



Università degli Studi di Palermo
Dipartimento di Ingegneria Informatica



Informatica per la Storia dell'Arte

Anno Accademico 2013/2014

Docente: ing. Salvatore Sorce

Basi di dati

Facoltà di Lettere e Filosofia

Sull'uso dell'e-mail...

- **SUBJECT:** <*testo significativo*>
- Testo chiaro e conciso
- Usare l'italiano *esteso* e grammaticalmente corretto
- Inserire i dati utili per ottenere l'informazione desiderata
- **Firmare** sempre il messaggio
 - In mancanza di **subject** e di **firma**, il messaggio rimarrà **senza risposta!!**
- Altri dettagli nella **sezione F.A.Q.** del mio sito

Notizie

- Docente:
- Ing. Salvatore Sorce, Ph.D.
- salvatore.sorce@unipa.it, 09123862609

- Lezioni:
- Mar e Mer, 15-17, aula Multimediale A del Polo Didattico

- Ricevimento:
- Martedì, 11-12, @ ex-Dip. Ing. Nucleare, edificio 6, II piano
- *Dopo il corso: per appuntamento*

- Sito web:
- <http://www.unipa.it/sorce> (LEGGERE LA SEZIONE F.A.Q.)

Sistema organizzativo

- Ogni entità (ente, azienda, persona) fa uso, più o meno consapevole, di un proprio sistema organizzativo
- Sistema organizzativo: Insieme di risorse e regole per lo svolgimento coordinato delle attività per il perseguimento dei propri scopi
- Sono risorse tipiche:
 - Persone
 - Denaro
 - materiali
 - Informazioni
 - Tempo
 - Strumenti



Sistema informativo

- Sistema informativo: componente (sottosistema) di una organizzazione che gestisce (acquisisce, elabora, conserva, produce) le informazioni di interesse (cioè utilizzate per il perseguimento 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, e va quindi studiato nel contesto in cui è inserito
 - Il sistema informativo è di solito suddiviso in sottosistemi (in modo gerarchico o decentrato), più o meno fortemente integrati



Sistemi informativi e automazione

- Il sistema informativo è parte del sistema organizzativo
- Il sistema informativo esegue/gestisce processi informativi (cioè i processi che coinvolgono informazioni)
- Il concetto di “sistema informativo” è indipendente da qualsiasi automatizzazione:
 - esistono organizzazioni la cui ragion d'essere è la gestione di informazioni (p. es. servizi anagrafici e banche) e che operano da secoli

