



Basi Di Dati e di conoscenza

Introduzione



Contenuti della Lezione

- Sistema informativo e informatico
- Informazioni e dati
- Base di dati

Dati, informazioni e sistemi informativi

- **Dato:** In informatica, la singola informazione codificabile o codificata (elementi di informazione costituiti da simboli che debbono essere elaborati). Ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione
- **Informazione:** notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere

Dati, informazioni e sistemi informativi

8 e Ferrari scritti su un foglio di carta sono due dati, da soli non significano nulla...

...se il foglio di carta è relativo alle ordinazioni in un ristorante la notte di Capodanno e sono note le regole che i camerieri devono seguire allora il foglio rappresenta l'ordinazione di una bottiglia di spumante marca Ferrari da addebitare alla stanza 8.

Con le indicazioni aggiuntive i dati diventano informazioni e arricchiscono la conoscenza

Dati, informazioni e sistemi informativi

- **L'evoluzione della tecnologia** permette oggi di raccogliere una grande quantità di dati, continuamente e in maniera pervasiva (social, smartphone, smartwatch, ecc).

Conseguenza: corretta e efficiente gestione dati, che è sempre stata fondamentale per la buona realizzazione di qualsiasi sistema informativo.

Le risorse di una organizzazione

- persone
- denaro
- materiali
- informazioni

Funzioni di un sistema informativo

- acquisizione delle informazioni
- conservazione delle informazioni
- elaborazione delle informazioni
- distribuzione, scambio di informazioni

Dati, informazioni e sistemi informativi

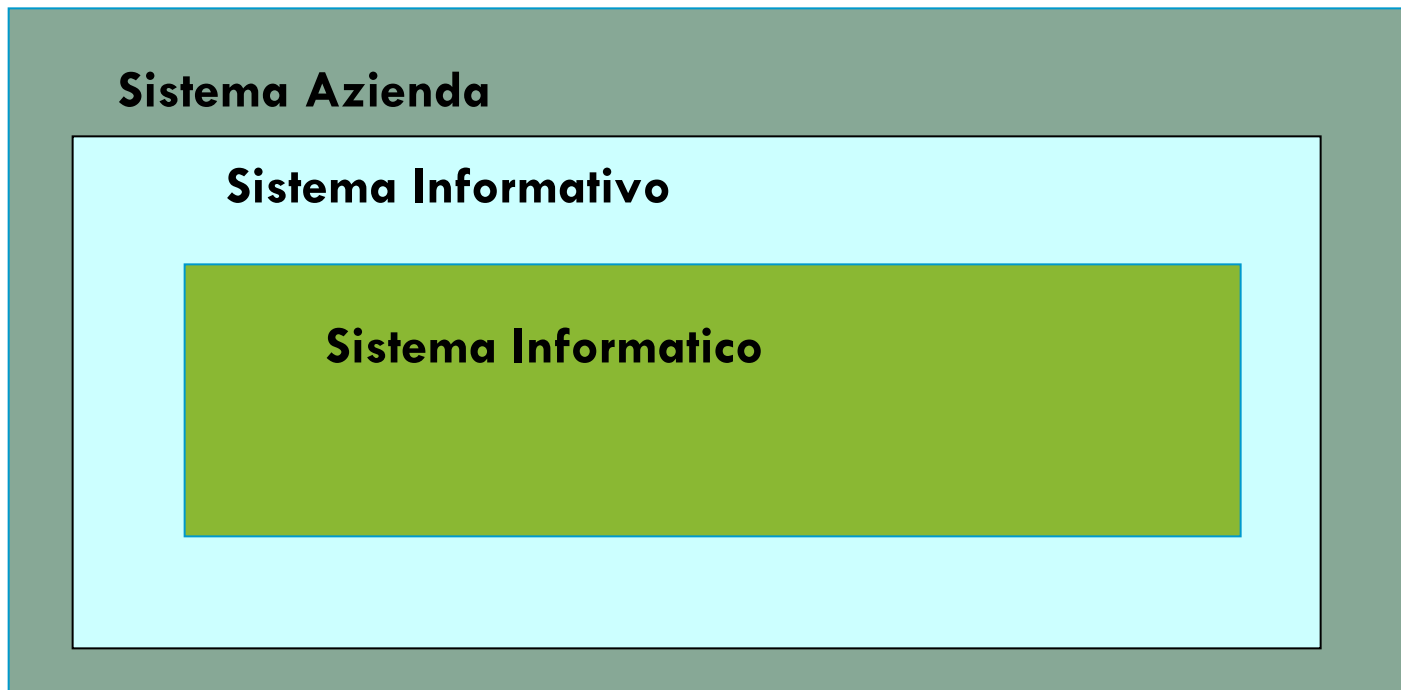
- **Sistema informativo**: componente (**sottosistema**) di un'organizzazione che gestisce (acquisisce, elabora, conserva, produce) le informazioni d'interesse (cioè utilizzate per il proseguimento degli scopi dell'organizzazione)
- Ogni organizzazione ha un sistema informativo, eventualmente non esplicitato nella struttura.
- Quasi sempre, il sistema informativo è di supporto ad altri sottosistemi, va quindi studiato nel contesto in cui è inserito
- Il sistema informativo è in genere suddiviso in sottosistemi più o meno fortemente integrati.

