



Corso di Laurea Magistrale in Scienze Infermieristiche e Ostetriche

# Sistemi di Elaborazione delle Informazioni

~ *Metodologia della Ricerca Applicata (R3)* ~  
A.A. 2019/20

**Docente: Gianni Pantaleo**

*Dipartimento di Ingegneria dell'Informazione – DINFO  
DISIT Lab – Distributed Systems and Internet Technology Lab  
Facoltà di Ingegneria  
Università degli Studi di Firenze  
Via S. Marta 3, 50139  
[gianni.pantaleo@unifi.it](mailto:gianni.pantaleo@unifi.it)*

DISIT Lab

<http://www.disit.org>

## *Sistemi di Elaborazione delle Informazioni – A.A. 2019/20*

# Sistemi di Elaborazione delle Informazioni

## **1. Introduzione**

- Definizione di «Informazione»
- Segnali e Codifica dell'Informazione

## **2. Sistemi di Elaborazione dell'Informazione**

- Informatica e Architettura dei Calcolatori
- Reti di Calcolatori e Tecnologie ICT

## **3. Basi di Dati**

- *Definizioni e Rappresentazioni*
- Modelli e Linguaggi di Rappresentazione della Conoscenza

## **4. Sistemi ICT in Ambito Medico Sanitario**

- Sistemi Informativi Sanitari
- Health Technology Assessment
- Direttive IVD - Sistemi Diagnostici in Vitro

## **5. Esercitazioni Pratiche**

Utilizzo di Sistemi di Gestione Dati e Database



## ***3. Basi di Dati***



### 3. Basi di Dati – Definizioni e Rappresentazioni

## La Piramide della Conoscenza - *Informazione*



### 3. Basi di Dati – Definizioni e Rappresentazioni

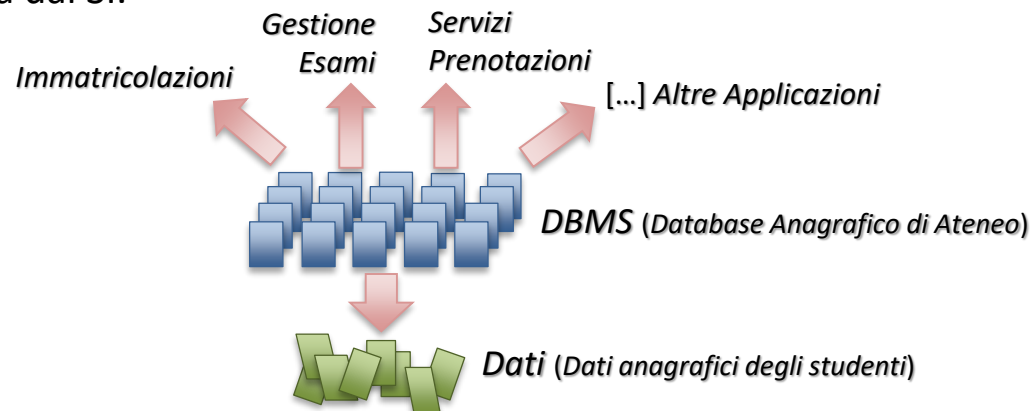
## Basi di Dati e DMBS: Definizioni

- Una **Base di Dati (*Database*)**, nella sua accezione più generica, è una collezione omogenea di dati strutturati.
- Una base di dati è una insieme di dati che può essere gestito da un sistema **DBMS (*Data Base Management System*)**.
- Un **DBMS** è un sistema software composto da sequenze di dati strutturati ed interconnessi tra di loro, e da una serie di programmi applicativi che consentono di effettuare su tali dati operazioni di accesso, modifica e interrogazione (*query* di ricerca).
- Un DBMS supporta diversi livelli di *astrazione sui dati*, a differenza di altri sistemi informativi analoghi non informatizzati (elenco telefonico, registro anagrafico cartaceo ecc.).
  - *Livello Fisico*
  - *Livello Logico/Concettuale*
  - *Livello Esterno*
- Un DBMS deve essere in grado di gestire basi di dati di grandi dimensioni (*Big Data*), *persistenti* (cioè con un ciclo di vita indipendente dal programma che le utilizza) e *condivise* da più utenti/sistemi/applicazioni.
- Un DBMS deve inoltre prevedere modalità di controllo differenziato di accesso ai dati, per garantire la privacy e la sicurezza di eventuali dati sensibili, riservati e/o personali.

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Problematiche di Completezza e Correttezza dei Dati

- Ogni SI è suddiviso generalmente in settori o sotto-sistemi, ognuno dei quali svolge varie funzionalità ed esegue vari processi su una base di dati spesso condivisa.
- Possono essere molteplici le applicazioni che utilizzano gli stessi dati o stesse porzioni della base di dati condivisa dal SI.



- Se la base di dati non è ben progettata, possono verificarsi problemi di integrità, ridondanza o incoerenza, a cui è necessario provvedere attraverso appositi metodi di *riconciliazione* dei dati:
  - **Integrità**: di solito dovuti ad errore umano o incompletezza nell'immissione di dati (referenziati tra di essi) all'interno del database.
  - **Ridondanza**: informazioni ripetute all'interno del database.
  - **Ambiguità**: incoerenza; differenti descrizioni di uno stesso dato o informazione non coincidono.



### 3. Basi di Dati – Definizioni e Rappresentazioni

## DBMS: Vantaggi e Limiti

➤ I vantaggi nell'uso di un DBMS sono molteplici:

- *Accesso ai dati indipendente* dalla loro rappresentazione e memorizzazione.
- *Tecniche di accesso ai dati ottimizzate* (anche attraverso l'uso di strutture aggiuntive, come gli *indici*), per migliorare le performance delle interrogazioni.
- Possibilità di *controllo dell'integrità* e della completezza dei dati.
- Possibilità di *controllo dell'accesso* ai dati e modalità di *accesso multi-utente*.

