

04. Le basi di dati relazionali

database

le basi di dati relazionali

Si deve realizzare un sistema informatico per gestire i dati di una piccola biblioteca; in questo scenario si possono individuare alcune ***classi di dati fondamentali*** insieme ai loro ***attributi***, cioè **le proprietà caratteristiche che li descrivono**:

- per i **libri**: il codice ISBN, il titolo, l'anno di pubblicazione, la lingua, il prezzo;
- per gli **autori**: il cognome e il nome, la nazionalità, la data di nascita, il sesso;
- per gli **editori**: la ragione sociale, l'indirizzo, la città;
- per i **generi**, cioè le categorie con cui si possono classificare i libri (romanzi, gialli, saggi, ecc.): la descrizione del genere stesso.

le basi di dati relazionali

Inoltre possiamo fare le seguenti osservazioni:

- un autore può aver scritto più di un libro, anche se ogni libro è stato scritto da un solo autore (ignoriamo per ora situazioni in cui un libro può essere scritto da più autori, come spesso succede per i libri di genere tecnico-scientifico);
- un editore normalmente pubblica numerosi libri, mentre un libro può essere pubblicato da un solo editore;
- ogni genere comprende un intero insieme di libri, ma ogni libro appartiene ad un solo genere.

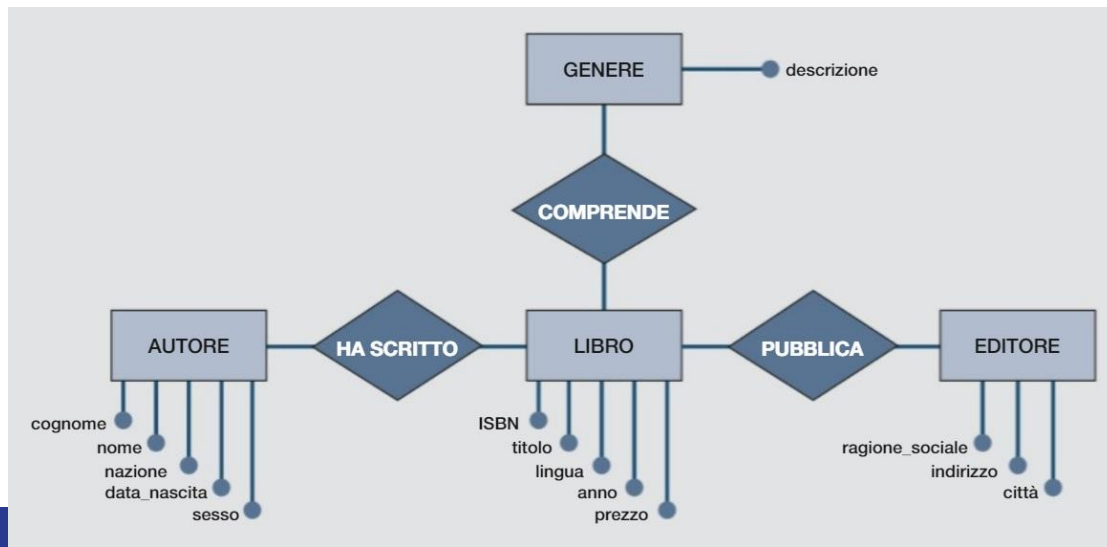
In pratica, dopo aver individuato le classi di dati (*entità/concetti*) che ci interessano, cominciamo a ragionare su eventuali interazioni tra esse.

le basi di dati relazionali

Un modo per rappresentare questo scenario è quello di fornire una rappresentazione grafica mediante un **diagramma Entità/Associazioni** (o Entità/Relazioni, nel seguito **E/R**) il cui formalismo è stato definito da Peter Chen nel 1976 (mostrato in figura).

