



UNIVERSITÀ DEGLI STUDI DI PALERMO

INFORMATIZZAZIONE DELLA DOCUMENTAZIONE SANITARIA

Anno Accademico 2019/2020

Docente: ing. Salvatore Sorce

Basi di dati I

Che cos'è l'informatica?

- *Scienza del **trattamento razionale**, specialmente **per mezzo di macchine automatiche**, dell'informazione, considerata come supporto alla conoscenza umana e alla comunicazione*

Due anime:

- **metodologica**
- **tecnologica**

Base di dati

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

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)

Sistema informativo, commenti

- 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

Sistema organizzativo

- Insieme di risorse e regole per lo svolgimento coordinato delle attività (processi) al fine del perseguimento degli scopi

Risorse

- le risorse di una azienda (o ente, amministrazione):
 - persone
 - denaro
 - materiali
 - informazioni

Sistema organizzativo vs. informativo

- Il sistema informativo è parte del sistema organizzativo
- Il sistema informativo esegue/gestisce processi informativi (cioè i processi che coinvolgono informazioni)

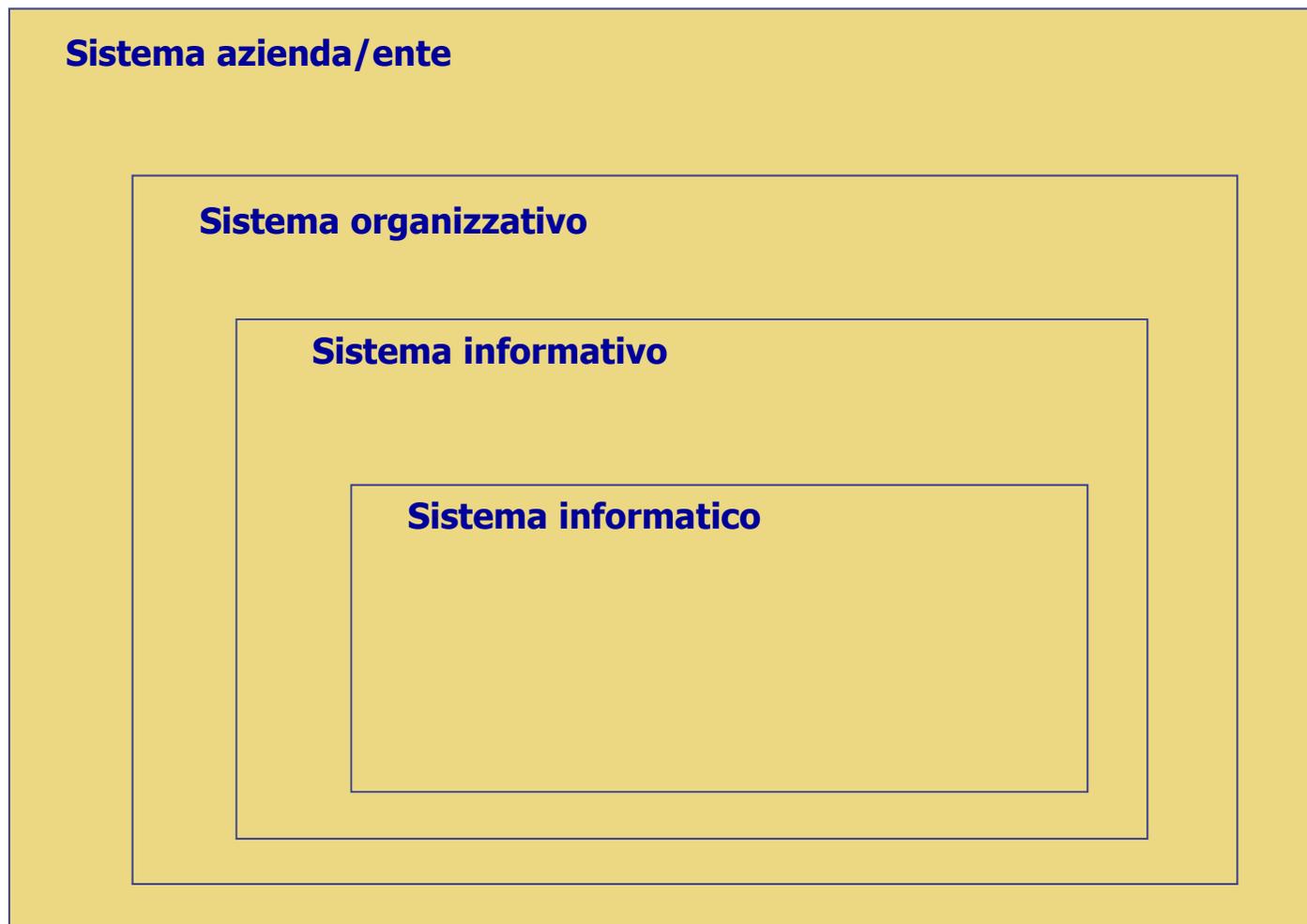
Sistemi informativi e automazione

- 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

Informazioni e dati

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

Informazioni e dati