➤ Alcuni limiti:

- Non permettono visualizzazioni grafiche dei dati o delle strutture dati.
- Implementano funzionalità matematiche e statistiche elementari di base.

### 3. Basi di Dati – Definizioni e Rappresentazioni

## DBMS: Livelli di Astrazione e Rappresentazione

- **Livello Fisico:** rappresentazione del DBMS per mezzo di strutture fisiche di memorizzazione, ovvero specifici files. Descrive come i dati sono memorizzati e organizzati su uno o più supporti di memoria secondaria.
- **Livello Logico:** rappresenta come vengono organizzati i dati secondo il modello logico adottato dal DBMS (*relazionale, gerarchico, ad oggetti ecc.*).
- **Livello Concettuale:** a volte distinto dal Livello Logico, descrive le tipologie di dati da rappresentare, in termini di *oggetti, attributi e relazioni*.
- **Livello Esterno:** descrive come vengono presentate porzioni della base di dati agli utenti, attraverso «*viste*» differenti, parziali o derivate da esigenze particolari degli utenti (regole di accesso, combinazioni di *query* ecc.).





### 3. Basi di Dati – Definizioni e Rappresentazioni

## Tipologie di Utenti

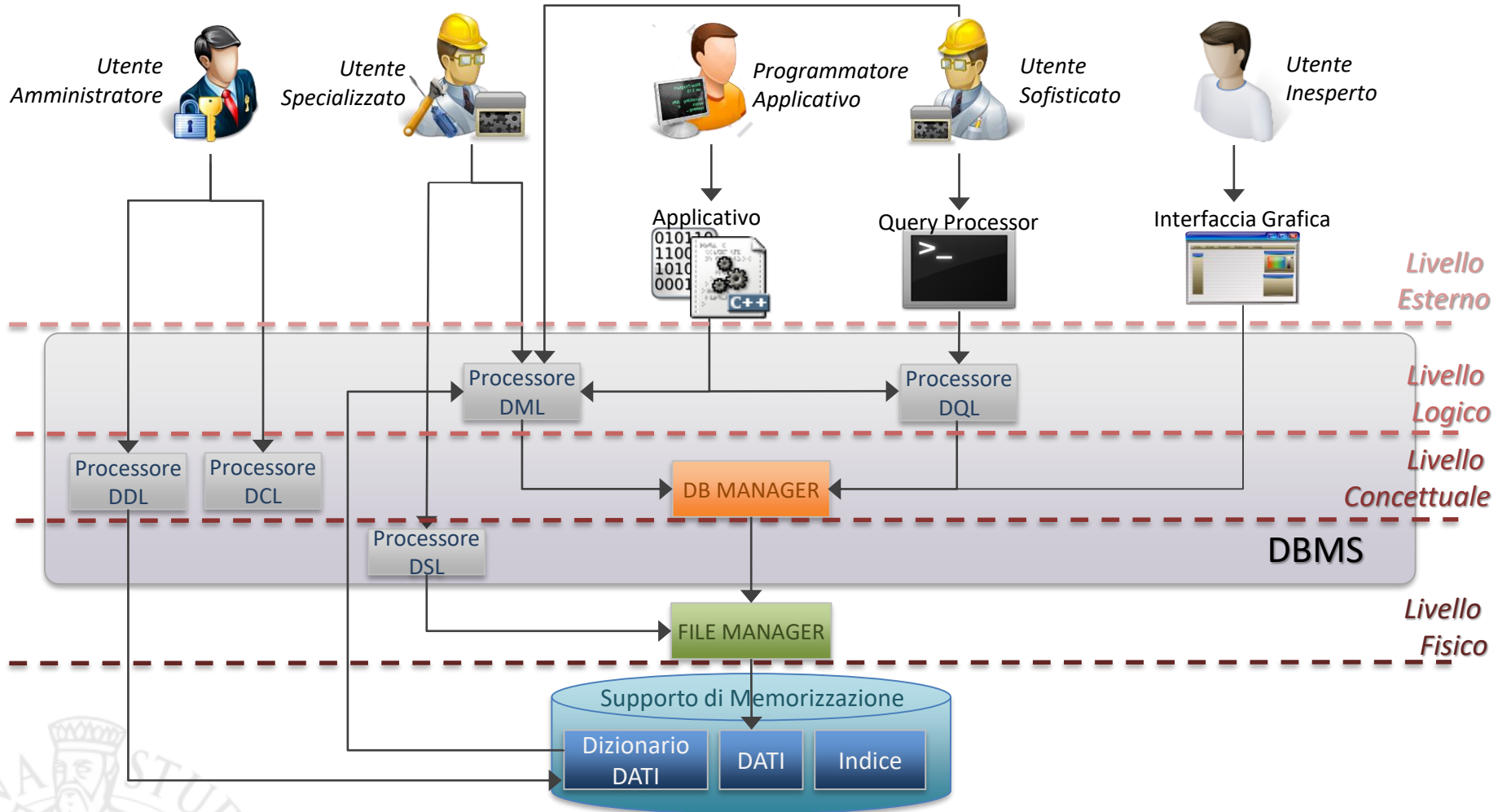
- **DBA - Database Administrator.** Utente amministratore che ha il controllo completo del sistema. Tra i suoi compiti principali:
  - Definizione e modifica degli schemi.
  - Gestione dei metodi di accesso e dei permessi per i vari utenti.
  - Specifica dei vincoli di integrità.

Le altre tipologie di utenti possono essere classificate in base alla loro esperienza nell'uso delle varie funzionalità del sistema.