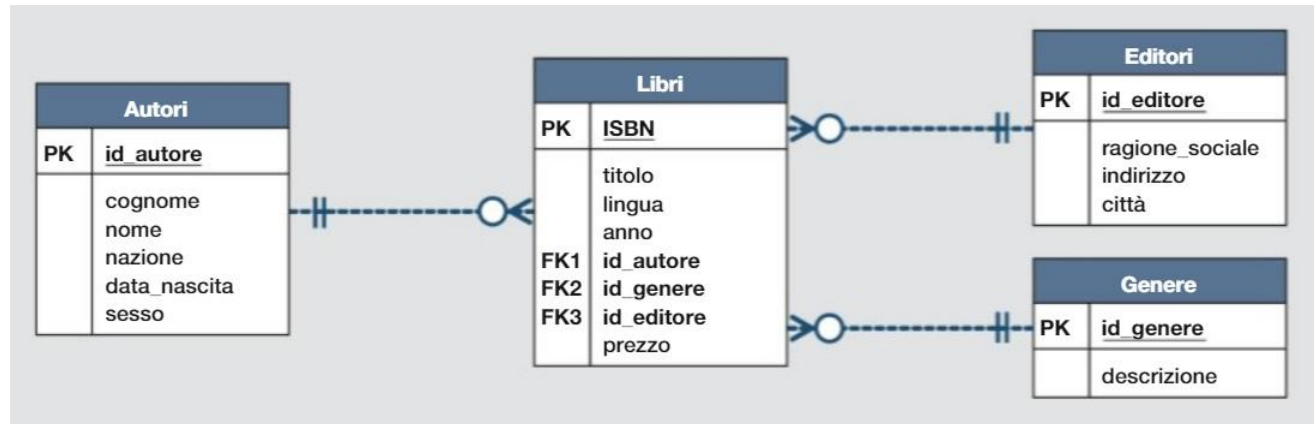
NB: In inglese *Entity/Relationship* - E/R, cioè Entità/Relazione.

Attenzione, però, è meglio tradurre la parola *relationship* in *associazione* per evitare confusioni con le “relazioni” del Modello Relazionale, che indicano tutt'altra cosa.