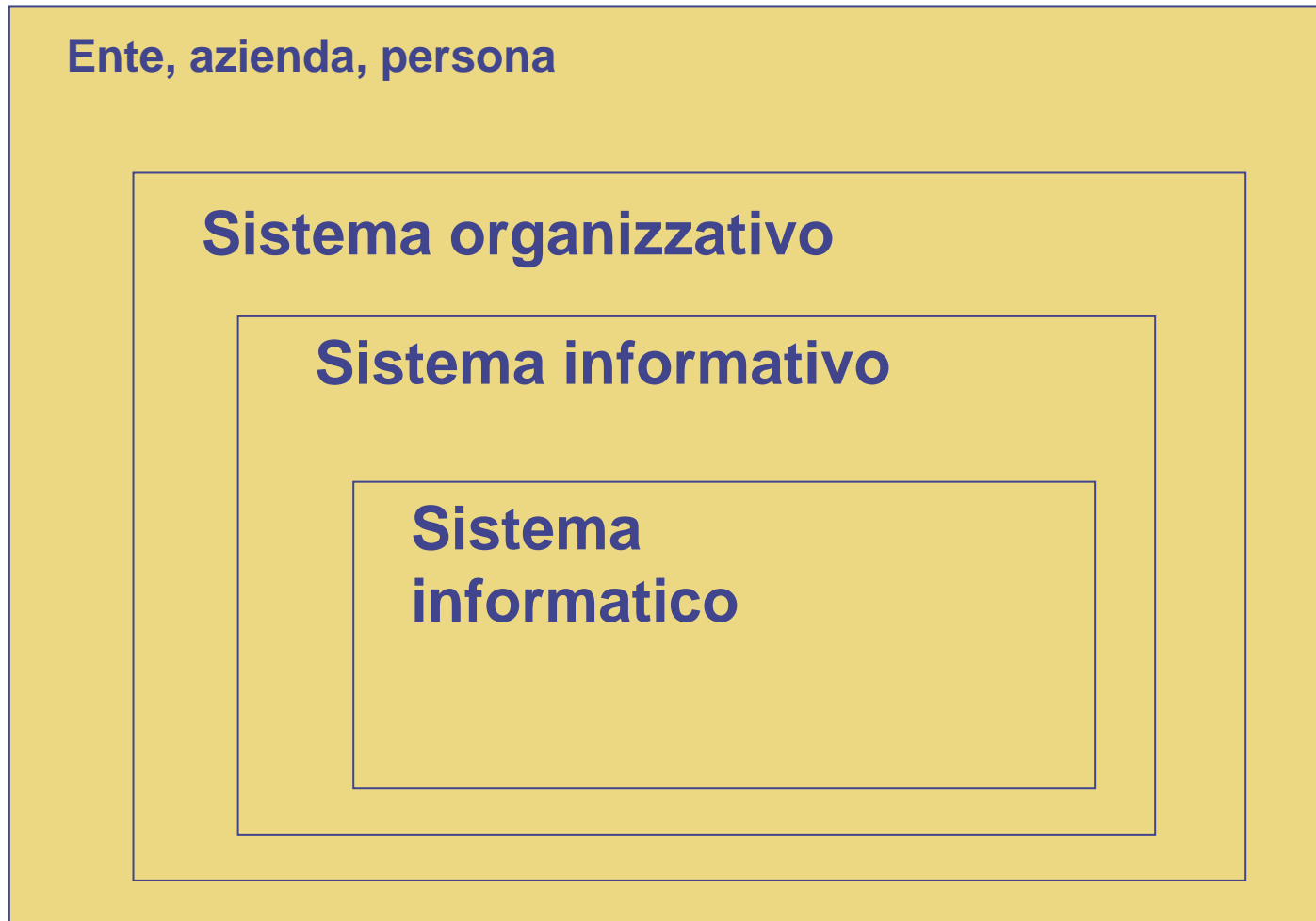


Sistema Informatico

- porzione automatizzata del sistema informativo:
la parte del sistema informativo che gestisce informazioni con tecnologia informatica



Sistema Informatico





Gestione delle informazioni

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



Gestione delle informazioni

- Nelle attività umane, le informazioni vengono gestite in forme diverse:
 - idee informali
 - linguaggio naturale (scritto o parlato, formale o colloquiale, in varie lingue)
 - disegni, grafici, schemi
 - numeri e codici
- e su vari supporti
 - mente umana, carta, dispositivi elettronici



Gestione delle informazioni

- Nelle attività standardizzate dei sistemi informativi complessi, sono state introdotte col 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 e Informazioni

- 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, ecc.
- Dato: ciò che è immediatamente presente alla conoscenza, *prima* di ogni forma di elaborazione
- Dati (informatica): elementi di un'informazione costituiti da simboli (numeri, lettere: dati numerici, alfabetici, alfanumerici) che *devono essere elaborati*, per lo più elettronicamente, secondo un determinato programma

Fonte: Vocabolario Treccani 2013



Dati e informazioni

- Un esempio:

Manzoni 1783

su un foglio di carta sono due *dati* e non significano molto

- Diventano *informazione* se contestualizzati:
 - Come si chiama l'impiegato che ha in carico la mia pratica? Qual è il suo numero di telefono diretto?
 - In che via e numero civico abita il mio amico?
 - Chi scrisse "I promessi sposi" e in che anno nacque?



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):
 - ad esempio, i dati delle banche o delle anagrafi



Basi di dati - definizione

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

- Alcune basi di dati "famosse":
 - Rubrica telefonica
 - Anagrafe
 - Conto corrente bancario

- N.B.: le basi di dati non sono necessariamente gestite da un computer

Un esempio: la catalogazione

- Tradizionalmente, gestione di un *registro* cartaceo, spesso strutturato in forma di *tabella*, in cui ogni *riga* corrisponde ad un "oggetto" catalogato
 - *oggetto* = persona, studente, disco, opera, manufatto, libro, edificio, monumento, evento, ...
- Esempio di una possibile struttura di catalogo per pinacoteca:

Codice	Autore	Titolo	Tecnica	Periodo	Collocazione	Dimensione	Note
01A01	Vincent Van Gogh	Ragazza in un bosco	Olio su tela	1882	Stanza 2, quadro 1	Cm 59x59 (LxH)	
01A02	Vincent Van Gogh	Tessitore al telaio	Olio su tela	1884	Stanza 3, quadro 2	Cm 80x64 (LxH)	In prestito
01A03	Renato Guttuso	Giocatore di carte	Olio su tela	1979	Stanza 1, quadro 1	Cm 40x24 (LxH)	

La catalogazione - problemi

- Ridondanza (dati ripetuti inutilmente)
 - Rischio di inserimenti errati o in forme espressive diverse (abbreviazioni, maiuscole/minuscole, spazi, ...)

Codice	Autore	Titolo	Tecnica	Periodo	Collocazione	Dimensione	Note
01A01	Vincent Van Gogh	Ragazza in un bosco	Olio su tela	1882	Stanza 2, quadro 1	Cm 59x59 (LxH)	
01A02	Vincent Van Gogh	Tessitore al telaio	Olio su tela	1884	Stanza 3, quadro 2	Cm 80x64 (LxH)	In prestito
01A03	Renato Guttuso	Giocatore di carte	Olio su tela	1979	Stanza 1, quadro 1	Cm 40x24 (LxH)	
...	...						
	...						
	...						
	Guttuso Renato						
	R. Guttuso						
	Guttuso						
	Vincenz Van Gog						
	V. Van Gogh						
	Van Gogh						

La catalogazione - problemi

- Incoerenza (dati di tipo diverso)
 - Rischio di inserire dati non omogenei

Codice	Autore	Titolo	Tecnica	Periodo	Collocazione	Dimensione	Note
01A01	Vincent Van Gogh	Ragazza in un bosco	Olio su tela	1882	Stanza 2, quadro 1	Cm 59x59 (LxH)	
01A02	Vincent Van Gogh	Tessitore al telaio	Olio su tela	1884	Stanza 3, quadro 2	Cm 80x64 (LxH)	In prestito
01A03	Renato Guttuso	Giocatore di carte	Olio su tela	1979	Stanza 1, quadro 1	Cm 40x24 (LxH)	
...	...						
	...						
	...						
	Guttuso Renato				I stanza, muro dx	200x230	
	R. Guttuso					1,5x1,20	
	Guttuso						
	Vincenz Van Gog						
	V. Van Gogh						
	Van Gogh						