(definizioni dal Vocabolario della lingua italiana 1987)

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

- Un esempio:

Mario 275

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

- 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

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

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

Sistema di gestione di basi di dati DataBase Management System (DBMS)

- Sistema informatico che gestisce collezioni di dati:
- grandi
 - persistenti
 - condivise
- garantendo
- riservatezza
 - affidabilità
 - efficienza
 - efficacia

DBMS

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

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

Le basi di dati sono ... persistenti

- hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano

Le basi di dati 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)

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

Orari di ricevimento - Netscape

File Edit View Go Communicator Help

UNIVERSITA' DEGLI STUDI CHISSADOVE

Corso di Studi in Ingegneria Informatica

Orario di ricevimento dei docenti

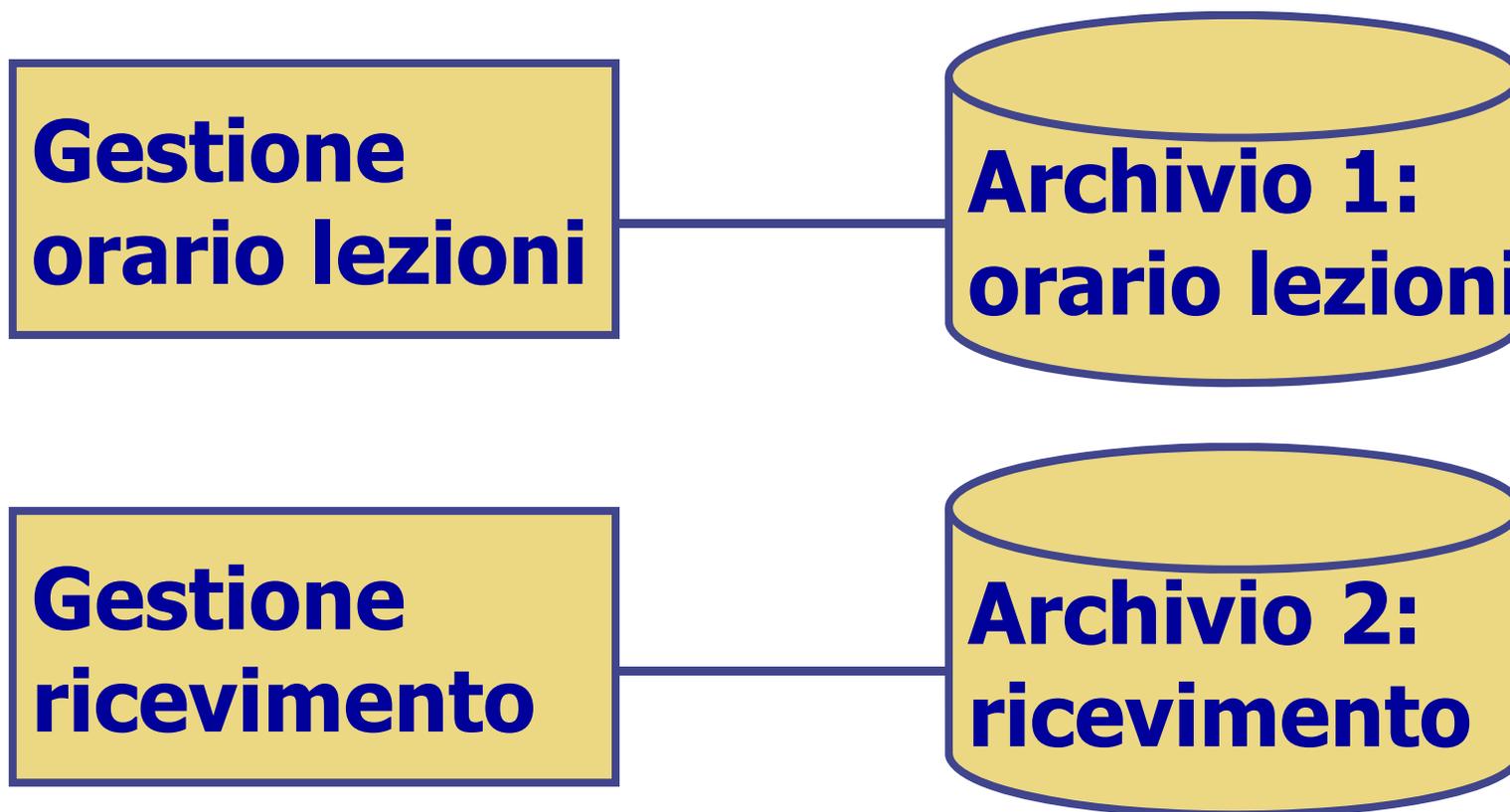
DOCENTE	INSEGNAMENTI	ORARIO
Mario BRUNI	Fisica I Fisica II	Martedì' 10-12
Luigi NERI	Analisi matematica I	Lunedì' 12-13
Piero ROSSI	Basi di dati Sistemi informativi	Giovedì' 11-13
Nicola MORI	Chimica	Martedì' 16-18