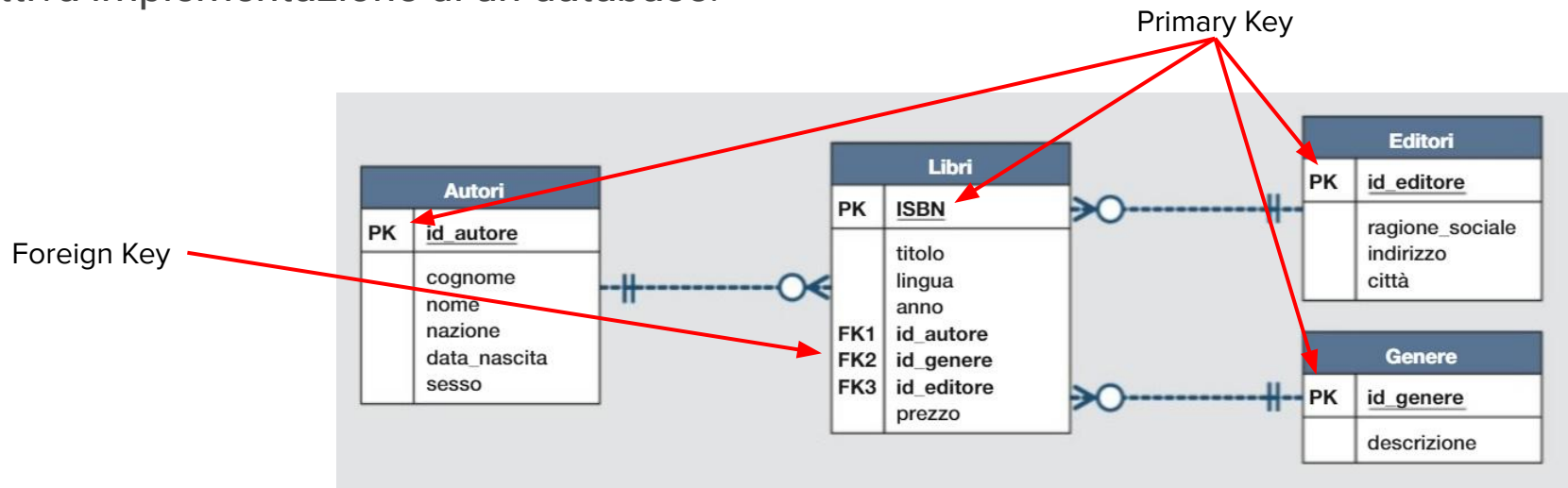
le basi di dati relazionali

Un altro tipo di formalismo grafico (diagramma delle tabelle derivato da UML) molto utilizzato dai progettisti di database e dagli sviluppatori software è invece quello illustrato in figura:



le basi di dati relazionali

In quest'ultimo caso, in aggiunta agli attributi indicati per le classi di dati, sono previsti elementi **identificativi** (contrassegnati dal simbolo «**PK**», acronimo di *Primary Key*) e **associativi** (contrassegnati dal simbolo «**FK**», acronimo di *Foreign Key*), necessari per l'effettiva implementazione di un database.



le basi di dati relazionali

In entrambi i diagrammi è immediato individuare gli **insiemi di dati** (*Libri, Autori, Editori, Genere*) e i relativi **attributi** che caratterizzano lo scenario della realtà da informatizzare oltre alle interrelazioni che esistono tra essi.



le basi di dati relazionali

In particolare il formalismo grafico E/R viene utilizzato per la realizzazione di modelli concettuali di scenari reali che **prescindono da come i dati saranno effettivamente memorizzati e organizzati nei computer di un sistema informatico.**

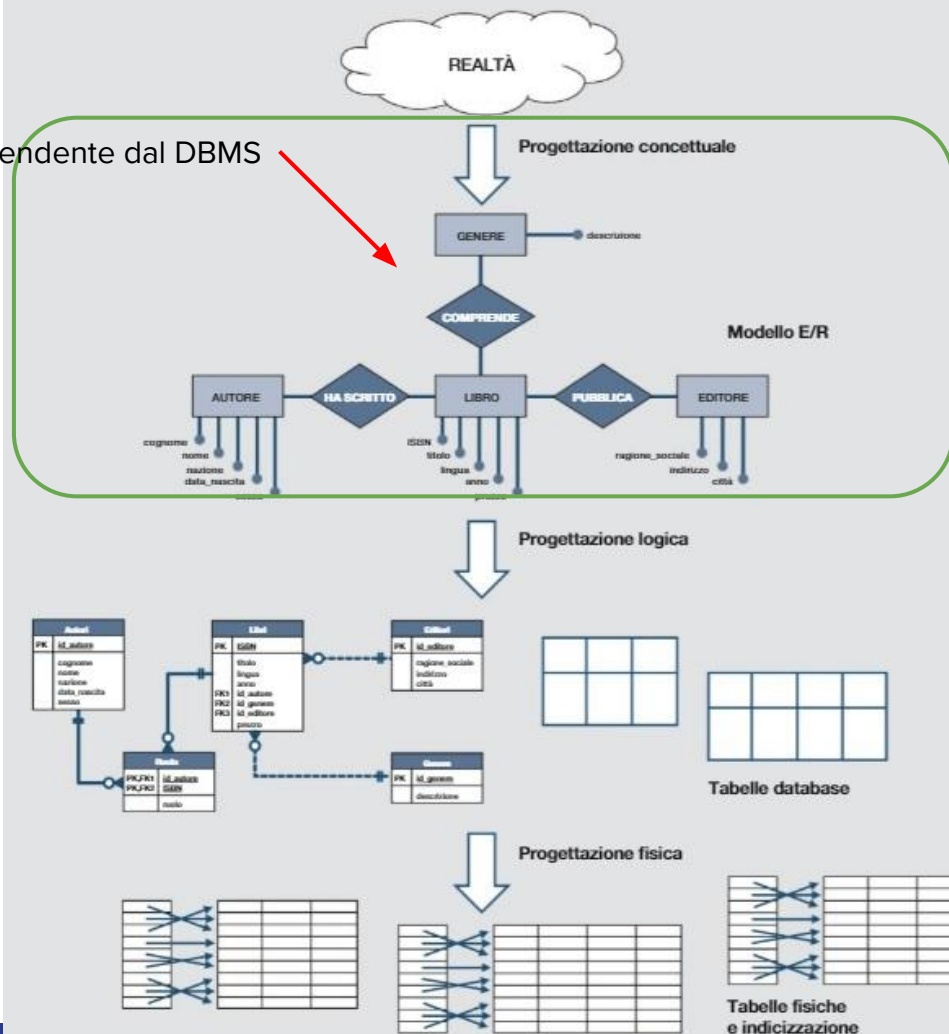
Il modello E/R è ancora oggi uno degli strumenti concettuali più sfruttati per la modellazione di dati in ambito informatico; viene impiegato non solo per la progettazione di database, ma anche come modello di riferimento per altri tipi di rappresentazione finalizzate alla modellazione: ad es. il linguaggio grafico UML ha ereditato i concetti introdotti con il modello E/R.