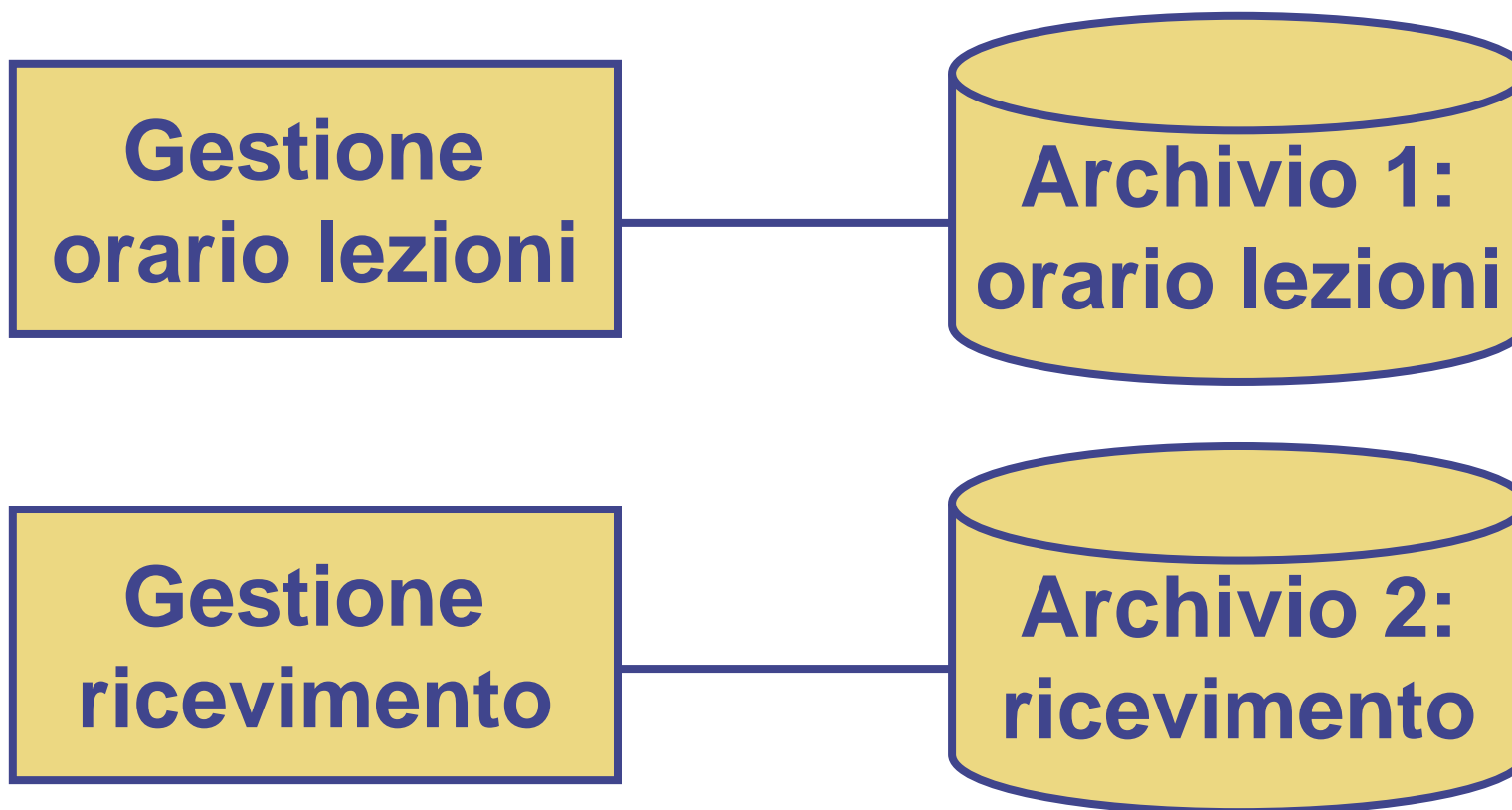
La catalogazione - problemi

- Incompletezza (dati mancanti o dimenticati)
 - Dati sconosciuti
 - Dati non inseriti per dimenticanza
 - Dati inesistenti per l'oggetto (riga)