Document: Done

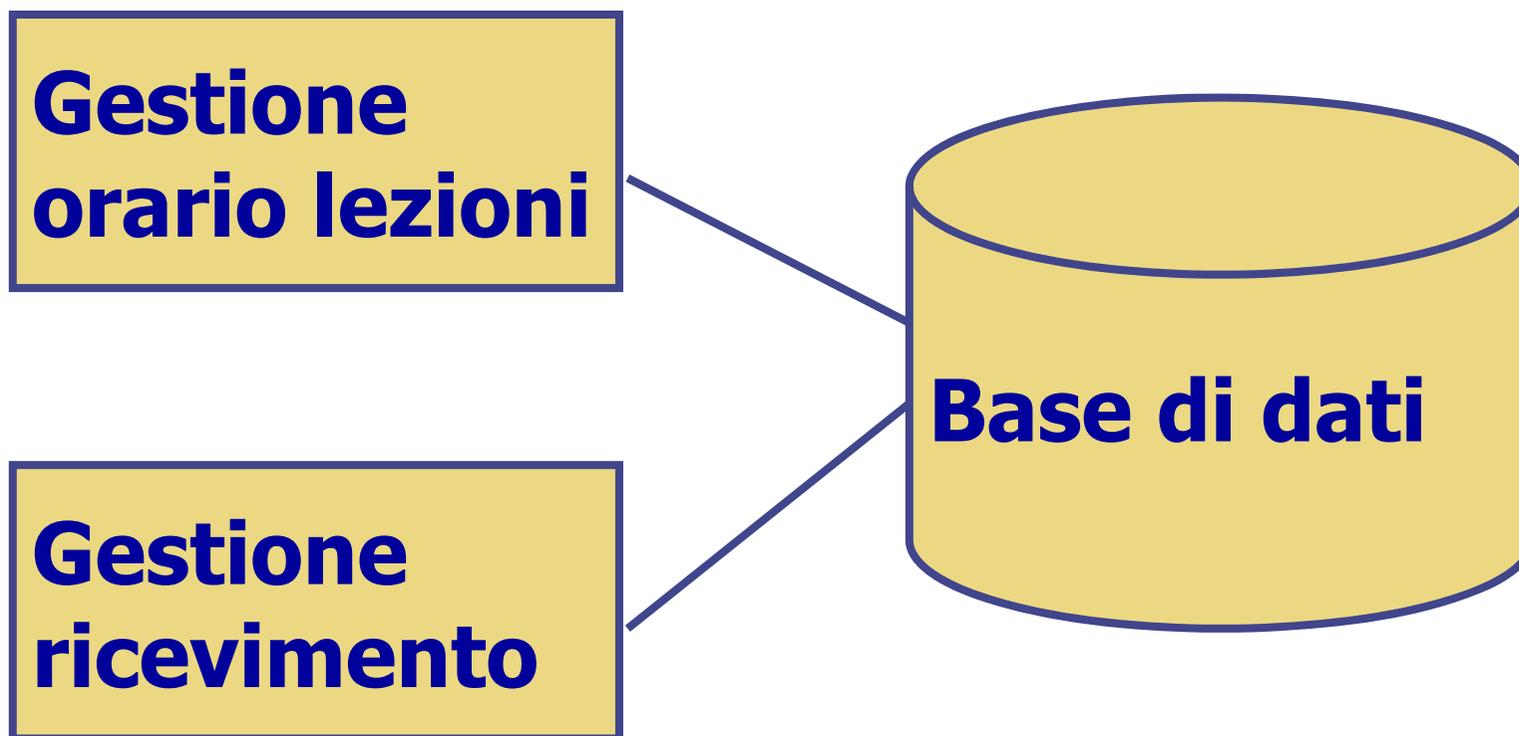
Problemi

- **Ridondanza:**
 - informazioni ripetute
- **Rischio di incoerenza:**
 - le versioni possono non coincidere

Archivi e basi di dati



Archivi e basi di dati



Le basi di dati sono condivise

- Una base di dati e' una risorsa **integrata, condivisa** fra applicazioni
- conseguenze
 - Attività diverse su dati condivisi:
 - ◆ meccanismi di **autorizzazione**
 - Accessi di più utenti ai dati condivisi:
 - ◆ controllo della **concorrenza**

I DBMS garantiscono ... riservatezza

- Si possono definire meccanismi di autorizzazione
 - l'utente A è autorizzato a leggere tutti i dati e a modificarne alcuni
 - l'utente B è autorizzato a leggere X e a modificare Y

I DBMS garantiscono... affidabilità

- **Affidabilità** (per le basi di dati):
 - resistenza a malfunzionamenti hardware e software
- Una base di dati è una risorsa pregiata e quindi deve essere conservata a lungo termine
- Tecnica fondamentale:
 - gestione delle **transazioni**

Transazione

- Insieme di operazioni da considerare indivisibile ("atomico"), corretto anche in presenza di concorrenza e con effetti definitivi

Le transazioni sono ... atomiche

- La sequenza di operazioni sulla base di dati viene eseguita per intero o per niente:
 - trasferimento di fondi da un conto A ad un conto B: o si fanno il prelievo da A e il versamento su B o nessuno dei due

Le transazioni sono ... concorrenti

- L'effetto di transazioni concorrenti deve essere coerente (ad esempio "equivalente" all'esecuzione separata)
 - se due assegni emessi sullo stesso conto corrente vengono incassati contemporaneamente si deve evitare di trascurarne uno

I risultati delle transazioni sono permanenti

- La conclusione positiva di una transazione corrisponde ad un impegno (in inglese **commit**) a mantenere traccia del risultato in modo definitivo, anche in presenza di guasti e di esecuzione concorrente