le basi di dati relazionali

Le **fasi del processo di progettazione di una base di dati** sono rappresentate nello schema di flusso della figura accanto.

A partire dai risultati dell'analisi viene realizzato un modello della realtà da informatizzare mediante uno schema E/R che descrive a livello concettuale le caratteristiche della base di dati da implementare.

Indipendente dal DBMS

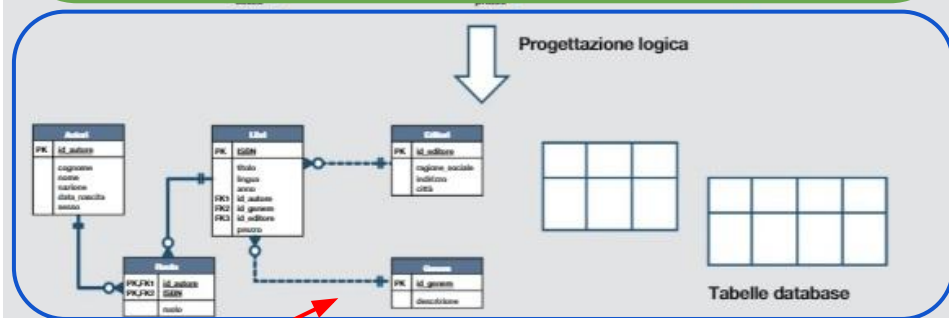
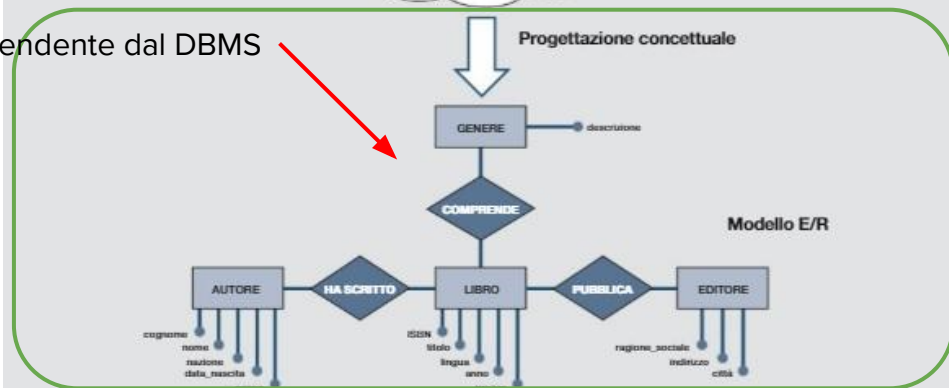