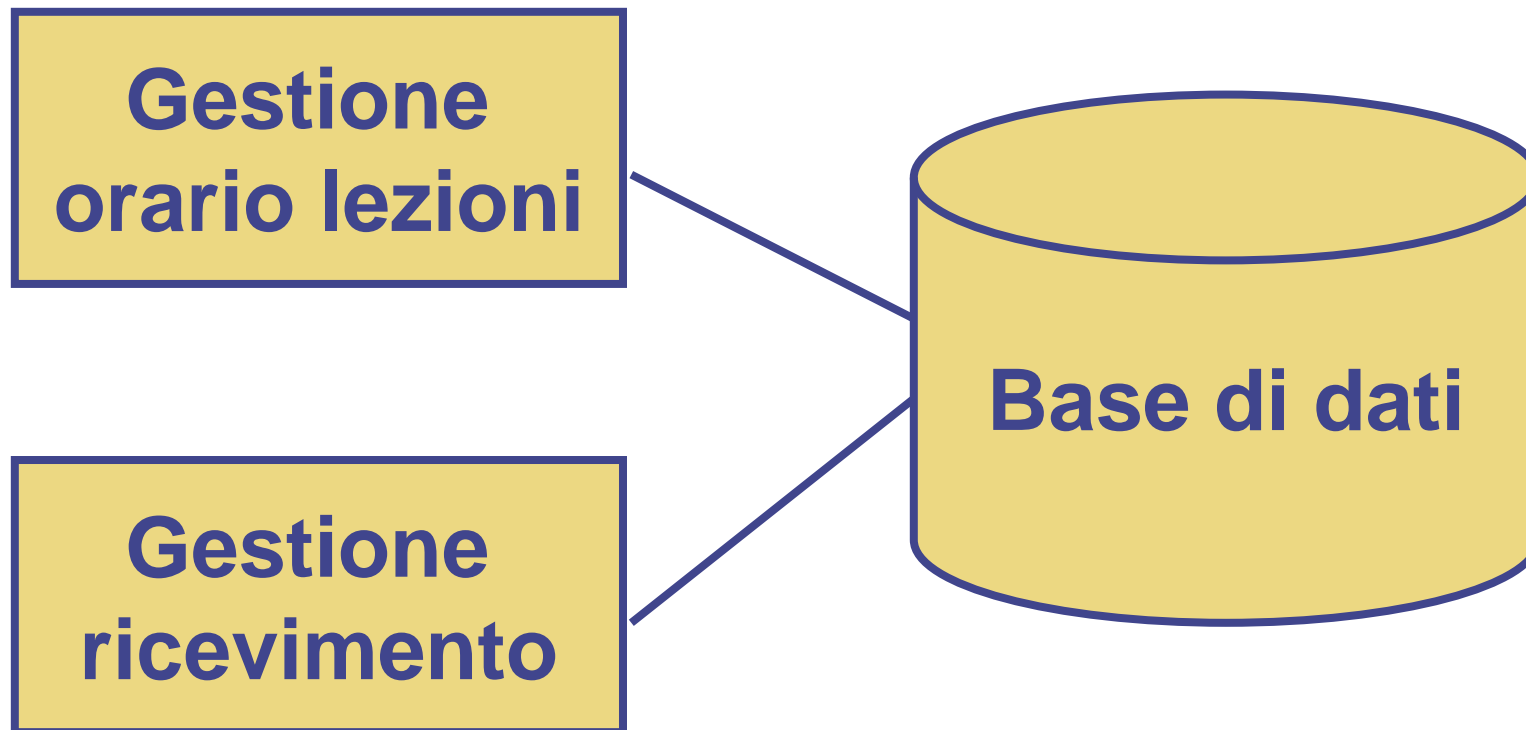
Codice	Autore	Titolo	Tecnica	Periodo	Collocazione	Dimensione	Note
01A01	Vincent Van Gogh	Ragazza in un bosco	Olio su tela	1882	Stanza 2, quadro 1	Cm 59x59 (LxH)	
01A02	Vincent Van Gogh	Tessitore al telaio	Olio su tela	1884	Stanza 3, quadro 2	Cm 80x64 (LxH)	In prestito
01A03	Renato Guttuso	Giocatore di carte	Olio su tela	1979	Stanza 1, quadro 1	Cm 40x24 (LxH)	
...	...						
	...						
	...						
	Guttuso Renato			1978	I stanza, muro dx	200x230	
	R. Guttuso					1,5x1,20	
	Guttuso			1975			
	Vincenz Van Gog						
	V. Van Gogh						
	Van Gogh						



Archivi e basi di dati



Archivi e basi di dati





Basi di dati

- In informatica: insieme di dati gestito da un sistema (informatico) di gestione per basi di dati (Data Base Management System, DBMS)
- DBMS, sistema informatico che gestisce collezioni di dati:
 - grandi
 - persistenti
 - condivise
- garantendo
 - privacy
 - affidabilità
 - efficienza
 - efficacia

Le basi di dati sono ...

- ...grandi
 - dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati
 - il limite deve essere solo quello fisico dei dispositivi
- ...persistenti
 - hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano
- ...condivise
 - Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività
 - Ciascun settore/attività ha un suo (sotto)sistema informativo che accede alla base di dati comune

I DBMS garantiscono ...

- ...privacy
 - Si possono definire meccanismi di autorizzazione
 - ◆ l'utente A è autorizzato a leggere tutti i dati e a modificare quelli sul ricevimento
 - ◆ l'utente B è autorizzato a leggere X e a modificare Y
- ...affidabilità
 - resistenza a malfunzionamenti hardware e software
 - ◆ Una base di dati è una risorsa pregiata e quindi deve essere conservata a lungo termine
- ...efficienza
 - Cercano di utilizzare al meglio le risorse di spazio di memoria (principale e secondaria) e tempo (di esecuzione e di risposta)
 - I DBMS, con tante funzioni, rischiano l'inefficienza e per questo ci sono grandi investimenti e competizione
 - L'efficienza è anche il risultato della qualità delle applicazioni
- ...efficacia
 - Cercano di rendere produttive le attività dei loro utilizzatori, offrendo funzionalità articolate, potenti e flessibili



