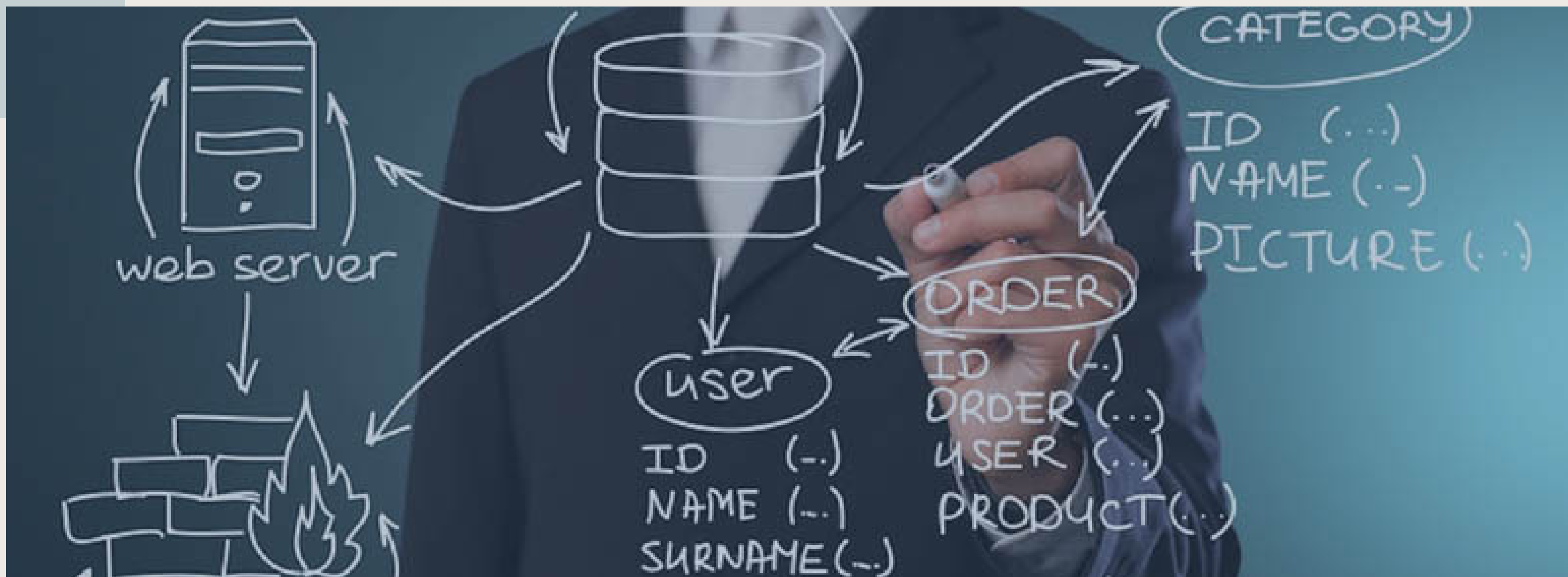


# DATABASE

## LEZIONE 6

ITIS PAOLO SARPI  
San Vito al Tagliamento  
INFORMATICA

**CRISTIAN VIRGILI**



# Individuazione degli oggetti del diagramma

Le prime operazioni che il modellatore esegue hanno il compito di **classificare gli oggetti come entità oppure attributi** partendo dalla documentazione del progetto.

Questa fase viene realizzata mediante due operazioni:

1. raccolta e **analisi della documentazione**
2. definizione del **glossario dei termini.**

# Analisi della documentazione

Durante l'**analisi dei requisiti** viene prodotta e analizzata una serie di **documenti**, che può essere classificata per **tipologia**.

- **Documentazione specifica prodotta per il progetto**, che comprende:
  - le note delle riunioni tecniche e le richieste del cliente;
  - gli appunti sulle interviste agli utenti finali;
  - la documentazione scritta predisposta appositamente.
- **Documentazione esistente**, composta da:
  - le normative generali e del settore;
  - i regolamenti interni;
  - le procedure aziendali.
- **Sistema esistente**, in particolare:
  - la documentazione del sistema da rimpiazzare;
  - le specifiche di integrazione con sistemi esistenti.

