

Introduzione alle basi di dati

(prima parte)

Università degli Studi di Salerno
Corso di Laurea in Scienze della Comunicazione
Informatica generale (matr. Dispari)
Docente: [Angela Peduto](#)
A.A. 2007/2008



Introduzione

- Le informazioni rappresentano una delle risorse strategiche più importanti, insieme a personale, denaro e materiali che ciascuna organizzazione o azienda deve gestire.
- Per lo svolgimento delle loro funzioni le aziende svolgono diversi processi. Tra tali processi quelli che prenderemo in esame sono quelli informativi che constano di:
 - raccolta, acquisizione delle informazioni;
 - archiviazione, conservazione delle informazioni;
 - elaborazione delle informazioni;
 - distribuzione, scambio delle informazioni.

Introduzione (2)



- In questo modulo vedremo come le informazioni vengono archiviate in modo tale da consentire di essere ritrovate per una loro successiva elaborazione o distribuzione.
- Inizialmente verranno introdotti i concetti di [sistema informativo](#), di [sistema informatico](#), di [base di dati](#) e di sistema di gestione di basi di dati (DBMS) e verrà chiarita la differenza tra informazioni e dati.
- Passeremo poi alla definizione dei vari modi di organizzare i dati introducendo il concetto di [modello dei dati](#) e vedremo che il [modello relazionale](#) garantisce l'indipendenza fisica dei dati stessi.

Introduzione (3)



- Nelle altre parti del modulo studieremo un particolare modello dei dati, denominato [modello dei dati relazionale](#), che organizza i dati in tabelle (o relazioni).
- Vedremo la [relazione](#) che intercorre tra [schema](#) di una base di dati ed [istanza](#) di una base di dati e come evitare che i dati memorizzati nella base di dati siano incoerenti, e rendano pertanto inutilizzabile la base di dati stessa.
- Introdurremo quindi il concetto di vincolo d'integrità, il cui compito è appunto di garantire l'integrità dei dati. Vedremo diversi tipi di vincoli, quelli definiti all'interno di una [tabella](#) e quelli definiti tra due tabelle.
- Discuteremo poi come è possibile manipolare le tabelle al fine di ottenerne una che contenga i valori che rispondono ad una certa [interrogazione](#). Questo processo consente di formulare delle interrogazioni per cui vedremo un particolare linguaggio, chiamato SQL, che consente di esprimere tutta una serie di operazioni che vanno dalla definizione dello schema agli aggiornamenti dei dati e all'interrogazione delle basi di dati.

Sistemi informativi e sistemi informatici



- Per comprendere cosa sono le basi di dati è opportuno inserirle nel contesto dei sistemi informativi.
- Un **sistema informativo** è un sistema di supporto ai processi informativi di un'organizzazione e pertanto non è legato in alcun modo all'informatica.
- I sistemi informativi esistono da quando l'uomo nelle pareti delle caverne disegnava dei cervi per ricordare una buona giornata di caccia.
- Ogni organizzazione utilizza al suo interno un sistema informativo. Quello classico si basa su lettere, moduli ed altri documenti cartacei.
- La porzione di sistema informativo che viene gestita in modo automatico mediante tecnologie informatiche prende il nome di **sistema informatico**.

Sistemi informativi e sistemi informatici (2)



- I sistemi informatici hanno il compito di raccogliere, organizzare e conservare le informazioni. Essi garantiscono che tali informazioni vengano conservate in modo permanente su dispositivi per la loro memorizzazione, aggiornate per riflettere rapidamente le loro variazioni e rese accessibili alle richieste degli utenti.
- Idealmente il sistema informatico dovrebbe coincidere con quello informativo.
- Purtroppo però i processi informativi gestiti in modo automatico costituiscono solo una porzione di gran parte dei sistemi informativi.
- Dal momento che un sistema informatico è di supporto ai processi di manipolazione automatica dell'informazione è opportuno chiedersi che cosa vuol dire "informazione" e come gestirla.