I DBMS debbono essere...efficienti

- Cercano di utilizzando 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

I DBMS debbono essere...efficaci

- Cercano di rendere produttive le attività dei loro utilizzatori, offrendo funzionalità articolate, potenti e flessibili

DBMS vs file system

- La gestione di insiemi di dati grandi e persistenti è possibile anche attraverso sistemi più semplici — gli ordinari **file system** dei sistemi operativi
- I file system prevedono forme rudimentali di condivisione: "tutto o niente"
- I DBMS estendono le funzionalità dei file system, fornendo più servizi ed in maniera integrata

DBMS vs file system (2)

- Nei programmi tradizionali che accedono a file, ogni programma contiene una descrizione della struttura del file stesso, con i conseguenti rischi di incoerenza fra le descrizioni (ripetute in ciascun programma) e i file stessi
- Nei DBMS, esiste una porzione della base di dati (il **catalogo** o **dizionario**) che contiene una descrizione centralizzata dei dati, che può essere utilizzata dai vari programmi

Descrizioni dei dati nei DBMS

- Rappresentazioni dei dati a livelli diversi
 - permettono l'indipendenza dei dati dalla rappresentazione fisica:
 - ◆ i programmi fanno riferimento alla struttura a livello più alto, e le rappresentazioni sottostanti possono essere modificate senza necessità di modifica dei programmi
 - Precisiamo attraverso il concetto di
 - ◆ **modello dei dati**

Modello dei dati

- Insieme di costrutti utilizzati per organizzare i dati di interesse e descriverne la dinamica
- Componente fondamentale: **meccanismi di strutturazione** (o **costruttori di tipo**)
- Come nei linguaggi di programmazione esistono meccanismi che permettono di definire nuovi tipi, così ogni modello dei dati prevede alcuni costruttori
- Esempio: il **modello relazionale** prevede il costruttore **relazione**, che permette di definire insiemi di record omogenei