# GLOSSARIO DEI TERMINI

Bisogna **eliminare le ambiguità del linguaggio naturale**, e si deve costruire, a partire dalle frasi, un glossario di termini che, ad esempio, contenga per ogni termine la descrizione, l'elenco dei sinonimi e l'elenco dei termini a cui si collega

Su di esso e/o mediante esso si eseguono due controlli.

1. Si verifica la **completezza** del glossario, cioè se tutti gli aspetti importanti sono stati considerati.
2. Si verifica la **consistenza** delle specifiche, se cioè tutti i termini sono stati definiti.

Esempio

Supponiamo di dover analizzare il funzionamento di un magazzino. Alcune frasi sintetiche che ne forniscono una descrizione sono le seguenti.

- Il magazzino è composto da scaffali.
- Gli scaffali contengono prodotti.
- I fornitori forniscono prodotti.
- Gli operai sono addetti agli scaffali.
- I clienti ordinano prodotti.

Se ne può ricavare il seguente **glossario**, che riportiamo in tabella.

Termine	Descrizione	Sinonimo	Legame
Fornitore	Partita IVA Denominazione Indirizzo Numero telefonico ...		Prodotto
Cliente	Partita IVA Denominazione Indirizzo Numero telefonico ...	Acquirente	Prodotto

Termine	Descrizione	Sinonimo	Legame
Prodotto	Codice Nome Genere ...	Articolo Voce	Fornitore Cliente Scaffale
Scaffale	Supporto numerico	Ripiano Armadio	Operaio Prodotto
Operaio	Dati anagrafici Matricola Qualifica	Addetto Magazziniere	Scaffale



# Definizione delle entità e degli attributi

## Individuare le entità

Le **entità** sono i primi elementi da determinare nella progettazione concettuale di un database ed è possibile individuarle aiutandosi con alcune definizioni e osservazioni.

Viene considerata un'**entità**:

- qualunque oggetto per il quale è necessario salvare alcuni **attributi**;
- qualsiasi oggetto che deve essere rappresentato in un **database**.

In particolare, sono **entità**:

- qualsiasi persona, luogo, cosa, evento o concetto distinguibile, sul conto del quale le informazioni vanno memorizzate;
- qualsiasi cosa per la quale salviamo informazioni (ad esempio fornitori, macchine, impiegati, numero posti in aereo ecc.).

## Definire gli attributi

Gli **attributi**, come abbiamo già detto, sono gli oggetti che **descrivono un'entità**: corrispondono ai campi dei record.

1. Gli attributi devono essere **atomici**, ovvero presentare un singolo fatto o una singola informazione, dove per singola informazione si intende proprio la più piccola informazione salvabile.

**Non utilizzare i dati in forma atomica** porta alla generazione di un **insieme di errori**: illustriamo di seguito i più frequenti.

**Aggregazioni semplici**: ad esempio concatenare il nome e il cognome, è un caso di aggregazione semplice.

**Codici complessi**: alcuni attributi hanno per valore codici composti di parti di informazioni concatenate che nella vita quotidiana sono utilizzate assieme, ma che risulta opportuno scomporre nei database. Un esempio di codice complesso è il numero telefonico completo di prefisso, che deve essere spezzato in due attributi.

**Attributi testuali:** ci sono attributi che **vengono definiti erroneamente di tipo testo** perché questa modalità di rappresentazione risulta essere la più semplice. Un esempio classico è quello di definire le date (gg-mm-aaaa) con attributi di tipo testo, e dover poi eseguire su di esse delle operazioni.

**Attributi numerici:** ci sono attributi che **vengono definiti erroneamente di tipo numerico** perché questa modalità di rappresentazione “sembra essere” la più indicata. Un esempio classico è quello di definire di tipo intero il CAP, la partita IVA, il numero civico, il numero di telefono ecc., quando invece andrebbero definiti di tipo testo, in quanto su di essi non verranno mai effettuate operazioni matematiche!