Sistemi informativi e sistemi informatici (4)



- Nelle attività umane più semplici le **informazioni** vengono rappresentate e scambiate secondo le tecniche naturali tipiche delle attività stesse: la lingua, scritta o parlata, disegni, figure, numeri.
- Mano a mano che le attività si sono andate sistemizzando, sono state sviluppate opportune forme di organizzazione e codifica delle informazioni.
- Nei sistemi informatici, per ragioni in parte tecnologiche e in parte legate alla semplicità dei meccanismi di gestione, il concetto di rappresentazione viene portato all'estremo: le informazioni vengono rappresentate per mezzo di dati, che hanno bisogno di essere interpretati per fornire informazioni.
- I dati da soli non hanno significato, ma, una volta interpretati e correlati opportunamente, essi forniscono informazione.

Informazione strutturata



- La struttura dell'informazione dipende dal suo utilizzo e può essere modificata nel tempo
- Esempio: per memorizzare dati su una persona, nel corso del tempo:
 - Nome e cognome (fino a qualche secolo fa non era ovvio neppure questo)
 - Nome, cognome, data di nascita e luogo di nascita
 - Codice Fiscale

Dati e Informazioni



- Nei sistemi computerizzati l'informazione è rappresentata sotto forma di dati
 - Dati: fatti grezzi che devono essere interpretati e correlati per fornire informazione
- **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

Dati e Informazioni (2)



- Esempio:
 - “Angela Peduto” e 089960022 sono una stringa e un numero, ossia due **dati**
 - Se sono restituiti in risposta alla domanda “Chi è il docente del corso e qual è il suo numero di telefono” allora costituiscono **informazione**

Basi di dati e sistemi per la gestione di basi di dati



- Una **base di dati** (o Database, DB) è una collezione di dati (di solito condivisa tra diversi applicazioni o utenti).
- Una base di dati può essere di *grandi dimensioni* (anche migliaia di miliardi di byte) e comunque di dimensioni molto maggiori della memoria primaria di un computer, di conseguenza è memorizzata in memoria secondaria.
- Ovviamente una base di dati può anche essere piccola, ma i sistemi non devono porre limiti alle dimensioni ed alla dinamicità dei dati memorizzati.
- Una base di dati è *condivisa*: applicazioni e utenti diversi devono poter accedere, secondo opportune modalità, a dati comuni. In questo modo si riduce la ridondanza dei dati, poiché si evitano ripetizioni e conseguentemente si riduce la possibilità di incoerenze.
- Una base di dati è *persistente*, cioè ha un tempo di vita superiore al tempo di esecuzione delle applicazioni che la utilizzano.

Sistema di gestione di basi di dati o DBMS



- Un **sistema di gestione di basi di dati (o Database Management System, DBMS)** è un sistema software in grado di gestire collezioni di dati che siano grandi, condivise e persistenti.
- I DBMS garantiscono *l'affidabilità*, cioè la capacità del sistema di conservare sostanzialmente intatto il contenuto della base di dati (o almeno di permetterne la ricostruzione) in caso di malfunzionamenti hardware e software. A questo scopo i DBMS forniscono specifiche funzionalità di salvataggio e ripristino (*backup e recovery*).
- I DBMS garantiscono la *privatezza* dei dati. Ciascun utente, riconosciuto in base ad un nome d'utente che è specificato all'atto di interagire con il DBMS, viene abilitato a svolgere solo determinate azioni sui dati, attraverso meccanismi di *autorizzazione*.

Sistema di gestione di basi di dati o DBMS (2)



- I DBMS gestiscono le basi di dati in modo *efficiente* ed *efficace*.
- Per *efficienza* si intende un utilizzo delle risorse (tempo e spazio) che sia accettabile dagli utenti. Ad esempio un utente che prenda un volo d'aereo attraverso Internet potrà accettare alcuni secondi di ritardo, ma un utente che controlla una centrale nucleare non potrà accettare così a lungo.
- Per *efficacia* si intende la capacità dei DBMS di rendere produttive le attività dei suoi utenti. L'attività di progettazione della base di dati e delle applicazioni che la utilizzano mira essenzialmente a garantire una buona efficacia complessiva del sistema.