Dati, informazioni e sistemi informativi

- Il concetto di “**sistema informativo**” è indipendente da qualsiasi automazione
 - Esistono organizzazioni la cui ragion d’essere è la gestione d’informazioni (p.e. servizi anagrafici e banche) e che esistono da secoli.
 - (ovviamente siamo maggiormente interessati a sistemi informativi prevalentemente automatizzati)

Dati, informazioni e sistemi informativi

- Per la porzione automatizzata del sistema informativo, al giorno d'oggi viene usato il termine **sistema informatico**, termine usato oggi per contraddistinguere tutti i sistemi informativi



Dati, informazioni e sistemi informativi

- Nelle attività standardizzate dei sistemi informativi complessi, sono state introdotte nel tempo forme di organizzazione e codifica delle informazioni
- Ad esempio, nei servizi anagrafici si è iniziato con registrazioni discorsive e poi:
 - **Nome e cognome**
 - **Estremi anagrafici**
 - **Codice Fiscale**

Dati, informazioni e sistemi informativi

- Nelle attività umane, le informazioni vengono gestite (registrate e scambiate) in forme diverse:
 - Idee informali
 - Linguaggio naturale
 - Disegni, grafici, schemi
 - Numeri e codici
- ...e su vari supporti:
 - Memoria umana
 - Carta
 - Dispositivi elettronici

Dati, informazioni e sistemi informativi

- Nei sistemi informatici, le informazioni vengono rappresentate in modo essenziale attraverso i dati.
- Approssimativamente i dati non hanno alcun significato, ma se correlati e interpretati forniscono informazioni che consentono di arricchire la nostra conoscenza del mondo.

Informazioni e dati (1)

- Nei sistemi informatici, le informazioni vengono rappresentate in modo essenziale, attraverso i **dati**
- Dal Vocabolario della lingua Italiana:
 - **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, informazioni e sistemi informativi

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

Esempio

- I dati bancari hanno una struttura invariata da decenni. Le applicazioni che operano su essi invece cambiano di frequente. La nuova procedura “eredita” i dati dalla vecchia con opportune trasformazioni
- I dati vista la loro stabilità costituiscono una risorsa per l'organizzazione, un patrimonio da sfruttare e proteggere

Dati, informazioni e sistemi informativi

- All'interno del sistema informativo, la collezione dei dati è chiamata **Base di Dati** o **Database**.
 - Compito della base di dati è non solo di memorizzare i dati ma di rappresentare le relazioni tra di essi.
- All'interno del sistema informativo il software atto specificatamente a gestire i dati è detto Sistema di Gestione della basi di dati o **Database Management System**
- Tradizionalmente adottato nei sistemi informativi di grandi dimensioni (solitamente composti da più programmi), oggi adottato anche da sistemi più semplici

Base di dati

(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

Data independence: La struttura di un DB deve dare garanzia che modifiche dei dati non richiedano modifiche ai programmi applicativi e/o alle tecniche di accesso ai dati stessi