le basi di dati relazionali

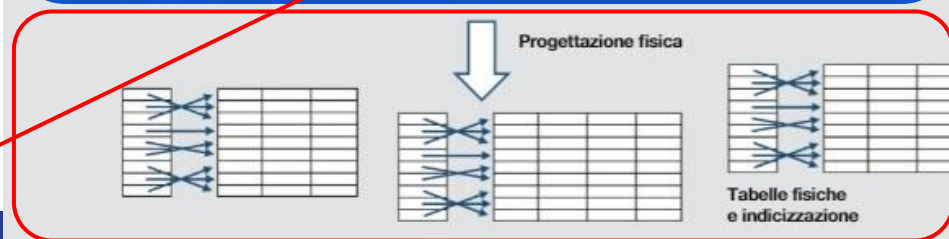
Tale modello viene trasformato in uno schema relazionale costituito da una collezione di tabelle.

Infine viene definita la struttura fisica dei dati (tipo e dimensioni dei campi) e vengono specificate le strutture ausiliarie (come gli indici di ricerca) per renderne efficiente la gestione dei dati.

Indipendente dal DBMS



Dipendente dal DBMS

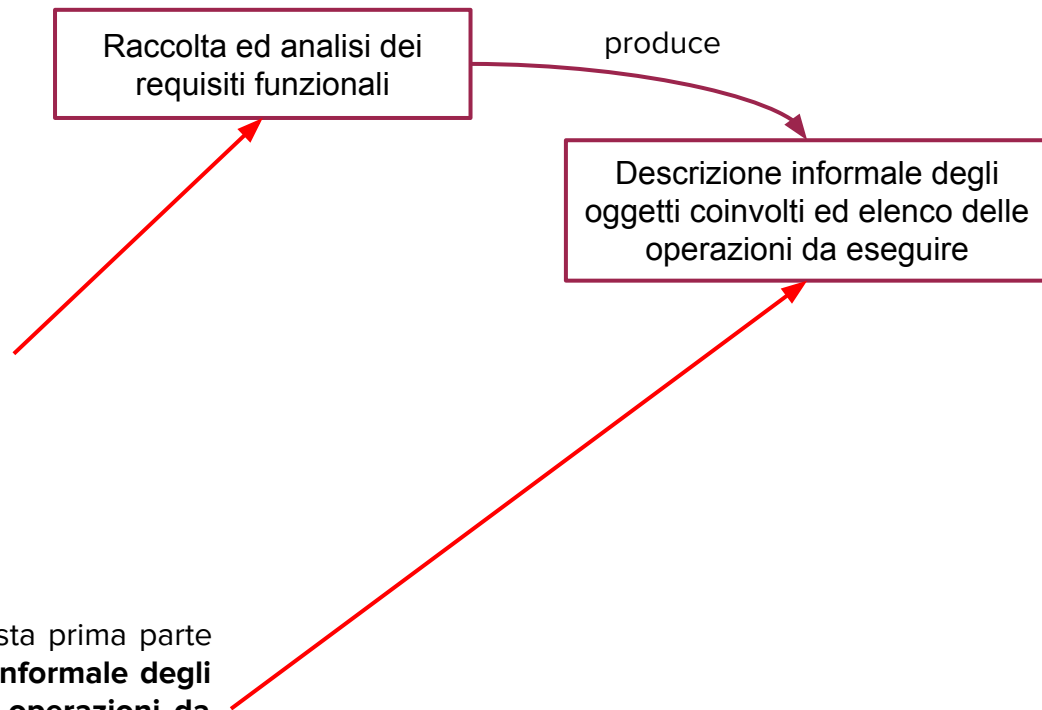


le basi di dati relazionali

Si parte con una fase preliminare di **raccolta ed analisi dei requisiti funzionali** che il sistema dovrà garantire.