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

Organizzazione dei dati in una 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

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

Schema e istanza

- In ogni base di dati esistono:
 - lo **schema**, sostanzialmente invariante nel tempo, che ne descrive la struttura
 - ◆ es.: le intestazioni delle tabelle
 - l'**istanza**, i valori attuali, che possono cambiare anche molto rapidamente
 - ◆ es.: il "corpo" di ciascuna tabella

Due tipi (principali) di modelli

- modelli logici
- modelli concettuali

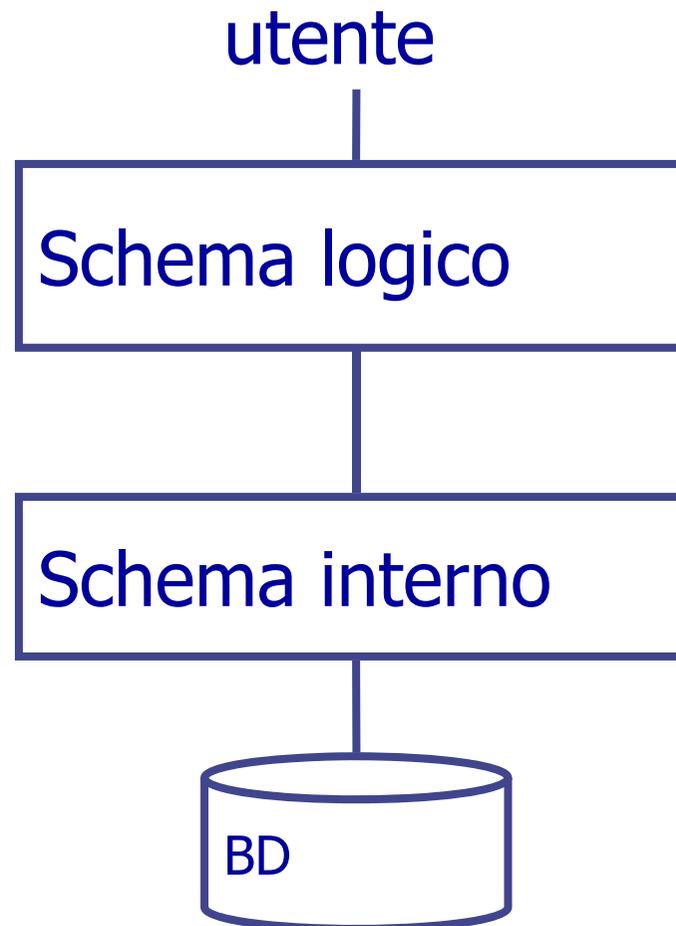
Modelli logici

- Adottati nei DBMS esistenti per l'organizzazione dei dati
 - utilizzati dai programmi
 - indipendenti dalle strutture fisiche
- esempi: **relazionale**, reticolare, gerarchico, a oggetti

Modelli concettuali

- Permettono di rappresentare i dati in modo indipendente da ogni sistema
 - cercano di descrivere i concetti del mondo reale
 - sono utilizzati nelle fasi preliminari di progettazione
- Il più diffuso è il modello **Entity-Relationship** (entità-relazione)