Basi di Dati

- **Basi di dati o Database:**

Collezione di informazioni registrate in formato leggibile dall'elaboratore elettronico e relativa ad un preciso dominio di conoscenze (azienda, università, mente di qualcuno, ecc.), organizzata allo scopo di poter essere consultata dai suoi utilizzatori. (Brunella Longo, 1993)

Sistema di gestione di basi di dati

DataBase Management System — DBMS

Sistema (**prodotto software**) in grado di gestire **collezioni di dati** che siano (anche):

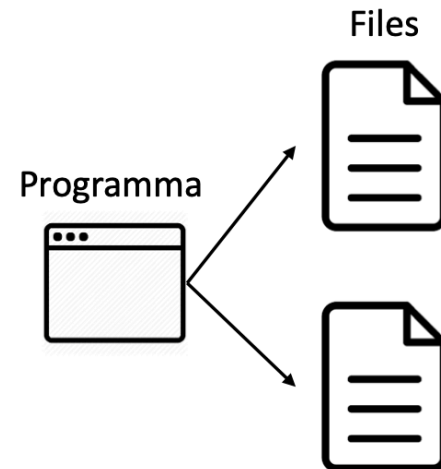
- **grandi** (di dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati)
- **persistenti** (con un periodo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano)
- **condivise** (utilizzate da applicazioni e utenti diversi)

garantendo **affidabilità** (resistenza a malfunzionamenti hardware e software) e **privatezza** (con una disciplina e un controllo degli accessi). Come ogni prodotto informatico, un DBMS deve essere **efficiente** (utilizzando al meglio le risorse di spazio e tempo del sistema) ed **efficace** (rendendo produttive le attività dei suoi utilizzatori).

Basi di Dati

- **Archivio di file**


- L'approccio classico usato dal/dai programma/i che compongono il sistema informativo per la gestione delle informazioni è un archivio basato su files
- Ogni programma ha accesso al file system gestito dal sistema operativo per creare uno o più files (**archivi**)



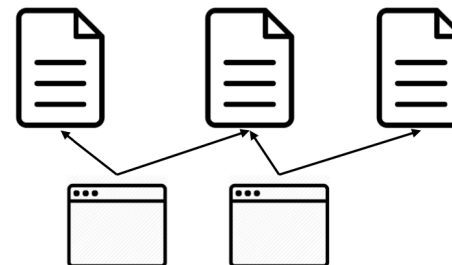
Basi di Dati

- **Archivio di file**

- Ogni file è un insieme di registrazioni (record) all'interno dei quali sono memorizzati i dati elementari (attributi e campi)
- Condivisione di dati tra più programmi può essere fatto tramite l'uso di file condivisi



316714	101001	42268	42268	2	2100.0000
316714	201003	133373	133373	12	12100.0000
316714	301005	27005	27005	5	5100.0000
316714	401007	19004	0	4	0.0000
316714	501009	46456	45426	5	5100.0000
316714	601011	11357	4250	2	1100.0000
316714	701013	21784	21784	4	4100.0000
316714	801015	117498	117498	7	7100.0000
316714	901017	36864	36864	4	4100.0000
316714	1001019	21923	3220	3	1100.0000
316714	1101021	37070	37070	5	5100.0000
316714	1201023	15902	13602	4	1100.0000
316714	1301025	28616	27027	5	4100.0000
316714	1401027	14027	11901	3	2100.0000
316714	1501029	14367	13873	2	2100.0000
316714	1601031	42610	42610	6	5100.0000
316714	1701033	52162	52162	7	7100.0000



Basi di Dati

- **Archivio di file**

- I file possono avere diversi **formati incompatibili** tra di loro, i programmi si devono adeguare a diverse convenzioni anche a distanza di parecchio tempo. Questo rende la **condivisione** dei dati attraverso applicazioni differenti **difficoltosa**
- I dati se non memorizzati su file condivisi sono **replicati** con spreco di risorse di memorizzazione e possibili problemi legati a inconsistenze
- L'accesso a file in condivisione porta a dover gestire la **concorrenza con soluzioni ad-hoc** (specialmente se due o più programmi vogliono modificarne il contenuto)