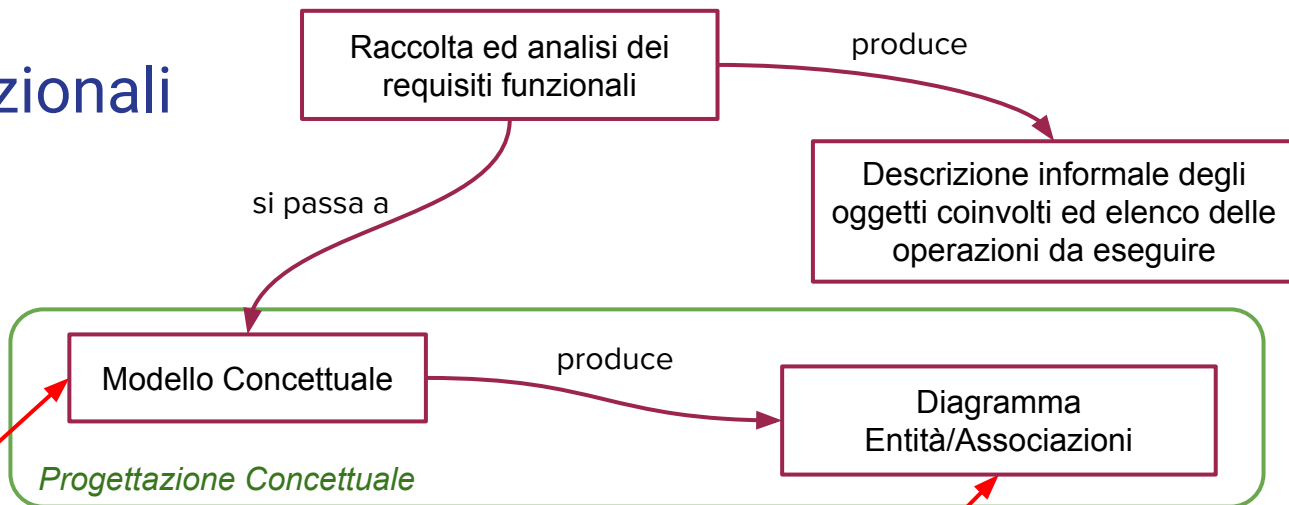
Ci domanderemo cioè quali saranno le informazioni che presumibilmente vorremo ottenere dal database, e quali i dati dovranno essere memorizzati per ottenere le informazioni desiderate.

La documentazione prodotta in questa prima parte comprenderà solo una **descrizione informale degli oggetti coinvolti** e un **elenco delle operazioni da eseguire** sia al livello di interrogazioni sul database, che al livello di manutenzione sui dati registrati.



le basi di dati relazionali

NB: Lo studio condotto fino a questo punto è di carattere generale e **non dipende dal particolare DBMS** che verrà utilizzato per implementare il database.



Il passo successivo dovrà produrre una rappresentazione formale delle idee e dei concetti espressi dalle specifiche informali appena descritte.

Ci serviremo in questa fase di un **Modello Concettuale** che ci permetta di ragionare sulla realtà di interesse, indipendentemente dagli aspetti realizzativi, e ci consenta di rappresentare i dati, scopo del nostro studio, come classi di oggetti (fatti, persone, cose, ecc.) con proprietà comuni ed esistenza autonoma.

La documentazione prodotta riguarda sostanzialmente una rappresentazione grafica del modello concettuale che individua gli oggetti di interesse, le loro proprietà caratteristiche (*attributi*) e le interazioni tra essi.