## 2. Gli attributi **derivati** non dovrebbero essere memorizzati



### 3. **Utilizzare codici** per classificare gli attributi ogni volta che si presenta tale opportunità .

Un valore codificato usa un numero ridotto di lettere o numeri per rappresentare un dato: ad esempio, i dati da memorizzare per il campo sesso potrebbero essere semplicemente le lettere M e F, piuttosto che la scritta “Maschio” o “Femmina”

Inoltre, è anche importante classificare gli **attributi** considerando se la situazione fotografata è:

- **permanente**, cioè vale per ogni giorno e mese dell'anno;
- **temporanea**, cioè se i dati variano in funzione del tempo, come i turni settimanali, le sospensioni estive, le fatture dell'anno in corso, i giustificativi per la dichiarazione dei redditi ecc.

## Nominare gli oggetti

Per non correre il rischio di trovarsi in presenza di **omonimi**, le definizioni devono essere più complete e accurate possibile affinché tutte le parti che concorrono alla modellazione dei dati sappiano quali concetti gli **oggetti** vogliono rappresentare.

## Scelta dei nomi

Tutti gli oggetti che fanno parte del nostro modello “dovrebbero” avere un nome e anche questo aspetto deve essere regolato da una serie di proprietà.

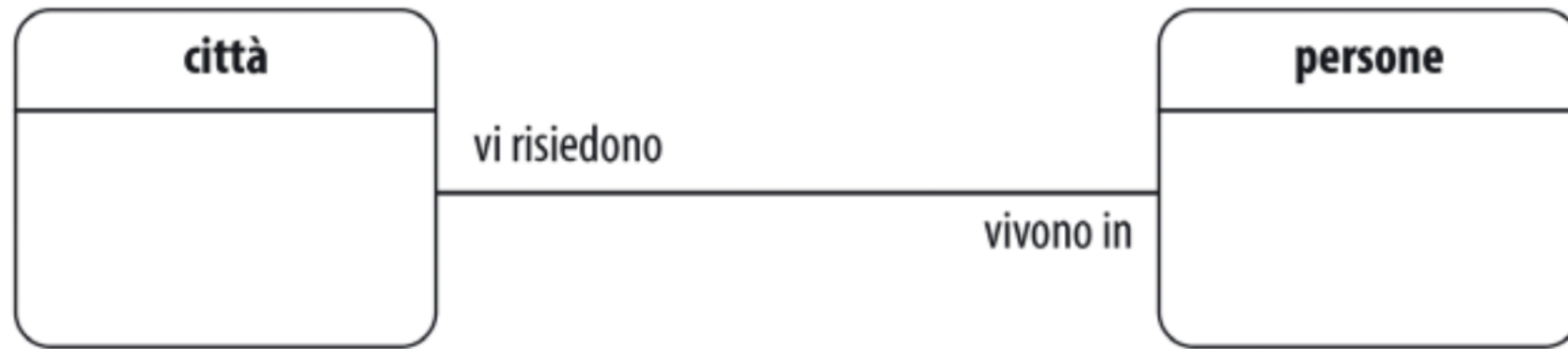
I **nomi** degli oggetti devono:

- essere **unici**;
- avere un **significato** per l'utente finale;
- contenere un **numero minimo di parole** di cui si ha bisogno **per descrivere univocamente** e accuratamente **l'oggetto**.

Per le **entità** e gli **attributi** i **nomi** sono usati generalmente **al singolare**, mentre le **relazioni** sono tipicamente descritte attraverso **verbi**.

## Esempio

Una semplice relazione tra *città* e *persone* può essere così rappresentata in **UML**.



Oppure con un unico termine per la relazione nel diagramma E-R.



## Documentare il progetto: matrici tra entità e attributi

Oltre al **glossario**, per tenere traccia degli oggetti che faranno parte del database si possono usare due schemi: la matrice **entità-entità** e la matrice **entità-attributo**.