Basi di Dati

- **Archivio di file**

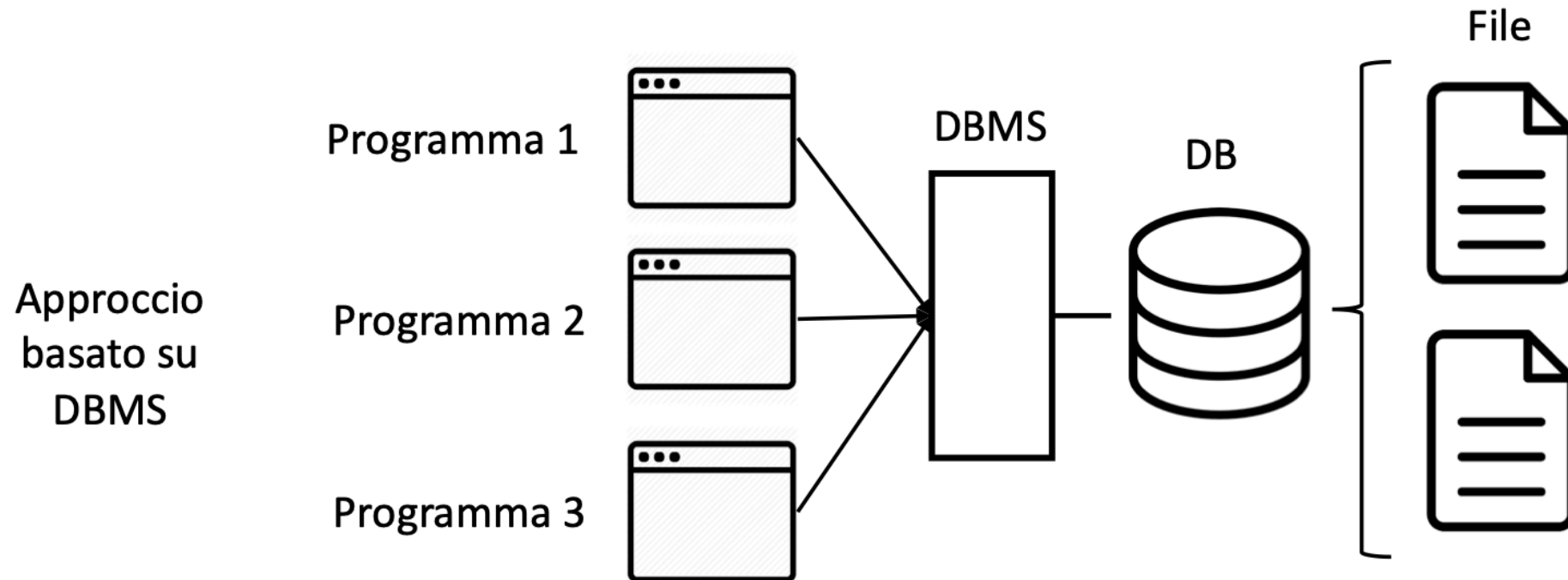
- I file possono avere diversi **formati incompatibili** tra di loro, i programmi si devono adeguare a diverse convenzioni anche a distanza di parecchio tempo. Questo rende la **condivisione** dei dati attraverso applicazioni differenti **difficoltosa**
- I dati se non memorizzati su file condivisi sono **replicati** con spreco di risorse di memorizzazione e possibili problemi legati a inconsistenze
- L'accesso a file in condivisione porta a dover gestire la **concorrenza con soluzioni ad-hoc** (specialmente se due o più programmi vogliono modificarne il contenuto)

Basi di Dati

- **Archivio di file**

- L'approccio basato su DBMS invece va oltre l'uso di file locali gestiti dalle singole applicazioni tramite l'adozione di un sistema di gestione dei dati che risulta **indipendente** dalle applicazioni e **specializzato** in tale funzione
- I dati non sono gestiti dalle singole applicazioni ma da un DBMS che offre **un'interfaccia comune** a tutte le applicazioni
- Si interpone fra le applicazioni e la memoria di massa
- I dati non appartengono ad una singola applicazione, ma esse vi accedono attraverso il DBMS

Basi di Dati



Database Management System

DBMS

- Sistema (**prodotto software**) in grado di gestire **collezioni di dati che siano** (anche):
 - **grandi** (di dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati)
 - **persistenti** (con un periodo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano)
 - **condivise** (utilizzate da applicazioni e utenti diversi)
- garantendo **affidabilità** (resistenza a malfunzionamenti hardware e software) e **privatezza** (con una disciplina e un controllo degli accessi). Come ogni prodotto informatico, un DBMS deve essere efficiente (utilizzando al meglio le risorse di spazio e tempo del sistema) ed efficace (rendendo produttive le attività dei suoi utilizzatori).

