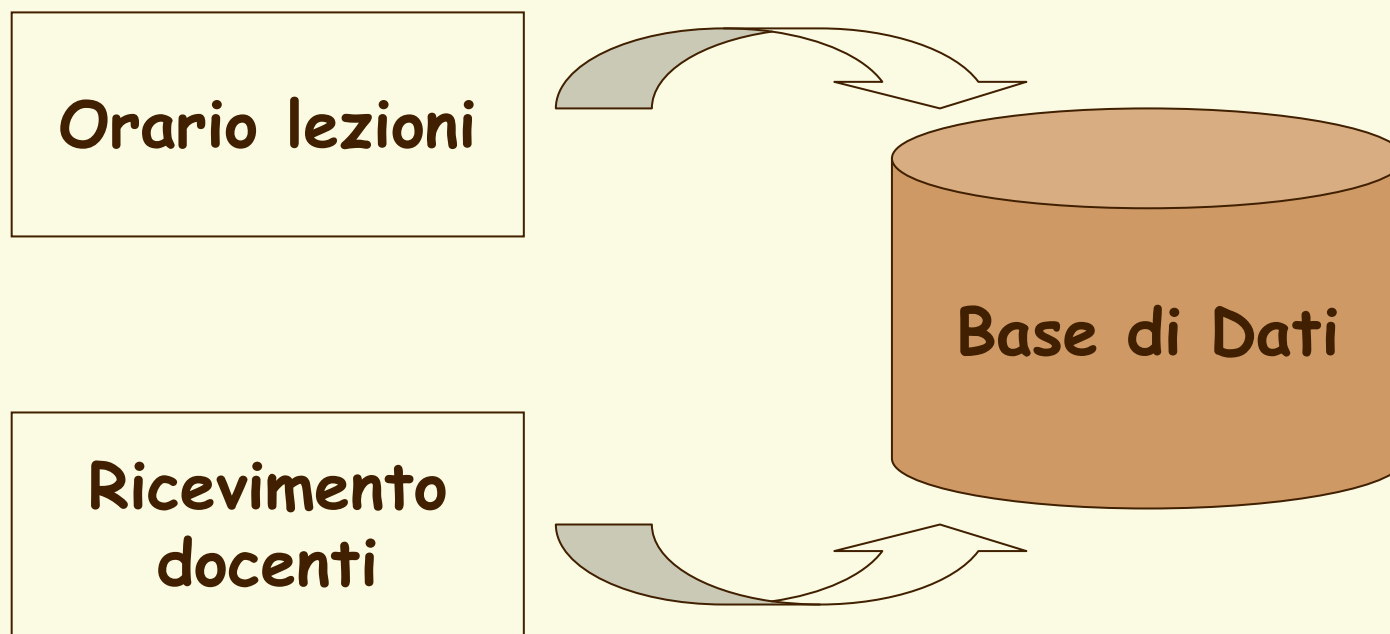
I possibili problemi di questi due archivi separati sono:

- **ridondanza** a causa della ripetizione di informazioni (ad esempio da entrambi gli archivi si può ricavare che Mario Bruni insegna Fisica I e Fisica II)
- **incoerenza** tra le informazioni presenti nei due archivi

# Dagli archivi tradizionali... ...alle Basi di Dati



# Dagli archivi tradizionali... ...alle Basi di Dati



Le informazioni dei due archivi precedenti confluiscono in un'unica Base di Dati a cui possono accedere entrambi gli uffici dell'università

# Le Basi di Dati

---

Ogni organizzazione (università, azienda, ufficio, etc.) è divisa in settori e svolge diverse attività.

A ciascun settore o attività corrisponde un sistema informativo.

Possono esistere sovrapposizioni tra i dati di interesse dei vari settori (vedere l'esempio precedente degli uffici orario lezioni e ricevimento docenti in un'università)

# Le Basi di Dati

---

Una **Base di Dati** è una risorsa di informazione **integrata e condivisa** da varie applicazioni.

Di conseguenza possono esistere attività diverse sui dati condivisi, il che richiede

meccanismi di **autorizzazione** (ad esempio per l'aggiornamento dei corsi e dei docenti), e attività multi-utente sui dati condivisi, il che richiede **controllo della concorrenza** (ad esempio nella gestione dei conti corrente in una banca).

Una Base di Dati necessita perciò di un sistema di gestione di Basi di Dati, ovvero di DBMS.

# Le Basi di Dati: DBMS

Un DBMS (DataBase Management System) è un prodotto software in grado di gestire collezioni di dati che siano:

- grandi
- condivise
- persistenti
- affidabili
- private
- efficienti
- efficaci

**Una Base di Dati è una collezione di dati gestita da un DBMS**



# Le Basi di Dati: DBMS

---

Una Base di Dati è **grande** e in generale ha dimensioni maggiori della memoria centrale. Il DBMS deve prevedere una gestione dei dati in memoria secondaria.

Una Base di Dati è **condivisa**: applicazioni e utenti diversi devono poter accedere a dati comuni. Ciò riduce ridondanze e ripetizioni. Il DBMS deve prevedere un meccanismo di controllo della concorrenza.