- **Utente Inesperto.** Utente finale, non ha conoscenza della struttura dei modelli e/o dei linguaggi del sistema, quindi vi accede ed interagisce solo mediante opportune interfacce (maschere) di alto livello fornite dalle applicazioni che dialogano con il DBMS.
- **Programmatore Applicativo.** Progettano applicativi in linguaggi di alto livello, al cui interno viene integrato un DBMS per la gestione dei dati.
- **Utente Sofisticato.** Interviene direttamente, formulando query oppure agendo sui dati.
- **Utente Specializzato.** E' un utente sofisticato che utilizza il sistema con metodi di accesso non convenzionali per scopi di test (prestazioni del DBMS, debug di applicazioni ecc.).

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Architettura di un DBMS



### 3. Basi di Dati – Definizioni e Rappresentazioni

## Linguaggi Utilizzati nei DBMS

Per descrivere i vari livelli funzionali si ricorre all'uso di specifici linguaggi:

- **DQL - Data Query Language.** Utilizzato per interrogare il database, formulando *query* o *combinazioni di query* ed estrarne così l'informazione desiderata (Esempi: *SQL*, *XQuery*).
- **DDL - Data Definition Language.** Utilizzato per la gestione, a livello logico, delle strutture e dei modelli dei dati. Consente di definire la strutture (*Schema*, memorizzati nel *Dizionario dei Dati*) del database Comprende le operazioni di creazione, modifica ed eliminazione di oggetti del database, ma non dei valori assunti dai dati al loro interno (Esempi: *SQL*).
- **DML - Data Manipulation Language.** Utilizzato per la gestione dei dati all'interno degli oggetti del database, senza coinvolgere le strutture o gli schemi a livello logico. Comprende operazioni di inserimento, aggiornamento o eliminazione di singoli dati (Esempi: *SQL*).
- **DCL - Data Control Language.** Utilizzato per definire e gestire le modalità di accesso degli utenti. Comprende i comandi per garantire o revocare l'accesso agli oggetti, ai dati, e all'utilizzo dei vari linguaggi di gestione (Esempi: *SQL*).
- **DSL - Data Storage and Definition Language.** Utilizzato per specificare le strutture e le modalità di memorizzazione a livello fisico.

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Modelli di Dati

- Un **modello di dati** è un insieme di costrutti utilizzati per organizzare i dati e descriverne la struttura. Vengono usati anche per descrivere i livelli di astrazione forniti da un DBMS.
- Una DBMS, per trasformare i dati in informazione, deve necessariamente avere un modello di organizzazione dei dati. I più comunemente utilizzati sono i seguenti:
  - Modello *Reticolare* - *Livello Logico* -
  - Modello *Gerarchico* - *Livello Logico* -
  - Modello *a Oggetti* (object-oriented) - *Livello Logico* -
  - Modello *Relazionale* - *Livello Logico* -
  - Modello *Entità-Relazione E/R* - *Livello Concettuale* -
- Come ogni modello può essere espresso attraverso diagrammi funzionali.



### 3. Basi di Dati – Definizioni e Rappresentazioni

## Modello Relazionale

- Nel Modello Relazionale, una base di dati è rappresentata per mezzo di un insieme di relazioni. Viene usata generalmente la forma *tabellare*.
- Le *colonne* della tabella sono anche dette **campi**, o **attributi** della tabella.
- I campi e il formato di dati da essi rappresentati sono predefiniti nello *Schema* della tabella.
- Le *righe* rappresentano i valori degli attributi, denominati **record**.
- Una tabella deve avere almeno un campo che permetta di identificare ogni record in modo *univoco*: tale campo è definito in fase di progettazione dello Schema come **Chiave Primaria**.

#### Studenti

<i>Cognome</i>	<i>Nome</i>	<i>Data di Nascita</i>	<i>Matricola</i>
Rossi	Mario	01/01/91	5876410
Bianchi	Anna	02/02/92	6029347
Verdi	Luigi	03/03/93	6459382

#### Studenti\_Schema

<i>Nome Campo</i>	<i>Tipo</i>	<i>Codifica</i>	<i>NULL</i>	<i>Primary</i>
<b>Cognome</b>	String(16)	UTF-8	No	No
<b>Nome</b>	String(16)	UTF-8	No	No
<b>Data di Nascita</b>	gg-mm-aa	ASCII	Yes	No
<b>Matricola</b>	INT(7)	ASCII	No	Yes



### 3. Basi di Dati – Definizioni e Rappresentazioni

## Relazioni tra Tabelle

#### Studenti

Cognome	Nome	Data di Nascita	Matricola
Rossi	Mario	01/01/91	5876410
Bianchi	Anna	02/02/92	6029347
Verdi	Luigi	03/03/93	6459382

#### Esami

Matricola Studente	Voto	Lode	Codice Modulo
6459382	28	NULL	B019874
5876410	30	Yes	B019874
6029347	26	NULL	B019872

#### Moduli

Codice	Titolo	Docente	Insegnamento
B019874	Sistemi di Elaborazione delle Informazioni	Gianni Pantaleo	B019867
[...]	[...]	[...]	[...]
B019872	Igiene Generale Applicata	NULL	B019867

#### Docenti

Cognome	Nome	Codice Fiscale	Cineca ID
Pantaleo	Gianni	XXXXXX	113774
Rasero	Laura	YYYYYY	89727
Nesi	Paolo	ZZZZZZ	83324

#### Insegnamenti

Codice	Titolo	Docente
B019867	Metodologia della Ricerca Applicata	Laura Rasero
B003734	Sistemi Distribuiti	Paolo Nesi
[...]	[...]	[...]