Architettura (semplificata) di un DBMS



Architettura semplificata di un DBMS: schemi

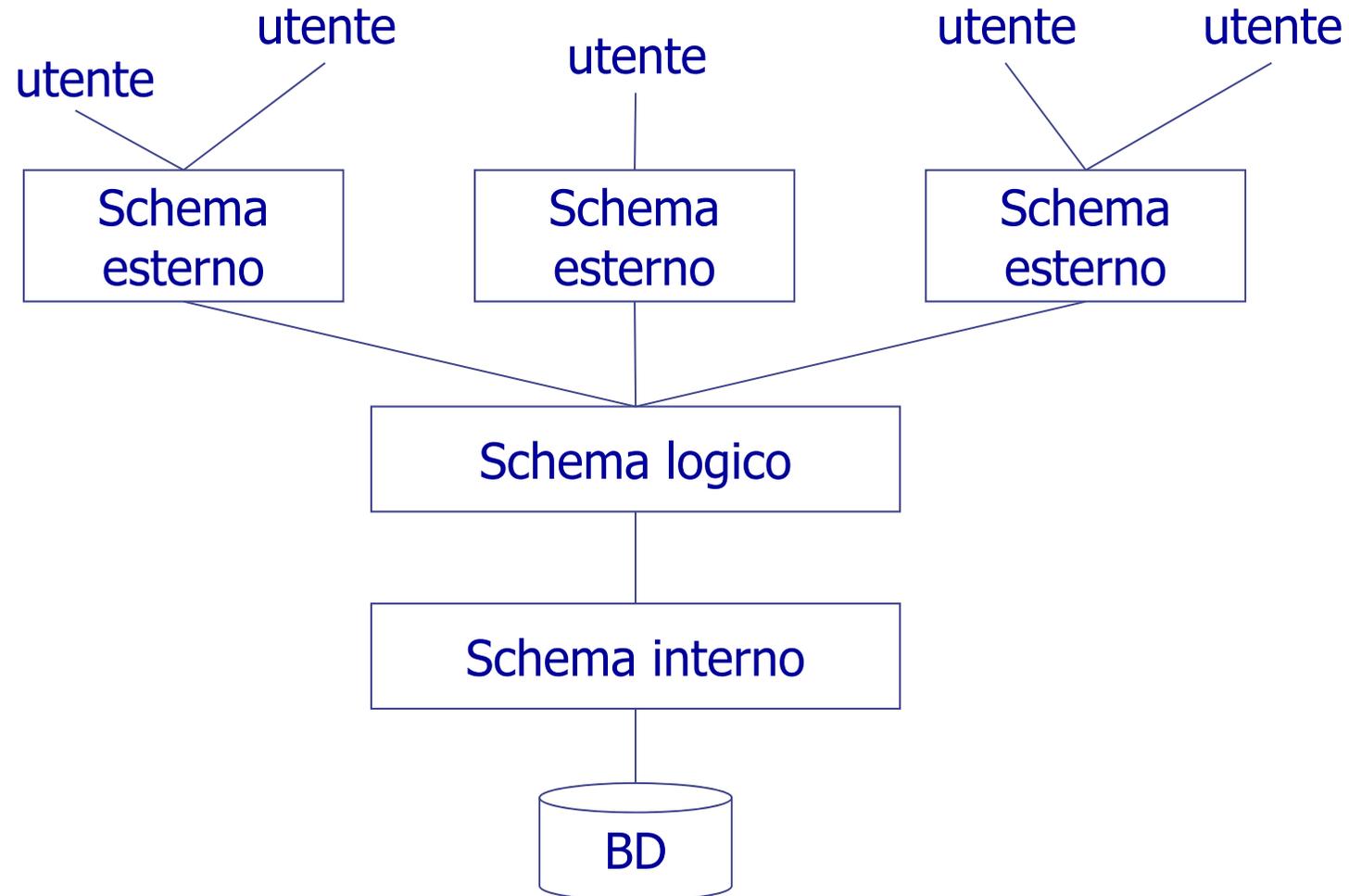
- **schema logico**: descrizione della base di dati nel modello logico (ad esempio, la struttura della tabella)
- **schema interno (o fisico)**: rappresentazione dello schema logico per mezzo di strutture memorizzazione (file; ad esempio, record con puntatori, ordinati in un certo modo)

Indipendenza dei dati

- Il livello 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)
- Perciò ci occuperemo solo del livello logico e non di quello fisico



Architettura standard (ANSI/SPARC) a tre livelli per DBMS



Architettura ANSI/SPARC: schemi

schema logico: descrizione dell'intera base di dati nel modello logico "principale" del DBMS

Schema interno (o fisico): rappresentazione dello schema logico per mezzo di strutture fisiche di memorizzazione

schema esterno: descrizione di parte della base di dati in un modello logico ("viste" parziali, derivate, anche in modelli diversi)

Una vista

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

CorsiSedi

Corso	Aula	Edificio	Piano
Sistemi	N3	OMI	Terra
Reti	N3	OMI	Terra
Controlli	G	Pincherle	Primo

Indipendenza dei dati

- conseguenza della articolazione in livelli
- l'accesso avviene solo tramite il livello esterno (che può coincidere con il livello logico)
- due forme:
 - indipendenza fisica
 - indipendenza logica

Indipendenza fisica

- il livello logico e quello esterno sono indipendenti da quello fisico
 - una relazione è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica
 - la realizzazione fisica può cambiare senza che debbano essere modificati i programmi

