

Progettazione di basi di dati

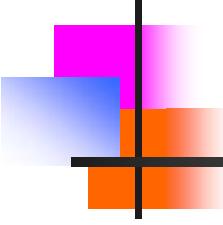
Sistemi Informativi L-B

Home Page del corso:

<http://www-db.deis.unibo.it/courses/SIL-B/>

Versione elettronica: progettazioneDB.pdf

Sistemi Informativi L-B

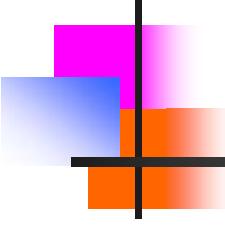


Progettazione di basi di dati

- È una delle attività del processo di sviluppo dei **sistemi informativi (SI)**
- Va quindi inquadrata in un contesto più generale:

Il ciclo di vita dei sistemi informativi:

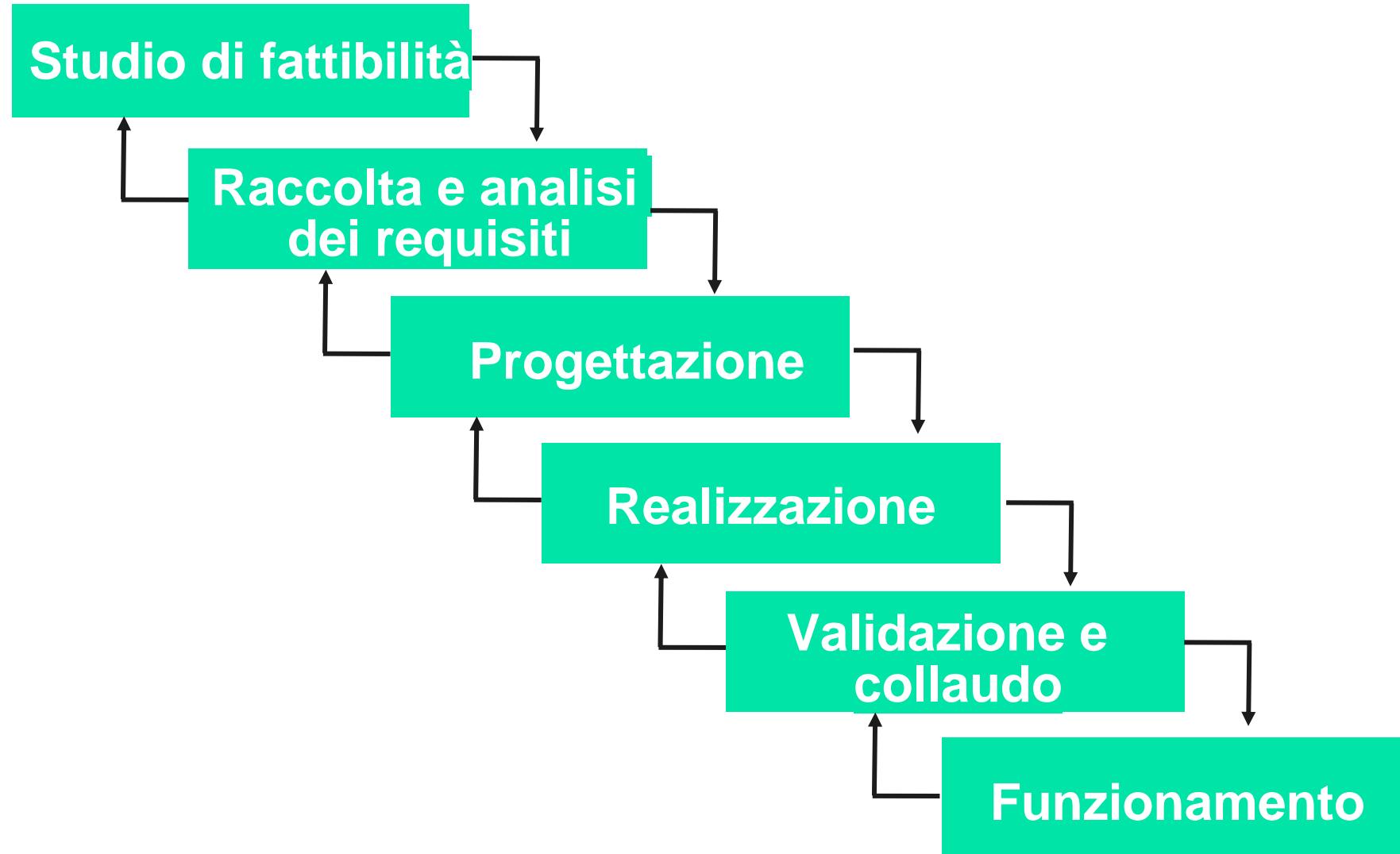
- Insieme e sequenzializzazione delle attività svolte da analisti, progettisti, utenti, nello sviluppo e nell'uso dei sistemi informativi
- attività iterativa, quindi **ciclo**

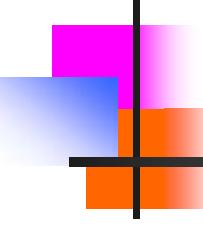


Fasi del ciclo di vita di un SI

- Studio di fattibilità (determinazione dei costi e delle priorità)
- Raccolta e analisi dei requisiti
- Progettazione della base di dati e delle applicazioni
- Realizzazione
- Validazione e verifica del corretto funzionamento
- Operatività (normale funzionamento)
 - ...e altre (prototipazione, installazione, ecc.)