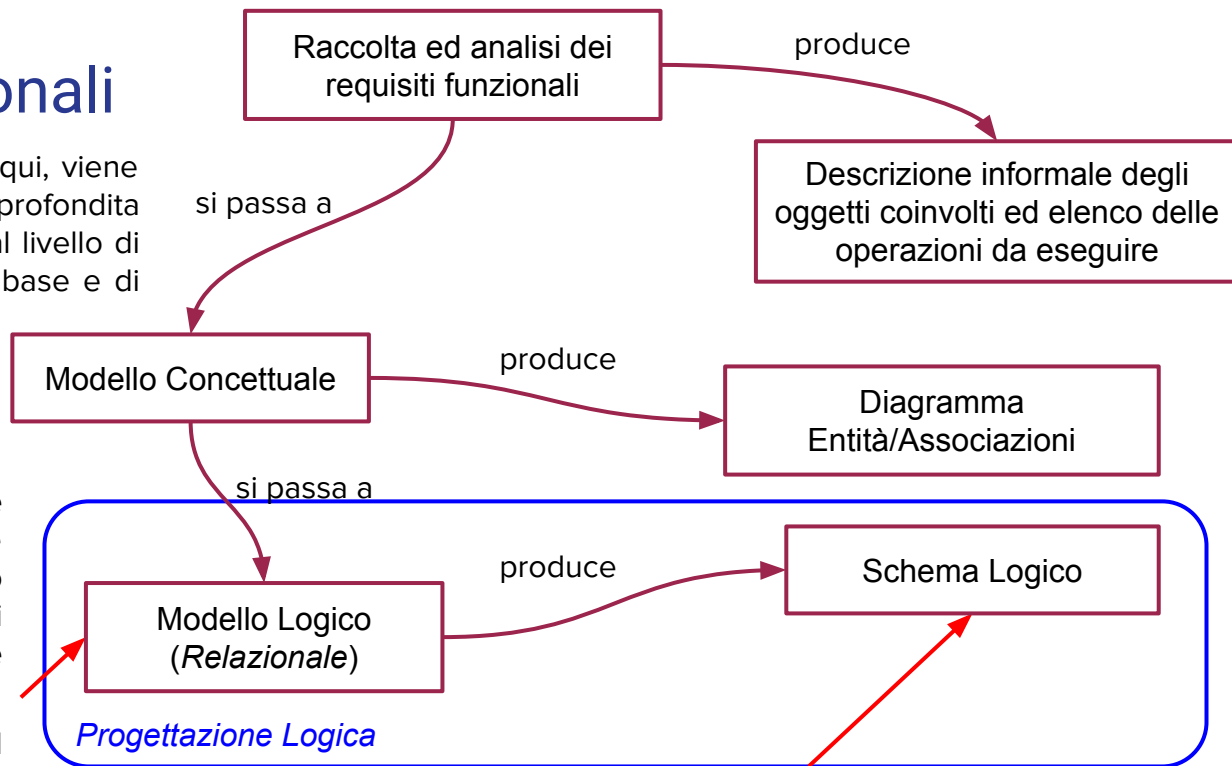
Viene indicata come **Diagramma Entità-Associazioni** (Entità/Relazione - E/R), e rappresenta il risultato della **Progettazione Concettuale**.

le basi di dati relazionali

NB: Contestualmente alle azioni viste fin qui, viene generalmente portata avanti un'analisi approfondita dei requisiti funzionali del sistema finale al livello di realizzazione delle interrogazioni sul database e di progettazione dell'interfaccia utente.

