

Basi di Dati - Appunti lezioni

Sara Angeretti

@Sara1798

2022/2023

Indice

1	Introduzione al corso	4
1.1	I professori	4
1.2	Organizzazione del corso	4
1.3	Organizzazione degli orari	4
1.4	Organizzazione degli esami	5
1.4.1	Esoneri	5
1.4.2	Totale	5
1.4.3	Laboratorio	5
1.4.4	Valutazione	5
1.5	Il corso	6
1.5.1	Il programma	6
1.5.2	Cosa vedremo	6
2	Basi di Dati	7
2.1	Introduzione e definizioni	7
2.1.1	Risorse	7
2.1.2	Basi di Dati	7
2.2	Sistemi informativi e sistemi informatici - una premessa	7
2.2.1	Sistema Informatico	9
2.3	Gestione delle informazioni	9
2.4	Informazioni e dati	10
2.5	Basi di Dati	11
2.5.1	Altra definizione	11
2.5.2	caratteristiche	11
2.5.3	Perché studiare le basi di dati?	12
2.5.4	Basi di dati multimediali	12
2.5.5	Data Base Management System - DBMS	12
2.5.6	Creazione di un database	12
2.5.7	Interrogazione di un database	12
2.6	Schemi e istanze	13

<i>INDICE</i>	3
3 Argomento 2	14
4 Nuova lezione	15

Capitolo 1

Introduzione al corso

Turno T1 - AL

Da cancellare, per Elia, turno T1 AL ma ammettono anche MZ alle lezioni. L'importante è che gli AL sostengono l'esame con il prof di AL e gli MZ con il prof di MZ.

1.1 I professori

- prof. Napoletano, di teoria (aula 1014 in U14).
Per eventuali comunicazioni, mail: paolo.napoletano@unimib.it, ma **tassativamente** bisogna aggiungere la sigla [DB].
- prof.ssa Damiani, di esercitazione.
- prof. Raganato, di laboratorio.

1.2 Organizzazione del corso

- lezioni teoriche ed analisi di casi studio (32 ore, pari a quattro crediti) PN
- esercitazioni ed analisi casi studio (20 ore, pari a due crediti). CD
- Laboratorio / esercitazioni (20 ore, pari a crediti). AR

1.3 Organizzazione degli orari

- mercoledì: comincia a 10:30 spaccate

- giovedì: da definire
- venerdì lez: non ho capito, ma per ora non li fa (marzo)
- venerdì lab: boh lo diranno.

1.4 Organizzazione degli esami

1.4.1 Esoneri

- Ci sono parziali, circa seconda metà di aprile (**due** es di progettazione concettuale e modello relazionale) e seconda metà di giugno (**tre** es di SQL, algebra relazionale, progettazione logica).
- Per ora, i parziali sono aperti a tutti, ma il prof Napoletano deve sentire il prof Schettini del turno MZ.

1.4.2 Totale

- Tutto il programma (quindi **cinque** es) è previsto per l'esame complessivo.

1.4.3 Laboratorio

- A frequenza facoltativa ma essenziale, prevede una prova unica (sempre facoltativa) alla fine del laboratorio che permettere di avere un punteggio $-1 \leq p \leq 3$ da aggiungere alla media dei parziali o al voto del totale.
- Si lavorerà su MySQL, e prevede la progettazione concettuale e logica di una base di dati assegnata utilizzando lo strumento di Data Modeling fornito da MySQL.
- Il voto del laboratorio rimane valido tutto l'anno accademico, ovvero fino a Febbraio 2024 compreso, e andrà sempre sommato al voto preso.

1.4.4 Valutazione

- Per superare l'esame, il voto minimo *per parziale* è di **15/30**.
- Per superare l'esame, la *media dei parziali* deve essere di **18/30**.
- Verrà eventualmente sommato il voto del laboratorio (se sostenuto).

1.5 Il corso

Le lezioni non verranno registrate, ma sono disponibili quelle dell'anno precedente.

1.5.1 Il programma

1. Introduzione → prof. Napoletano
2. Metodologie e modelli per il progetto delle basi di dati → prof. Napoletano
3. La progettazione concettuale → prof. Napoletano
4. Il modello relazionale (*pausa compitini*) → prof. Napoletano
5. SQL → prof.ssa Damiani
6. Algebra relazionale → prof.ssa Damiani
7. La progettazione logica → prof. Napoletano

L'ordine degli argomenti è diverso da quello suggerito dal libro.

1.5.2 Cosa vedremo

Il corso e' dedicato a capire come e' organizzata una base di dati, a cosa serve, come si progetta, come si interroga e si crea.