Nel dettaglio:

- la **matrice entità-entità** è una tabella bidimensionale per indicare le **relazioni tra le entità**. I nomi di tutte le entità identificate sono elencate sui due assi (verticale e orizzontale). Una relazione viene indicata con una “X” posizionata nelle caselle di intersezione tra due entità. Questa tabella sarà il punto di partenza per la fase di progettazione concettuale dove ogni relazione deve essere descritta mediante uno **schema E-R**;
- la **matrice entità-attributo** viene usata per indicare l'**assegnazione degli attributi alle entità**. È simile nella forma alla matrice entità-entità, tranne che i nomi degli attributi sono elencati in righe: questa tabella viene utilizzata successivamente, in fase di definizione del modello logico, dove per ogni tabella vengono indicati i singoli campi (**tracciato record**).



### Esempio

Si vogliono definire le associazioni e gli attributi per organizzare i dati relativi ai **dipendenti** di una società allo scopo di gestire le informazioni concernenti i **progetti** seguiti dagli **impiegati**, le **attività** da svolgere all'interno di un progetto e le **conoscenze** necessarie per i progetti.

Matrice entità-entità				
	Impiegato	Progetto	Attività	Conoscenze
Impiegato		X		
Progetto	X		X	X
Attività		X		
Conoscenze	X			

Matrice entità-attributo				
	Impiegato	Progetto	Attività	Conoscenze
Matricola impiegato	X			
Nome impiegato	X			
Mansione				
Codice progetto		X		
Nome progetto		X	X	
Attività			X	
Descrizione attività			X	
Conoscenza				X
Specializzazione				X



## ■ Individuazione delle relazioni

Come abbiamo già detto, per individuare le **relazioni** si analizzano i verbi che figurano nelle frasi scritte in fase di analisi, dato che ogni relazione è indicata generalmente da un **verbo** che permette di **connettere due o più entità**.

A partire dalla matrice **entità-entità** si deve:

- definire e individuare tutte le **relazioni**;
- classificare le **relazioni** in termini di:
  - **cardinalità**, per quantificare le **relazioni** tra le entità misurando come molte istanze di una entità sono relazionate a singole istanze di un'altra entità;
  - **opzionalità**, per verificare se la relazione deve esistere oppure è opzionale;
  - **direzione**, per individuare la direzione della relazione (ad esempio da **studente** a **voto**);
  - **dipendenza**, per stabilire se una relazione dipende da una o più altre relazioni.

Una volta individuate le entità, si definiscono le possibili **gerarchie** in termini di:

- **generalizzazione**, se un caso è più generale di un altro;
- **specializzazione**, se un caso è più particolare di un altro.

## Attributi delle relazioni

La rappresentazione grafica dello **schema** ci permette di poter aggiungere gli eventuali **attributi** nel caso che questi siano legati specificamente alla relazione e non alle entità. Vediamo un primo esempio.

### Esempio

Descriviamo la relazione che deriva dalla frase “*lo studente è interrogato in una materia*”. Se analizziamo la **relazione**, che in questo circostanza è espressa da **interrogazione**, ci rendiamo immediatamente conto che per specificare l’interrogazione è necessario indicare in quale data si è svolta e il voto che è stato attribuito: sono due **attributi** legati proprio alla **relazione**, in quanto non appartengono né allo *studente* né alla *materia*.



## ■ Conclusione

Lo **schema E-R** deve essere verificato accuratamente affinché risponda a requisiti di:

- **correttezza**: non devono essere presenti errori (sintattici o semantici);
- **completezza**: tutti i dati di interesse sono specificati;
- **leggibilità**: anche in relazione ad aspetti prettamente estetici dello schema;
- **minimalità**: è importante capire se esistono elementi ridondanti nello schema e se queste situazioni costituiscono un problema oppure sono dovute a una scelta di progettazione volta a favorire l'esecuzione di certe operazioni.

**Domande?**