Una Base di Dati è **persistente**, cioè la sua vita non è limitata a quella dei programmi che la utilizzano.

# Le Basi di Dati: DBMS

---

Una Base di Dati è **affidabile** e il contenuto viene conservato intatto anche in caso di malfunzionamenti. Il DBMS fornisce funzionalità di ripristino e salvataggio (recovery e backup).

Una Base di Dati è **privata** e ogni utente viene riconosciuto e autorizzato a svolgere solo determinate attività sui dati (tramite un meccanismo di autorizzazione).

Una Base di Dati è **efficiente**, cioè permette di svolgere le operazioni utilizzando risorse di tempo e di spazio che siano accettabili per gli utenti.

# Le Basi di Dati: DBMS

---

Una Base di Dati è **efficace**, nel senso che rende produttiva l'attività degli utenti che ne fanno uso.

Alcuni esempi di DBMS sul mercato:

- Access
- DB2
- Oracle
- Informix
- Sybase
- SQLServer

# Il modello logico dei dati

---

Il modello logico dei dati è l'insieme dei concetti e dei costrutti usati per organizzare i dati e renderli comprensibili ad un elaboratore. Il modello logico dei dati è indipendente dal **modello fisico**, ovvero non dipende da come effettivamente i dati sono memorizzati sull'elaboratore.

I DBMS, in base al modello logico dei dati a cui fanno riferimento, si possono suddividere in:

- relazionali
- gerarchici
- reticolari
- a oggetti

# Il modello relazionale

---

Il **modello relazionale** è attualmente il modello logico più diffuso e utilizza il costrutto di **relazione** inteso qui in senso matematico e non come legame (relationship). Organizza i dati secondo relazioni. Una relazione si può anche vedere come una tabella a struttura fissa, in cui le righe rappresentano specifici record e le colonne rappresentano i campi (attributi) dei record. Ad esempio nella Base di Dati relazionale di un'università, i dati relativi agli studenti possono essere organizzati nella relazione/tabella della slide successiva.

# Il modello relazionale

Un esempio di relazione (che chiamiamo **Studenti**) che organizza i dati relativi agli studenti di un'università.

Matricola	Cognome	Nome	Nascita
545	Rossi	Marco	15-06-1987
876	Verdi	Andrea	13-01-1986
454	Bianchi	Luca	16-04-1985
111	Rossi	Chiara	25-08-1988

Gli attributi di ogni studente sono le intestazioni delle colonne (Matricola, Cognome, Nome, Nascita). Ogni riga corrisponde al record di un particolare studente.

# Il modello relazionale

---

Lo **schema** di una Base di Dati è la parte che rimane invariata nel tempo e che definisce le caratteristiche dei dati.

Ad esempio la relazione **Studenti** vista nella slide precedente ha 4 colonne (4 attributi) che sono rispettivamente **Matricola**, **Cognome**, **Nome**, **Nascita**. Il suo schema è il seguente:

**Studenti(Matricola, Cognome, Nome, Nascita)**

cioè è composto dal nome della relazione seguito dagli attributi tra parentesi tonde.

# Il modello relazionale

L'istanza di una Base di Dati è la parte che varia nel tempo ed è costituita dai dati (valori) contenuti nella Base di Dati.

Ad esempio la relazione **Studenti** della slide precedente ha 4 righe che corrispondono ai dati che riguardano 4 studenti. L'istanza di una relazione è l'insieme delle sue righe e può chiaramente variare nel tempo in quanto possono essere aggiunte o tolte righe. Nel caso della relazione **Studenti** si possono aggiungere nuovi studenti che si immatricolano e togliere quelli che si laureano.



# I modelli concettuali

I **modelli concettuali** permettono di descrivere la realtà da rappresentare in una Base di Dati in maniera indipendente da ogni sistema.

Essi sono utilizzati nella fase preliminare di progettazione della Base di Dati.

Il più noto è il **modello Entità-Relazione (E-R)** in cui il concetto di Relazione è in questo caso da intendersi nel senso di legame logico (relationship). Il modello E-R fondamentale descrive la realtà attraverso **entità** tra cui esistono **relazioni**. Ad esempio...

# I modelli concettuali

---

...il modello E-R per descrivere una realtà universitaria può avere le entità **Studente** e **Corso** tra cui esiste la relazione (relationship) **Esame**. Si vedrà poi che alle entità e alle relazioni nel modello E-R corrisponderanno relazioni/tabelle nel modello relazionale che definisce il modello logico dei dati.

# Linguaggi per Basi di Dati

Il linguaggi per Basi di Dati si possono distinguere in due categorie:

➤ linguaggi di definizione dei dati (Data Definition Language, **DDL**) per definire lo schema della Base di Dati

➤ linguaggi di manipolazione dei dati (Data Manipulation Language, **DML**) per l'interrogazione e l'aggiornamento delle istanze della Base di Dati.

Il linguaggio **SQL** (Structured Query Language) presenta le funzionalità di entrambe le categorie.