Le informazioni contenute in tabelle diverse dello stesso database sono spesso correlate. Le **chiavi esterne** esprimono un vincolo di **integrità referenziale** tra due o più tabelle.





### 3. Basi di Dati – Definizioni e Rappresentazioni

## Il Linguaggio SQL

- **SQL (*Structured Query Language*)** è un linguaggio standard per la definizione e la manipolazione dei dati di DBMS relazionali.
- Originariamente sviluppato presso il laboratorio IBM a San Jose (California) nel 1974, è diventato standard ufficiale nel 1986. L'ultimo standard è SQL:1999 (noto anche come SQL3).
- Tra i più diffusi DBMS relazionali che utilizzano il linguaggio SQL:
  - MySQL
  - Oracle
  - Informix
  - PostgreSQL
  - Microsoft SQLServer
- Permette di creare, eliminare e modificare tabelle; inserire, modificare, eliminare e cercare dati; può essere inoltre utilizzato anche da programmi esterni al DBMS.



### 3. Basi di Dati – Definizioni e Rappresentazioni

## Cenni di Sintassi SQL

- SQL contiene i linguaggi che descrivono tutti i livelli di astrazione di un DBMS.
- Le principali funzionalità (essenzialmente la definizione e la manipolazione delle tabelle di record) sono espletate dai moduli di *Data Definition Language (DDL)* e *Data Manipulation Language (DML)*.
- Con il termine «*query*» non si intende soltanto una generica interrogazione a un database, ma un'istruzione SQL, anche complessa, con cui si possono produrre diverse azioni, tra cui:
  - recuperare informazioni dal database;
  - modificare record esistenti;
  - aggiungere record a una tabella esistente;
  - creare tabelle nuove o modificarne di esistenti.

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Creazione, Uso ed Eliminazione di Un Database

- La creazione di un database si esegue con il comando `CREATE DATABASE`:

```
CREATE DATABASE [db_name];
```

- Esempio:

```
CREATE DATABASE AnagraficaStudenti;
```

- La prima volta dopo la creazione, ed ogni volta che si vogliono eseguire operazioni sul database appena creato, si deve specificare l'intenzione di utilizzo di quest'ultimo con la direttiva `USE [db_name]`;

```
USE AnagraficaStudenti;
```

- L'eliminazione di un database si esegue con il comando `DROP DATABASE`:

```
DROP DATABASE [db_name];
```



### 3. Basi di Dati – Definizioni e Rappresentazioni

## Creazione, Eliminazione e Svuotamento di Una Tabella

- La creazione di una tabella si esegue con il comando `CREATE TABLE`:

```
CREATE TABLE [table_name]  
(  
    [column_name1] [data_type(size)],  
    [column_name2] [data_type(size)],  
    [column_name3] [data_type(size)],  
    ....  
);
```

```
CREATE TABLE Studente  
(  
    NumMatricola int(11),  
    Cognome varchar(255),  
    Nome varchar(255),  
    DataNascita date,  
);
```

```
CREATE TABLE Studente  
(  
    NumMatricola NOT NULL PRIMARY KEY int(11),  
    Cognome NOT NULL varchar(255),  
    Nome NOT NULL varchar(255),  
    DataNascita DEFAULT NULL varchar(255),  
);
```

Il tipo `varchar(255)` identifica  
una stringa di 256 caratteri.



- Per eliminare completamente la tabella (dati e schema): `DROP TABLE Studenti;`
- Per svuotare la tabella senza eliminare lo schema: `TRUNCATE TABLE Studenti;`

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Modifica dello Schema di Una Tabella

- Comando base per modificare la schema di una tabella: `ALTER TABLE`:
- Per aggiungere una colonna aggiungere la direttiva `ADD COLUMN [column_name]`:  

```
ALTER TABLE [table_name] ADD COLUMN [column_name] [data_type(size)];  
ALTER TABLE Studenti ADD COLUMN Indirizzo varchar(255);
```
- Per eliminare una colonna aggiungere la direttiva `DROP COLUMN [column_name]`:  

```
ALTER TABLE Studenti DROP COLUMN Indirizzo;
```
- Per modificare tipo di dati una colonna:  

```
ALTER TABLE [table_name] MODIFY COLUMN [column_name] [data_type];  
ALTER TABLE Studenti MODIFY COLUMN Cognome varchar(31);
```



### 3. Basi di Dati – Definizioni e Rappresentazioni

## Modifica dei Record di Una Tabella (1 di 3)

➤ Per inserire un record si usa il comando `INSERT INTO TABLE`:

```
INSERT INTO [table_name] ([column_name1], [column_name2], ...)  
VALUES ([value1], [value2], ...);
```

Ogni formato di dato da inserire  
va racchiuso tra apici '...'



```
INSERT INTO Studenti (NumMatricola, Cognome, Nome)  
VALUES ('5876410', 'Rossi', 'Mario');
```

<i>NumMatricola</i>	<i>Cognome</i>	<i>Nome</i>	<i>DataNascita</i>
5876410	Rossi	Mario	NULL

Nella definizione dello schema della  
Tabella *Studente*, era stato definito l'attributo  
`DEFAULT NULL` per il campo `DataNascita`





### 3. Basi di Dati – Definizioni e Rappresentazioni

## Modifica dei Record di Una Tabella (2 di 3)

➤ Per modificare/aggiornare un record si usa il comando UPDATE:

```
UPDATE [table_name]
SET [column_name1] = [value1], [column_name2] = [value2], ...
WHERE [una_certa_Colonna] = [un_certo_valore] AND|OR
      [una_altra_certa_Colonna] = [un_altro_certo_valore];
```



Espressione condizionale per la presenza del costrutto WHERE e degli operatori logici AND e OR

```
UPDATE Studenti
SET DataNascita = '1991-01-01'
WHERE Cognome='Rossi' AND Nome='Mario';
```

```
UPDATE Studenti
SET DataNascita = '1991-01-01'
WHERE NumMatricola = '5876410';
```

NumMatricola	Cognome	Nome	DataNascita
5876410	Rossi	Mario	1991-01-01



Questa direttiva alternativa produce lo stesso risultato, ma elimina i rischi di ambiguità, usando nella espressione condizionale una PRIMARY KEY !