Descrizione dei dati nei DBMS

- Rappresentazioni dei dati a livelli diversi
 - permettono l'indipendenza dei dati dalla rappresentazione fisica
 - ◆ singolo file vs. N files
 - ◆ file locali vs. distribuiti
 - i programmi fanno riferimento alla struttura a livello più alto, e le rappresentazioni sottostanti possono essere modificate senza necessità di modifica dei programmi

- Ogni DBMS fa riferimento ad uno specifico *modello di dati*

Modello di dati

- Insieme di costrutti utilizzati per organizzare i dati di interesse e descriverne la dinamica
- Componente fondamentale: meccanismi di strutturazione (o costruttori di tipo)
- Esempio: il *modello relazionale* che prevede la definizione di insiemi omogenei (oggetti che stanno tra loro in una data relazione)
 - strutturazione in tabelle
 - ◆ Ogni *tabella* rappresenta una relazione
 - ◆ Ogni *riga* della tabella (record) rappresenta un oggetto
 - ◆ Ogni *colonna* della tabella (campo) rappresenta un dettaglio per l'oggetto



Modello relazionale - esempio

Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00



Schema e istanza

- *Schema* di una base di dati
 - Descrizione della struttura, sostanzialmente invariante nel tempo
 - ◆ es.: le intestazioni delle tabelle

- *Istanza* di una base di dati
 - i valori attuali dei dati, che possono cambiare anche molto rapidamente
 - ◆ es.: il contenuto di ciascuna tabella