Capitolo 2

Basi di Dati

Una cosa che impareremo sarà **organizzare** il lavoro che ci viene presentato, per esempio da un eventuale cliente. Per prima cosa ci servirà un *foglio dei requisiti*, ciò che dobbiamo progettare e sviluppare. Poi dovremo tradurre il passaggio da input ad output in una schematizzazione o mappa.

2.1 Introduzione e definizioni

2.1.1 Risorse

Le risorse di una azienda (o ente, amministrazione):

- persone
- denaro
- materiali
- informazioni

2.1.2 Basi di Dati

Insieme organizzato di dati utilizzati per il supporto allo svolgimento di attività (di un ente, azienda, ufficio, persona).

2.2 Sistemi informativi e sistemi informatici - una premessa

Che cos'è l'informatica? Una definizione:

- *Scienza del trattamento razionale, specialmente per mezzo di macchine automatiche, dell'informazione, considerata come supporto alla conoscenza umana e alla comunicazione (Academie Francaise).*
- L'informatica ha due anime:
 - metodologica: i metodi per la soluzione di problemi e la gestione delle informazioni;
 - tecnologica: i calcolatori elettronici e i sistemi che li utilizzano;

N.B.: **sistema informativo** \neq **sistema informatico**

Sistema informativo Componente (sottosistema) di una organizzazione che gestisce le informazioni di interesse (cioè utilizzate per il perseguimento degli scopi dell'organizzazione), le cui funzioni sono:

- acquisizione/memorizzazione
- aggiornamento
- interrogazione
- elaborazione

N.B.: Il concetto di “sistema informativo” è indipendente da qualsiasi automatizzazione!! Anche prima di essere automatizzati, molti sistemi informativi si sono evoluti verso una razionalizzazione e standardizzazione delle procedure e dell'organizzazione delle informazioni.

Sistema informatico Porzione automatizzata del sistema informativo: in pratica è la parte del sistema informativo che gestisce le informazioni con tecnologia informatica.

Ma perché le basi di dati sono così importanti? Proviamo a definirle con degli aggettivi o caratteristiche che spieghino come mai sono così interessanti:

- accessibili: le informazioni sono archiviate in modo ordinato;
- capienti: sono storate grandi quantità di dati;
- (facili da modificare;)
- ottimizzate: rapida ricerca delle informazioni; (comune a più sistemi)
- possibilità di raggruppare/filtrare le informazioni e schematizzarle/modellizzarle;
- sicurezza dei dati;;
- facilità di relazione dei dati;

- interfaccia per visualizzare in diversi modi;
- personalizzabili;
- scalabilità (Elia questa me la spieghi);
- interoperabilità: lavorabile con più linguaggi e strumenti;
- accesso concorrente alle informazioni: più persone possono lavorare allo stesso database o alla stessa sottosezione senza andare incontro ad inconsistenza dei dati;
- facilità di gestione delle ridondanze, che aiuta a ridurre al minimo l'inconsistenza dei dati;
- limitazione dell'inconsistenza dei dati: devono sempre essere consistenti, ovvero accessibili solo a chi ha diritto di farlo, il ruolo giusto;

Ma perché non usare una cosa più semplice come un FileSystem invece di un Database? Il primo mi aiuta con un'organizzazione logica, ma i Database sono dotati di strumenti (tipo la progettazione modulare) che sono più efficienti.

2.2.1 Sistema Informatico

Gestisce un sistema informativo in modo automatizzato.

Garantisce che i dati siano conservati in modo permanente sui dispositivi di memorizzazione.

Permette un rapido aggiornamento dei dati per riflettere rapidamente le loro variazioni.

Rende i dati accessibili alle interrogazioni degli utenti.

Può essere distribuito sul territorio.

2.3 Gestione delle informazioni

Parole chiave:

- Raccolta, acquisizione
- Archiviazione, conservazione
- Elaborazione, trasformazione, produzione
- Distribuzione, comunicazione, scambio

Nelle attività umane, le informazioni vengono gestite (registrate e scambiate) in forme diverse:

- idee informali
- linguaggio naturale (scritto o parlato, formale o colloquiale, in una lingua o in un'altra)
- disegni, grafici, schemi
- numeri e codici

e su vari supporti

- memoria umana, carta, dispositivi elettronici

Nelle attività standardizzate dei sistemi informativi complessi, sono state introdotte col tempo forme di organizzazione e codifica delle informazioni via via più precise (e in un certo senso artificiali).