Database Management System

DBMS

- Garantendo (anche)
 - **affidabilità** (resistenza a malfunzionamenti hardware e software)
 - **privatezza** (con una disciplina e un controllo degli accessi).
- Come ogni prodotto informatico, un DBMS deve essere **efficiente** (utilizzando al meglio le risorse di spazio e tempo del sistema) ed **efficace** (rendendo produttive le attività dei suoi utilizzatori).

Database Management System

Grandi quantità di dati

- I dati gestiti da una base di dati sono:
 - Di solito più della memoria centrale: vanno gestiti in memoria secondaria
 - L'unico limite deve essere la dimensione della memoria secondaria... e nelle basi di dati distribuite neanche questo rappresenta un problema!!

Database Management System

Persistenza

- I dati hanno un **ciclo di vita** che dura nel tempo...
...un'altra ragione per la gestione in memoria secondaria

Database Management System

Condivisione

- Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività
- A ciascun settore o attività corrisponde un (sotto)sistema informativo
- Possono esistere sovrapposizioni fra i dati di interesse dei vari settori
- Una base di dati è una risorsa **integrata**, condivisa fra i vari settori

Database Management System

Condivisione

- Ridurre la **ridondanza**: una base di dati centralizzata permette di ridurre la replica della stessa informazione che si avrebbe se le diverse applicazioni gestissero i dati tramite file locali
- Ridurre **l'inconsistenza**: l'eliminazione della presenza di varie copie dello stesso dato elimina la possibilità di inconsistenze, la gestione attraverso una componente specializzata permette di introdurre controlli sui dati per garantirne la consistenza

Database Management System

DBMS sono componenti software specializzati nel gestire grandi quantità di dati e implementano procedure basate sulle best-practices (solitamente non implementate nelle soluzioni basate su file) per la gestione di:

- **Efficacia e efficienza:** le tecniche di memorizzazione adottate permettono di migliorare le prestazioni di memorizzazione e accesso alle informazioni (che altrimenti dovrebbero essere implementate in ogni programma)
- **Affidabilità:** tecniche di salvaguardia e verifica dell'integrità dei dati in caso di malfunzionamenti hardware e software (crash recovery) sono solitamente implementate

Database Management System

DBMS sono componenti software specializzati nel gestire grandi quantità di dati e implementano procedure basate sulle best-practices (solitamente non implementate nelle soluzioni basate su file) per la gestione di:

- **Concorrenza:** i sistemi DBMS implementano delle metodologie per garantire un accesso concorrente ai dati minimizzandone l'impatto sulle prestazioni di accesso (esempio limitando i tempi di attesa in seguito alla mutua esclusione su un dato)
- **Privatezza:** tecniche di sicurezza per garantire accesso ristretto sono implementate in modo da garantire a ciascun utente accesso solo al sottoinsieme dei dati a cui è autorizzato

Database Management System

DBMS sono componenti software specializzati nel gestire grandi quantità di dati e implementano procedure basate sulle best-practices (solitamente non implementate nelle soluzioni basate su file) per la gestione di:

- **Riduzione del tempo di sviluppo:** invece di implementare le funzionalità di gestione dei dati ogni applicazione si appoggia su quelle fornite dal DBMS
- **Semplificazione e standardizzazione dello sviluppo:** il processo di realizzazione delle applicazioni viene semplificato dato che la memorizzazione e la gestione dei dati è demandato ad una componente con la quale l'applicazione interagisce tramite un'interfaccia standard

Database Management System

DBMS, alcuni prodotti:

- DB2
- Oracle
- SQLServer
- Postgres
- MySQL
- SQLite