Sistemi di archiviazione e banche dati



- L'attenzione alle informazioni ha caratterizzato le applicazioni dell'informatica fin dalle sue origini attraverso la realizzazione di sistemi software specificamente dedicati alla gestione delle informazioni.
- Tali sistemi spesso vengono chiamati in vari modi. A volte si parla di sistemi di archiviazione, altre di banche dati ed altre ancora di basi di dati rendendo confusa le loro funzionalità e quindi il contesto applicativo.
- Un sistema di archiviazione consente la memorizzazione e la ricerca di informazioni che sono invarianti nel tempo.
 - La caratteristica di tali sistemi è la loro staticità. Ad esempio un compact disk contenente il dizionario delle lingua italiana, oppure un'enciclopedia sono esempi di sistemi di archiviazione che tipicamente vengono aggiornati mediante la realizzazione di un nuovo sistema di archiviazione che rimpiazza il vecchio.

Sistemi di archiviazione e banche dati (2)



- Anche una **banca dati** consente la memorizzazione e la ricerca di informazioni.
- Le banche dati a differenza dei sistemi di archiviazione possono crescere nel tempo mediante aggiunta di altre informazioni e spesso sono condivise mediante una rete tra più utenti.
- Una banca dati contenente le Gazzette Ufficiali contiene informazioni testuali e l'aggiornamento non prevede la rimozione o la modifica di una Gazzetta, ma solo l'inserzione di una nuova "Gazzetta".
- Quindi né i sistemi di archiviazione né le banche dati sono in grado di gestire aggiornamenti frequenti delle informazioni. Come vedremo questo compito viene svolto dalle basi di dati.

Basi di dati e sistemi per la gestione di basi di dati



- Una **base di dati** (o Database, DB) è una collezione di dati (di solito condivisa tra diverse applicazioni o utenti).
- Una base di dati può essere di **grandi dimensioni** (anche migliaia di miliardi di byte) e comunque di dimensioni molto maggiori della memoria primaria di un computer, di conseguenza è memorizzata in memoria secondaria. Ovviamente una base di dati può anche essere piccola, ma i sistemi non devono porre limiti alle dimensioni ed alla dinamicità dei dati memorizzati.
- Una base di dati è **condivisa**: applicazioni e utenti diversi devono poter accedere, secondo opportune modalità, a dati comuni. In questo modo si riduce la ridondanza dei dati, poiché si evitano ripetizioni e conseguentemente si riduce la possibilità di incoerenze.
- Una base di dati è **persistente**, cioè ha un tempo di vita superiore al tempo di esecuzione delle applicazioni che la utilizzano.

Sistema di gestione di basi di dati o DBMS



- Un **sistema di gestione di basi di dati (o Database Management System, DBMS)** è un sistema software in grado di gestire collezioni di dati che siano grandi, condivise e persistenti.
- I DBMS garantiscono l'*affidabilità*, cioè la capacità del sistema di conservare sostanzialmente intatto il contenuto della base di dati (o almeno di permetterne la ricostruzione) in caso di malfunzionamenti hardware e software. A questo scopo i DBMS forniscono specifiche funzionalità di salvataggio e ripristino (*backup* e *recovery*).
- I DBMS garantiscono la *privatizza* dei dati. Ciascun utente, riconosciuto in base ad un nome d'utente che è specificato all'atto di interagire con il DBMS, viene abilitato a svolgere solo determinate azioni sui dati, attraverso meccanismi di *autorizzazione*.

Sistema di gestione di basi di dati o DBMS (2)



- I DBMS gestiscono le basi di dati in modo *efficiente* ed *efficace*.
- Per *efficienza* si intende un utilizzo delle risorse (tempo e spazio) che sia accettabile dagli utenti. Ad esempio un utente che prenda un volo d'aereo attraverso Internet potrà accettare alcuni secondi di ritardo, ma un utente che controlla una centrale nucleare non potrà accettare così a lungo.
- Per *efficacia* si intende la capacità dei DBMS di rendere produttive le attività dei suoi utenti. L'attività di progettazione della base di dati e delle applicazioni che la utilizzano mira essenzialmente a garantire una buona efficacia complessiva del sistema.