2.4 Informazioni e dati

Informazioni \neq Dati

Nei sistemi informatici (e non solo), le informazioni vengono rappresentate in modo essenziale, spartano: **attraverso i dati**.

- **informazione**: notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere.
- **dato**: ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione; (in informatica) elementi di informazione costituiti da simboli che debbono essere elaborati.

I dati hanno bisogno di essere interpretati.

Esempio:

‘Mario’ e ‘275’ su un foglio di carta sono due dati.

Se il foglio di carta viene fornito in risposta alla domanda “A chi mi devo rivolgere per il problema X; qual è il suo numero di telefono?”, allora i dati possono essere interpretati per fornire informazione e arricchire la conoscenza.

Ma perché i dati?

La rappresentazione precisa di forme più ricche di informazione e conoscenza è difficile.

I dati costituiscono spesso una risorsa strategica, perché più stabili nel tempo di altre componenti (processi, tecnologie, ruoli umani)

I dati rimangono gli stessi nella *migrazione* da un sistema al successivo.

2.5 Basi di Dati

DB: Data Base Collezione di dati utilizzati per rappresentare le informazioni di interesse di un sistema informativo.

DBMS: Data Base Management System. Sistema software capace di gestire collezioni di dati che siano grandi, condivise e persistenti, assicurando la loro affidabilità e privacy.

Accezione generica, **metodologica**: insieme organizzato di dati utilizzati per il supporto allo svolgimento delle attività di un ente (azienda, ufficio, persona).

Accezione specifica, **metodologica** e **tecnologica**: insieme di dati gestito da un DBMS.

2.5.1 Altra definizione

Possiamo definire una BdD anche come: insieme di archivi in cui ogni dato e' rappresentato logicamente una sola volta e puo' essere utilizzato da un insieme di applicazioni da diversi utenti secondo opportuni criteri di riservatezza.

2.5.2 caratteristiche

- i dati sono molti
- i dati hanno un formato definito
- i dati sono permanenti
- i dati sono raggruppati per insiemi omogenei di dati
- esistono relazioni specifiche tra gli insiemi di dati
- la ridondanza è minima e controllata: è assicurata la consistenza delle informazioni
- i dati sono disponibili per utenze diverse e concorrenti (anche contemporanee)
- i dati sono controllati: protetti da malfunzionamenti hardware e software
- indipendenza dei dati dal programma

2.5.3 Perché studiare le basi di dati?

Copia.

2.5.4 Basi di dati multimediali

Mi sono persa tutto.

2.5.5 Data Base Management System - DBMS

Un DBMS è un insieme di programmi che permettono di creare, usare e gestire una base di dati.

Quindi un DBMS è un sistema software general purpose che facilita il processo di definizione, costruzione e manipolazione del database per varie applicazioni.

2.5.6 Creazione di un database

Tre fasi:

- definizione
- creazione/popolazione
- manipolazione

2.5.7 Interrogazione di un database

ziofrass

```
SELECT [Nome], [Cognome], [Indirizzo],  
      [Citta]  
FROM Studenti  
WHERE [Cognome]="Rossi";
```

L'efficacia della query dipende da:

- conoscenza del contenuto del db
- esperienza del linguaggio di interrogazione

oppure

- semplicità ed efficacia dell'interfaccia di interrogazione

DataBase Management System (DBMS) • Sistema che gestisce collezioni di dati: - grandi - persistenti - condivise garantendo - privatezza - affidabilità - efficienza - efficacia

Hanno grandi dimensioni: dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati; il limite deve essere solo quello fisico dei dispositivi. Sono persistenti: hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano. Sono condivise: ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività; ciascun settore/attività ha un (sotto)sistema informativo (non necessariamente disgiunto).

MI SONO PERSA UNA QUINDICINA DI SLIDES

2.6 Schemi e istanze

In ogni base di dati esistono: - lo schema, sostanzialmente invariante nel tempo, che ne descrive la struttura, il significato (aspetto intensionale). • nell'esempio, le intestazioni delle tabelle - l'istanza, i valori attuali, che possono cambiare anche molto rapidamente (aspetto estensionale) • nell'esempio, il "corpo" di ciascuna tabella

lo schema costituisce l'aspetto intensionale, ovvero la descrizione "astratta" delle proprietà, ed è invariante nel tempo. • L'istanza (i valori degli attributi) costituiscono invece l'aspetto estensionale "concreto", che varia nel tempo al variare della situazione di ciò che stiamo descrivendo

Perso altre cinque SLIDES ho rinunciato ma cerca il sito DB engines rankings

Capitolo 3

Argomento 2

Capitolo 4

Nuova lezione