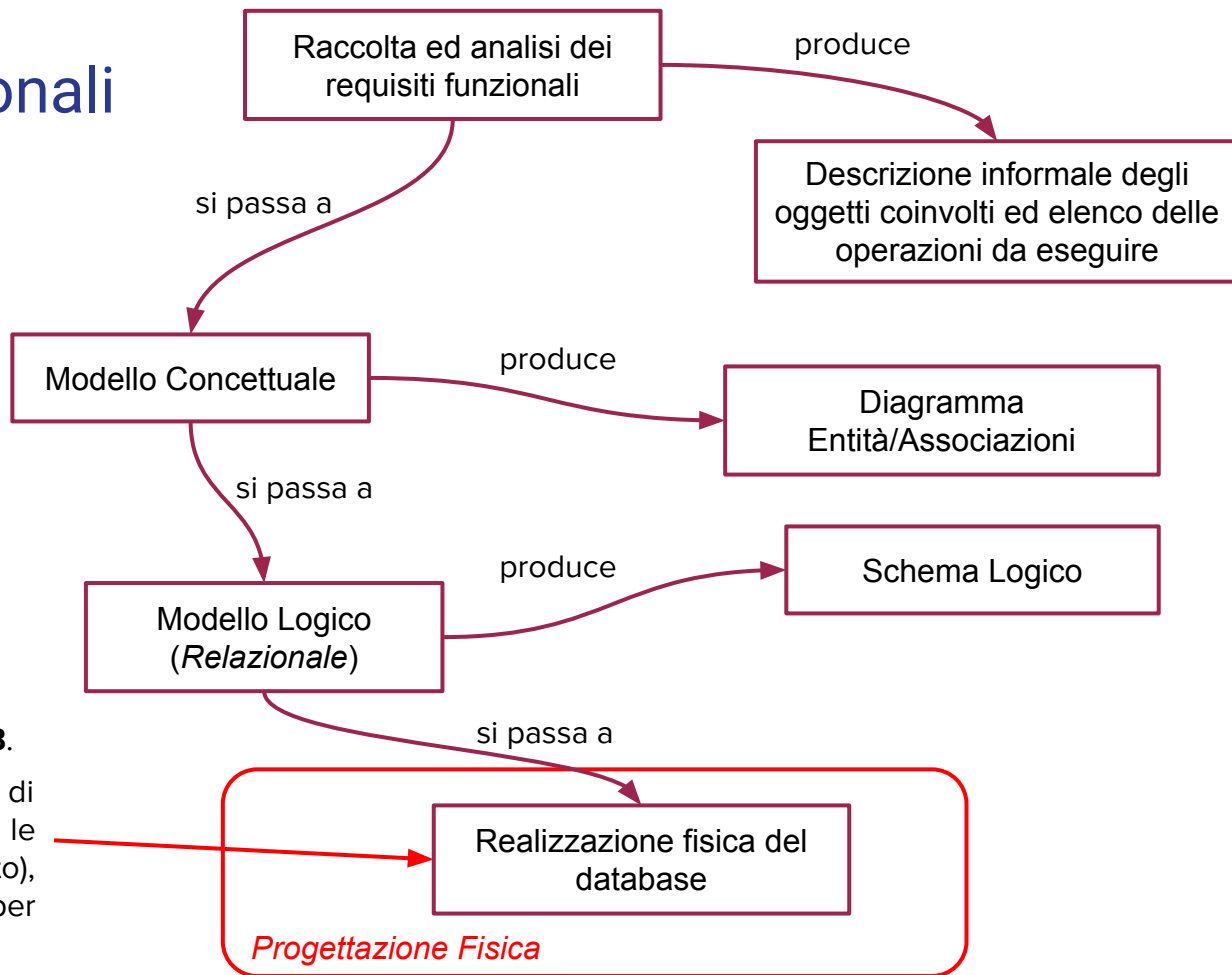
Il diagramma E/R prodotto deve essere fornito di un **Modello Logico** adeguato, che garantisca un formalismo matematico rigido per la definizione della struttura degli archivi e delle operazioni di manutenzione e di interrogazione sulla base di dati.

In questo contesto siamo interessati al **Modello Relazionale**, ed è a questo che faremo riferimento, visto che la maggior parte dei moderni DBMS gestiscono DB relazionali.



Il risultato di questa fase di **Progettazione Logica** è uno **Schema Logico** che fornisce una descrizione concreta del contenuto del database al fine di ottimizzarne gli aspetti funzionali sia al livello di interrogazioni che di aggiornamento dei dati.

le basi di dati relazionali



L'ultima fase è la **realizzazione fisica del DB**.

Vengono messe a punto le diverse tabelle di cui si compone il database (seguendo le indicazioni del Modello Relazionale prodotto), le interrogazioni previste, le maschere per l'interfaccia utente, ecc.

le basi di dati relazionali

Riassumendo, il risultato della progettazione di un DB deve dunque fornire la seguente documentazione:

- un'analisi preliminare dei requisiti funzionali del sistema informativo risultante, con l'elenco delle interrogazioni da realizzare;
- lo Schema E/R;
- lo Schema Logico;
- uno studio sulla possibile organizzazione dell'interfaccia utente.