Attenzione! Se si omette la clausola WHERE verranno modificati tutti i record!

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Modifica dei Record di Una Tabella (3 di 3)

- Per eliminare uno o più record si usa il comando DELETE:

```
DELETE FROM [table_name]
WHERE [una_certa_Colonna] = [un_certo_valore] AND|OR
      [una_altra_certa_Colonna] = [un_altro_certo_valore];
```

```
DELETE FROM Studenti
WHERE DataNascita IS NULL;
```

- Per eliminare tutte le righe da una tabella (ovvero tutti i record, il che equivale allo svuotamento della tabella), in alternativa al comando TRUNCATE TABLE, si può eseguire il comando DELETE con la seguente sintassi:

```
DELETE FROM [table_name];
```

### 3. Basi di Dati – Definizioni e Rappresentazioni

## Ricerche all'Interno di Una Tabella

- Per eseguire ricerche all'interno di una tabella si usa il comando SELECT:

```
SELECT [column_name1], [column_name2], ... FROM [table_name]
WHERE [una_certa_Colonna] = [un_certo_valore] AND|OR ...
ORDER BY [column_name1] ASC|DESC;
```

### Esami

NumMatricola	Voto	Lode	CodModulo
6459382	28	NULL	B019874
5876410	30	Yes	B019874
6029347	26	NULL	B019872

```
Q:> SELECT NumMatricola
      FROM Esami WHERE Voto > 28
      AND CodModulo = 'B019874';
```

```
A:> 5876410
```

- Per selezionare TUTTI i campi di un record, si usa il comando SELECT e il carattere speciale \* per specificare la selezione di TUTTE le colonne:



```
Q:>SELECT * FROM Studenti WHERE CodModulo = 'B019874' ORDER BY Voto DESC;
```

```
A:> 5876410      30      Yes      B019874
      6459382      28      NULL     B019874
```

## *Sistemi di Elaborazione delle Informazioni – A.A. 2019/20*

# Sistemi di Elaborazione delle Informazioni

## **1. Introduzione**

- Definizione di «Informazione»
- Segnali e Codifica dell'Informazione

## **2. Sistemi di Elaborazione dell'Informazione**

- Informatica e Architettura dei Calcolatori
- Reti di Calcolatori e Tecnologie ICT

## **3. Basi di Dati**

- Definizioni e Rappresentazioni
- *Modelli e Linguaggi di Rappresentazione della Conoscenza*

## **4. Sistemi ICT in Ambito Medico Sanitario**

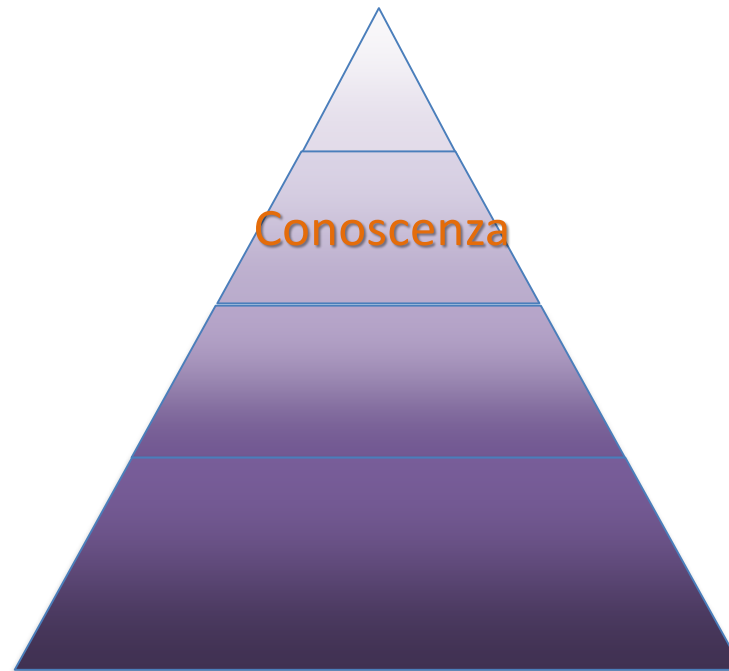
- Sistemi Informativi Sanitari
- Health Technology Assessment
- Direttive IVD - Sistemi Diagnostici in Vitro

## **5. Esercitazioni Pratiche**

Utilizzo di Sistemi di Gestione Dati e Database

### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## La Piramide della Conoscenza - Conoscenza



### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Il Modello Entità-Relazione E/R

- Il modello **Entità-Relazione** è un modello *concettuale*. I modelli concettuali permettono di rappresentare i dati in modo *indipendente* da ogni sistema, cercando di descrivere i concetti del mondo reale.
- I modelli concettuali sono utilizzati nelle fasi di *progettazione* di un DBMS.
- Il modello **Entità-Relazione** si basa su tre elementi:
  - Le *Entità*
  - Le *Relazioni (o Associazioni)*
  - Gli *Attributi*



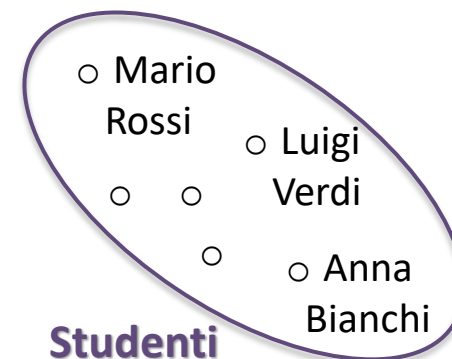
### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Le Entità

- Un' **Entità** rappresenta una classe di oggetti concreti del mondo reale (e.g., persone, oggetti, luoghi ecc., o anche astratti, che abbiano proprietà comuni, ed esistenza autonoma.
- Un tipo di Entità viene di solito definito attraverso un nome e lo si rappresenta graficamente con un rettangolo.
- Un oggetto specifico appartenente ad una Entità viene definito **Istanza** di quell'Entità.