Ex



Basi di dati: schema e istanza

Lo *schema* della base di dati

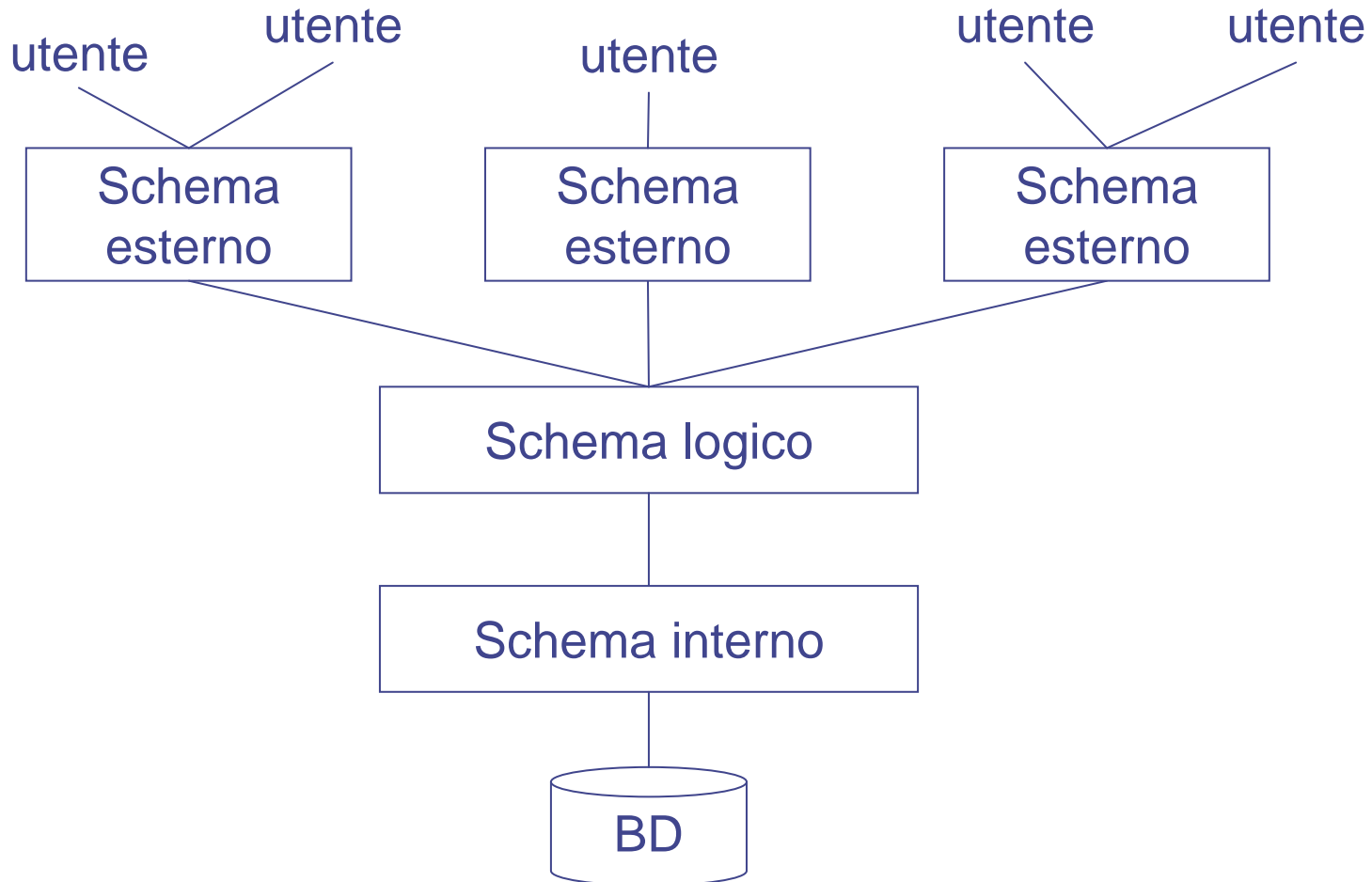
Orario

Insegnamento	Docente	Aula	Ora
Analisi matem. I	Luigi Neri	N1	8:00
Basi di dati	Piero Rossi	N2	9:45
Chimica	Nicola Mori	N1	9:45
Fisica I	Mario Bruni	N1	11:45
Fisica II	Mario Bruni	N3	9:45
Sistemi inform.	Piero Rossi	N3	8:00

L'*istanza* della base di dati



Architettura standard a tre livelli per i DBMS





Architettura semplificata di un DBMS

- schema esterno: descrizione (visualizzazione) di parte della base di dati in un dato modello logico (“viste” parziali, derivate anche da modelli diversi)
 - Esempio 1: visualizzazione in tabella di un DB reticolare
 - Esempio 2: visualizzazione in unica tabella di dati provenienti da più tabelle
- schema logico: descrizione della base di dati nel modello logico usato
 - Esempio: la struttura della tabella nel modello relazionale
- schema interno (o fisico): rappresentazione dello schema logico per mezzo di strutture di memorizzazione
 - Esempio: uno o più file
- Lo schema logico è indipendente da quello fisico:
 - una tabella è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica (che può anche cambiare nel tempo)