Indipendenza logica

- il livello esterno è indipendente da quello logico
- aggiunte o modifiche alle viste non richiedono modifiche al livello logico
- modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti

Linguaggi per basi di dati

- Un altro contributo all'efficacia: disponibilità di vari linguaggi e interfacce
 - ⇒ linguaggi testuali interattivi (**SQL**)
 - ⇒ comandi (SQL) immersi in un linguaggio **ospite** (Pascal, Java, C ...)
 - ⇒ comandi (SQL) immersi in un linguaggio ad hoc, con anche altre funzionalità (p.es. per grafici o stampe strutturate)
 - ⇒ con interfacce amichevoli (senza linguaggio testuale)

SQL, un linguaggio interattivo

- "Trovare i corsi tenuti in aule a piano terra"

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

SQL, un linguaggio interattivo

```
SELECT Corso, Aula, Piano  
FROM Aule, Corsi  
WHERE Nome = Aula  
AND Piano = "Terra"
```

Corso	Aula	Piano
Sistemi	N3	Terra
Reti	N3	Terra

SQL immerso in linguaggio ospite

```
write('nome della citta"?'); readln(citta);
EXEC SQL DECLARE P CURSOR FOR
    SELECT NOME, REDDITO
    FROM PERSONE
    WHERE CITTA = :citta ;
EXEC SQL OPEN P ;
EXEC SQL FETCH P INTO :nome, :reddito ;
while SQLCODE = 0 do begin
    write('nome della persona:', nome, 'aumento?');
    readln(aumento);
    EXEC SQL UPDATE PERSONE
        SET REDDITO = REDDITO + :aumento
        WHERE CURRENT OF P
    EXEC SQL FETCH P INTO :nome, :reddito
end;
EXEC SQL CLOSE CURSOR P
```

SQL in linguaggio ad hoc (Oracle PL/SQL)

```
declare Stip number;
begin
  select Stipendio into Stip from Impiegato
  where Matricola = '575488' for update of Stipendio;
  if Stip > 30 then
    update Impiegato set Stipendio = Stipendio * 1.1
    where Matricola = '575488';
  else
    update Impiegato set Stipendio = Stipendio * 1.15
    where Matricola = '575488';
  end if;
  commit;
exception
  when no_data_found then
    insert into Errori values('Matricola inesistente',sysdate);
end;
```


Una distinzione terminologica (separazione fra dati e programmi)

data manipulation language (DML)

per l'interrogazione e l'aggiornamento di (**istanze** di) basi di dati

data definition language (DDL)

per la definizione di **schemi** (logici, esterni, fisici) e altre operazioni generali

Un'operazione DDL (sullo schema)

```
CREATE TABLE orario (  
    insegnamento    CHAR(20) ,  
    docente          CHAR(20) ,  
    aula             CHAR(4)  ,  
    ora              CHAR(5) )
```

Personaggi e interpreti

- progettisti e realizzatori di **DBMS**
- progettisti della **base di dati** e amministratori della base di dati (DBA)
- progettisti e programmatori di **applicazioni**
- **utenti**
 - utenti **finali** (terminalisti): eseguono applicazioni predefinite (transazioni)
 - utenti **casuali**: eseguono operazioni non previste a priori, usando linguaggi interattivi

Database administrator (DBA)

- Persona o gruppo di persone responsabile del controllo centralizzato e della gestione del sistema, delle prestazioni, dell'affidabilità, delle autorizzazioni
- Le funzioni del DBA includono quelle di progettazione, anche se in progetti complessi ci possono essere distinzioni

Transazioni (per l'utente)

- Programmi che realizzano attività frequenti e predefinite, con poche eccezioni, previste a priori.
- Esempi:
 - versamento presso uno sportello bancario
 - emissione di certificato anagrafico
 - dichiarazione presso l'ufficio di stato civile
 - prenotazione aerea
- Le transazioni sono di solito realizzate in linguaggio ospite (tradizionale o ad hoc)

Transazioni, due accezioni

- Per l'utente:
 - programma a disposizione, da eseguire per realizzare una funzione di interesse
- Per il sistema:
 - sequenza indivisibile di operazioni (cfr. affidabilità)

Vantaggi e svantaggi dei DBMS, 1

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)

Vantaggi e svantaggi dei DBMS, 2

Contro

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