Tipologie di basi di dati



- Le basi di dati possono distinguersi a seconda delle dimensioni, ossia numero di utenti contemporanei, frequenza delle operazioni e sistema per la gestione di base di dati suggerito.
- Ad esempio per una base di dati personale con alcune decina di megabyte di dati, qualche utente ed una frequenza di alcune operazioni al secondo, come sistema per la gestione di basi di dati si potrebbe scegliere sia Access che MySQL.
- Per basi di dati di grandi aziende contenenti migliaia di miliardi di byte ed in cui migliaia di utenti contemporaneamente svolgono decine di milioni di operazioni come sistema per la gestione di basi di dati si potrebbero scegliere sia Oracle che DB2.

Tipologie di basi di dati (2)



- Di seguito riportiamo una tabella esemplificativa relativa alle varie tipologie di basi di dati e quindi alle relative dimensioni, numero di utenti contemporanei, frequenza delle operazioni e sistema per la gestione di base di dati suggerito.
- Questi numeri non sono assoluti, ma hanno lo scopo di esemplificare i [campi](#) di utilizzo dei vari DBMS.

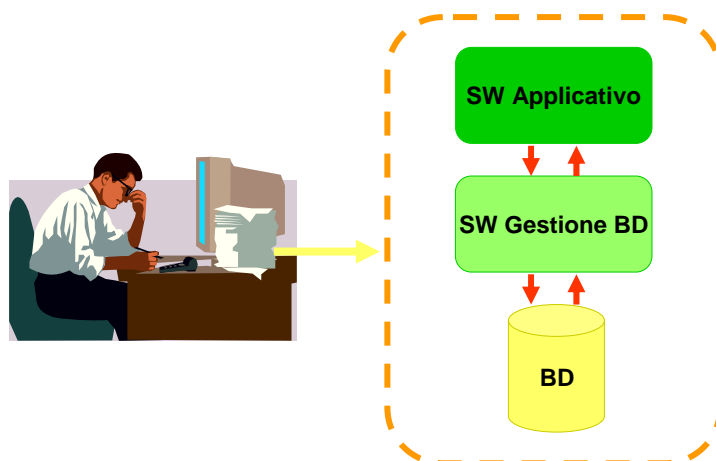
Tipologia	Dimensioni	Utenti	Oper/s	DBMS
Personali	10MB-100MB	1-3	<3	Access/MySQL
Gruppo/Dip	100MB-10GB	3-100	<300	SQLserver/MySQL
Aziendali	10GB-100GB	100-1000	<30.000	SQLServer/Oracle/DB2
Grandi Aziende	100GB-10TB	>1000	<300.000	Oracle/DB2

Componenti di un Sistema Informativo basato su computer



- Base di dati (BD)
- Software di gestione della Base di Dati
- Software applicativo
- Hardware del computer (es. dispositivi di memorizzazione)
- Personale che sviluppa, gestisce o usa il sistema

Componenti di un Sistema Informativo basato su computer



DB e DBMS



- DB (Database = Base di Dati): collezione di dati logicamente correlati di interesse per il Sistema Informativo
- DBMS (Database Management System = Sistema di gestione della Base di Dati): componente software che interagisce con la Base di Dati e con i programmi applicativi degli utenti

Riepilogando un DBMS è...



- Sistema che gestisce collezioni di dati:
 - grandi
 - persistenti
 - condivisegarantendo
 - privatezza
 - affidabilità
 - efficienza
 - efficacia

Alcuni DBMS in commercio



- Prodotti software (complessi) disponibili sul mercato; esempi:
 - Access
 - DB2
 - Oracle
 - Informix
 - Sybase
 - SQLServer