Personaggi e interpreti

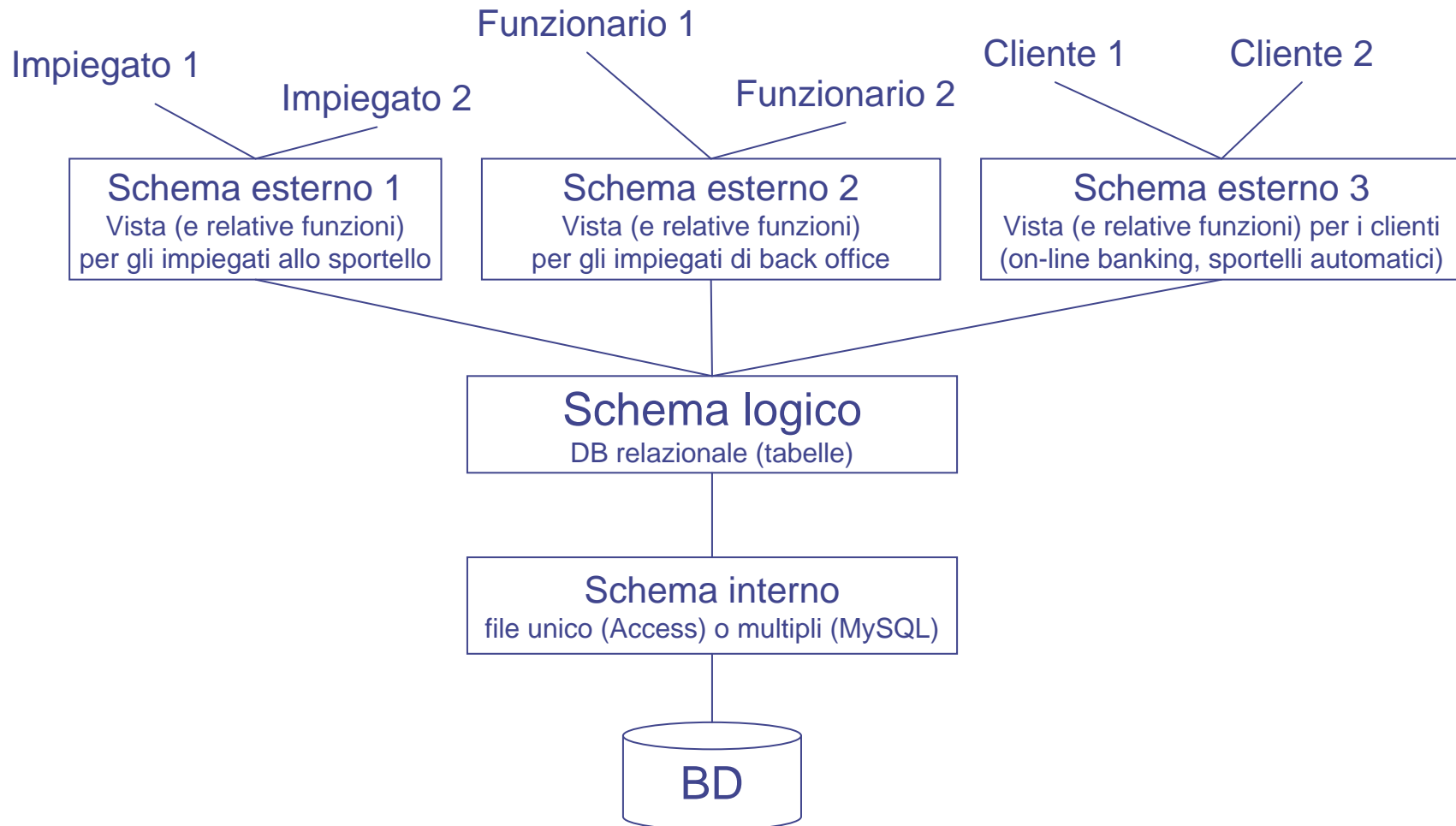
- progettisti e realizzatori di DBMS
 - Esempio: i programmatori di Microsoft per il DBMS "Access"
- progettisti della base di dati e amministratori della base di dati
 - DBA: Data Base Administrators
- progettisti e programmatori di applicazioni
 - Esempio: programmatori di un istituto bancario per le loro applicazioni basate su MS Access
- utenti
 - utenti finali (terminalisti): eseguono applicazioni predefinite (transazioni)
 - utenti casuali: eseguono operazioni non previste a priori, usando linguaggi interattivi

Transazioni

- Programmi che realizzano attività frequenti e predefinite, con poche eccezioni, previste a priori.
- programma a disposizione, da eseguire per realizzare le funzioni di interesse

- Esempi:
 - versamento presso uno sportello bancario
 - emissione di certificato anagrafico
 - dichiarazione presso l'ufficio di stato civile
 - prenotazione aerea

Esempio a tre livelli per un DB bancario



Dai dati alle informazioni

- Una volta strutturata e popolata una base di dati, come si estraggono le informazioni?
- I DBMS forniscono strumenti (linguaggi) per comporre al DB le richieste (*query*), anche complesse, di aggregazione
 - Criteri di selezione e visualizzazione di dati, provenienti anche da tabelle diverse
- Esempio:
 - Mostra tutti gli studenti che hanno un'età inferiore ai 30 anni e residenti fuori sede (tabella "Studenti")
 - Mostra tutti i clienti della banca che hanno il conto in rosso (tabelle "Clienti" e "Conti correnti")
 - Mostra tutti gli studenti della Facoltà di Lettere e Filosofia che hanno più di 24 anni, che sono iscritti al Corso di LM in Storia dell'Arte, I anno (tabelle "Studenti", "Corsi di Laurea", "Iscrizioni", "Insegnamenti")



Vantaggi e svantaggi dei DBMS

Pro

- dati come risorsa comune, base di dati come modello della realtà
- gestione centralizzata con possibilità di standardizzazione ed "economia di scala"
- disponibilità di servizi integrati
- riduzione di ridondanze e inconsistenze
- indipendenza dei dati (favorisce lo sviluppo e la manutenzione delle applicazioni)

Contro

- costo dei prodotti e della transizione verso di essi
- non scorporabilità delle funzionalità (con riduzione di efficienza)