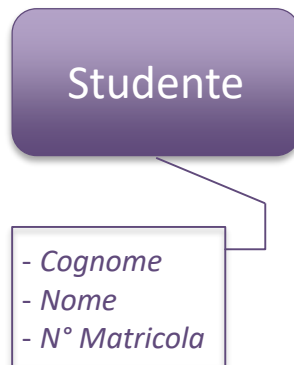
- Mario Rossi
- Anna Bianchi
- Luigi Verdi
- Xxxx Yyyy



### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Gli Attributi

- Gli **Attributi** descrivono le *proprietà* delle Entità.
- Gli Attributi sono caratterizzati da i seguenti parametri:
  - **Formato**: indica il formato dell'attributo (numero, stringa, formati data *gg-mm-aaaa* ecc.).
  - **Dimensione**: specifica il range entro cui varia l'attributo (a seconda del formato), e di conseguenza il numero di caratteri o cifre per rappresentarlo.
  - **Opzionalità**: indica se la sua presenza è necessaria oppure facoltativa.

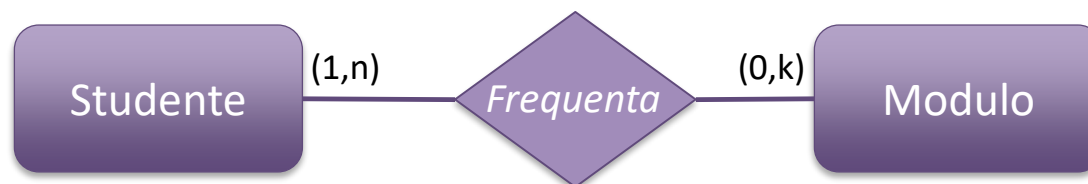


<i>Attributo</i>	<i>Formato</i>	<i>Dimensione</i>	<i>Opzionalità</i>
Cognome	<i>Stringa</i>	String(16)	<i>Opzionale</i>
Nome	<i>Stringa</i>	String(16)	<i>Opzionale</i>
Matricola	<i>Numero</i>	INT(7)	<i>Necessario</i>

### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Le Relazioni (1 di 3)

- Una **Relazione** è il legame che definisce un'interazione tra due Entità.

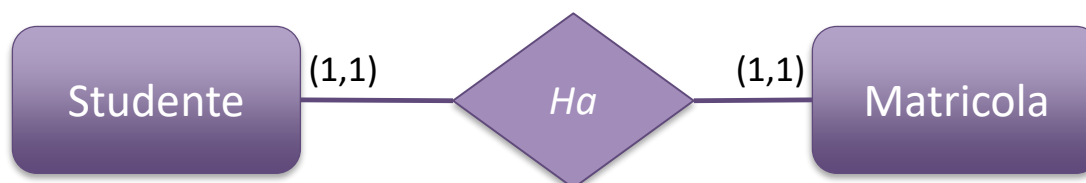


- Ogni Relazione può avere uno o due *versi* di percorrenza, a seconda che esista e sia definita o meno la sua *Relazione Inversa*. Per ogni verso di percorrenza si ha un'Entità di partenza e un'Entità di arrivo.
- Ogni Relazione è anche identificata da una *cardinalità* minima e massima.
- Le Relazioni permettono di esplicitare i collegamenti logici esistenti fra le varie Entità.

### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Le Relazioni (2 di 3) – Cardinalità

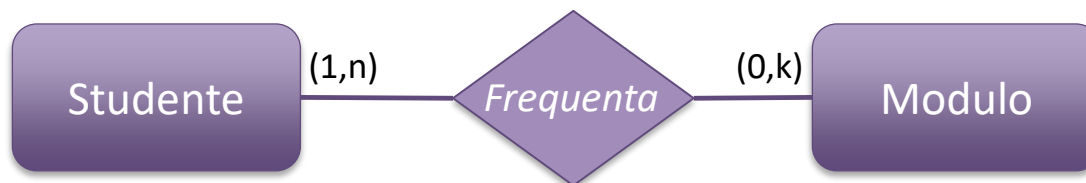
#### ➤ Cardinalità **Uno-A-Uno**:



#### ➤ Cardinalità **Uno-A-Molti** o **Molti-A-Uno**:



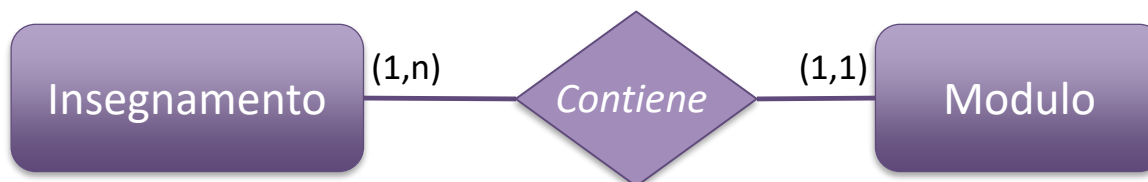
#### ➤ Relazione **Molti-A-Molti** :



### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

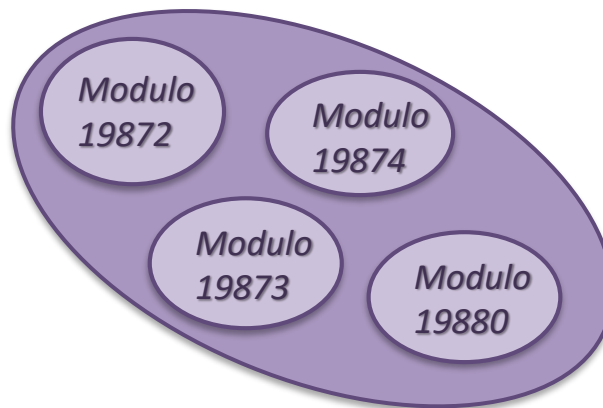
## Le Relazioni (3 di 3)

- Se un'entità rappresenta una classe di oggetti del mondo reale, è possibile che essa debba essere «*specializzata*» o «*generalizzata*» in diversi contesti.



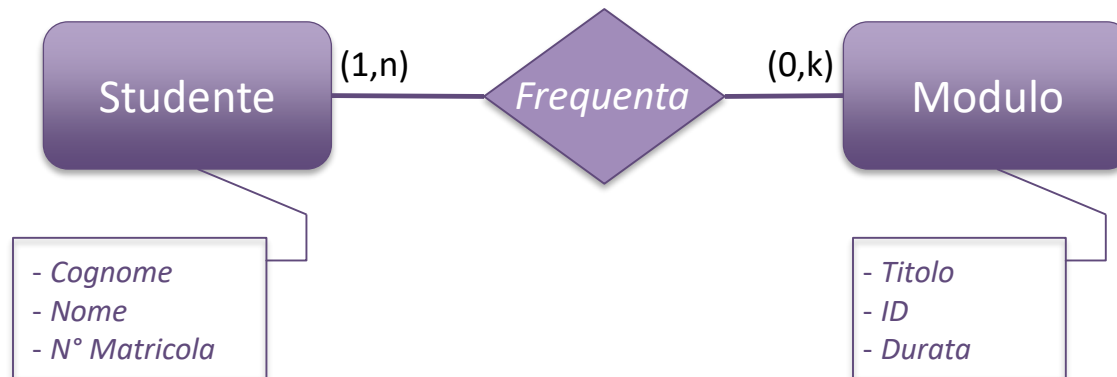
- Le Relazioni di generalizzazione e specificazione tra Entità richiama un parallelo con la proprietà insiemistica dell'*Inclusione*, e anche con il concetto di *Ereditarietà* nella programmazione ad oggetti (*Superclass* e *Subclass*).

**Insegnamento**  
**B019867**



### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

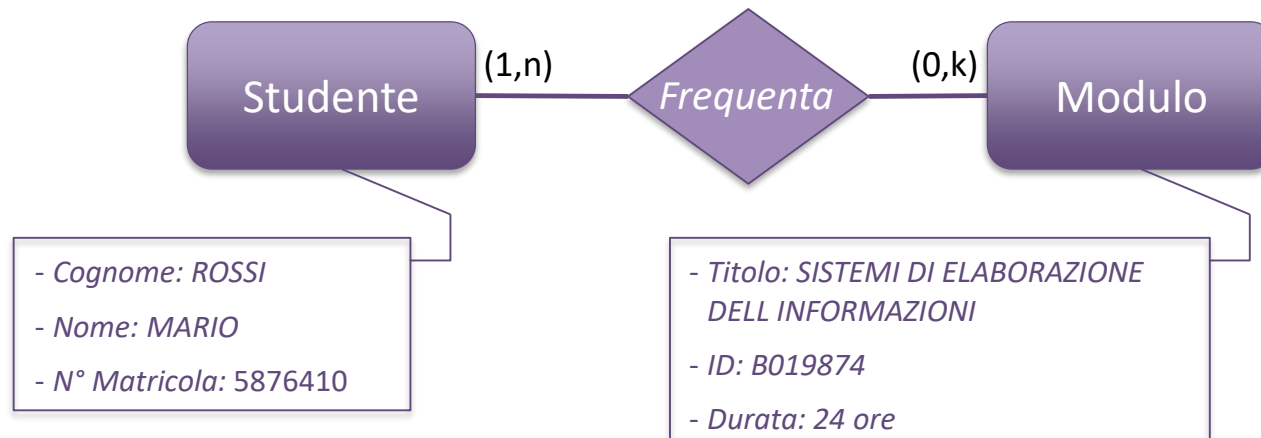
## Un Esempio





### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Un Esempio



### 3. Basi di Dati – Modelli di Rappresentazione della Conoscenza

## Linguaggi per Rappresentare la Conoscenza: XML

- **XML (eXtensible Markup Language)** è un meta-linguaggio che consente la creazione di linguaggi di mark-up (come HTML).
- XML consente la rappresentazione di documenti e dati strutturati su supporto digitale.
- XML è uno dei più potenti e versatili sistemi per la creazione, archiviazione, preservazione e disseminazione di documenti digitali.
- XML è indipendente dal tipo di piattaforma hardware e software su cui viene utilizzato.
- XML è indipendente dai dispositivi di archiviazione e visualizzazione:
  - un documento XML può essere archiviato su qualsiasi tipo di supporto digitale;
  - un documento XML può essere visualizzato su qualsiasi dispositivo di output;
- Concretamente, un documento XML è un file di testo che contiene una serie di tag, attributi e testo, secondo regole sintattiche ben definite.

### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Esempio di Documento XML

```
<?xml version="1.0" ?>
```

```
<articolo titolo="Titolo dell'articolo">
```

```
<paragrafo titolo="Titolo del primo paragrafo">
```

```
<testo>
```

```
  Blocco di testo del primo paragrafo
```

```
</testo>
```

```
<immagine file="immagine1.jpg"></immagine>
```

```
</paragrafo>
```

```
<paragrafo titolo="Titolo del secondo paragrafo">
```

```
<testo>
```

```
  Blocco di testo del secondo paragrafo
```

```
</testo>
```

```
<codice>
```

```
  [Esempio di codice del Programma]
```

```
</codice>
```

```
<testo>
```

```
  Altro blocco di testo
```

```
</testo>
```

```
</paragrafo>
```

```
<paragrafo tipo="bibliografia">
```

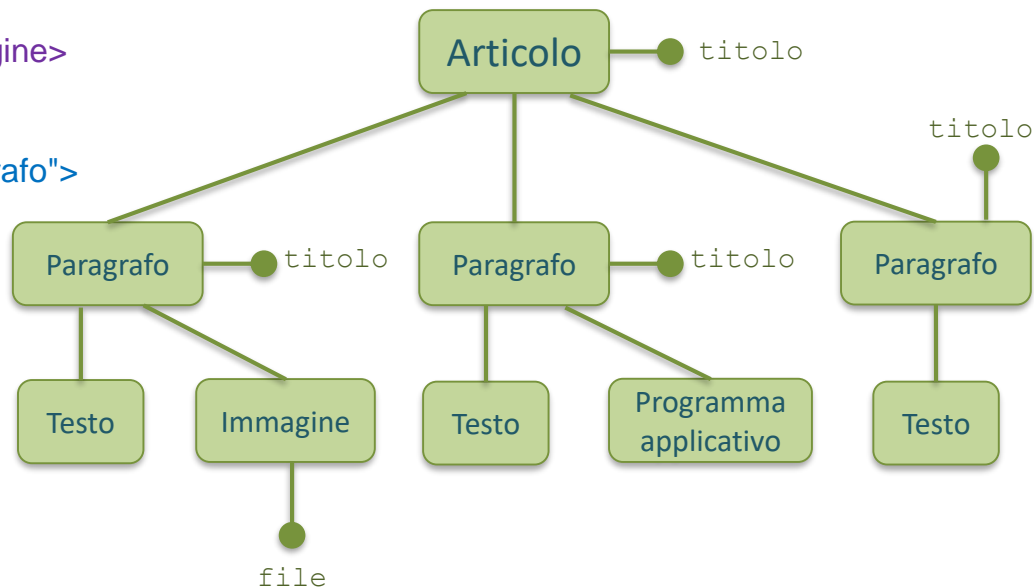
```
<testo>
```

```
  Riferimento ad un articolo
```

```
</testo>
```

```
</paragrafo>
```

```
</articolo>
```



### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Applicazioni di XML: il Web Semantico

- «*Semantic Web*» è un termine coniato da Tim Berners-Lee:

*“Ho un sogno per il Web [in cui i computer] diventeranno capaci di analizzare tutti i dati sul Web - il contenuto, i link e le interazioni tra persone e macchine. Un **Web Semantico**, che dovrebbe rendere questo possibile, deve ancora emergere, ma quando succederà, i quotidiani meccanismi del commercio, della burocrazia e la nostra vita quotidiana saranno affidati a macchine che comunicano con altre macchine. Gli "agenti intelligenti" che sono stati per anni propagandati, finalmente si materializzeranno.”*

Tim Berners Lee, 1999

- Estende l'attuale World Wide Web da *machine-representable* a *machine-understandable*.
- Non lo sostituisce... ma lo migliora associando informazioni di carattere semantico alle risorse della rete.
- Il Semantic Web è una ragnatela di informazioni connesse tramite relazioni semantiche.
- Logiche di connessione più elaborate rispetto ai collegamenti ipertestuali di HTML.
- L'idea è di generare dei documenti che possano essere letti al tempo stesso da esseri umani e da agenti automatici.
- Massiccio utilizzo di linguaggi e grammatiche XML per rendere il formato machine-readable.

### 3. Basi di Dati – Modelli e Linguaggi di Rappresentazione della Conoscenza

## Limiti del linguaggio XML

- Nell'ambito del *Web Semantico*, le applicazioni devono essere in grado di interagire ed essere *interoperabili* tra di loro. Questo richiede uniformità di linguaggio e/o di regole.
- Le tecnologie basate su XML sono necessarie ma non sufficienti: infatti XML codifica e standardizza la **sintassi** dei dati ma non la loro **semantica**.
- Ad esempio, due documenti XML possono contenere le seguenti stringhe:

"<spesa>85</spesa>"      "<costo>85</costo>"

- Una applicazione informatica non può sapere che si tratta di due diverse codifiche della stessa informazione, a meno che non sia disponibile la conoscenza del fatto che “prezzo” e “costo” sono sinonimi (conoscenza terminologica), almeno nel contesto dell'applicazione in esame.

### 3. Basi di Dati – Riferimenti e Link Utili

## Riferimenti e Link Utili

- ❖ D. Sciuto, G. Buonanno, L. Mari, «Introduzione ai sistemi informatici». McGraw Hill (4° Ed., 2008)
- ❖ P. Vittorini, «L'Informatica per la medicina e la sanità pubblica», Ed. L'Una, 2009.
- ❖ A. Rosotti, «Informatica medica». McGraw Hill (Ed. aggiornata 2018)
- ❖ Documentazione W3C su XML: <http://www.w3.org/XML>
- ❖ Documentazione W3C su RDF: <http://www.w3.org/RDF>
- ❖ Documentazione W3C su OWL: <http://www.w3.org/OWL>
- ❖ OSIM – Open Space Innovative Mind (DISIT Lab – UniFI): <http://openmind.disit.org>