Ciclo di vita di un sistema informativo

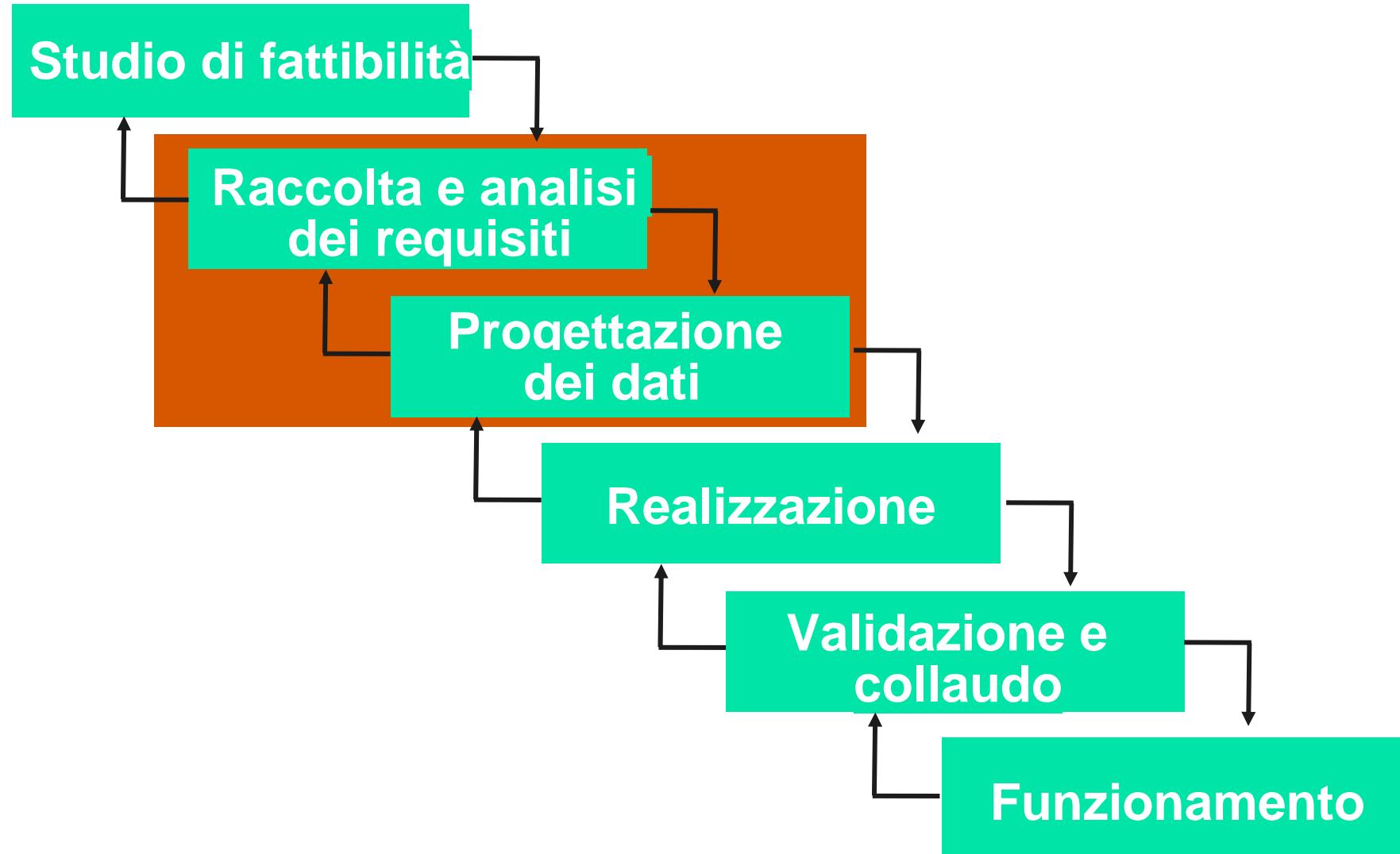


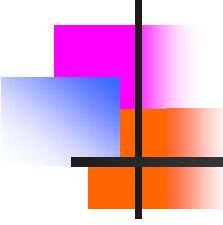


Progettazione guidata dai dati

- La progettazione di un sistema informativo riguarda 2 aspetti:
 - Progettazione dei dati
 - Progettazione delle applicazioni
- Il ruolo primario viene svolto dai dati, in quanto:
 - Sono (strutturalmente) più stabili
 - Sono condivisi da più applicazioni
- È quindi opportuno **progettare innanzitutto la base di dati e** successivamente le applicazioni

Ciclo di vita guidato dai dati





Metodologia di progettazione

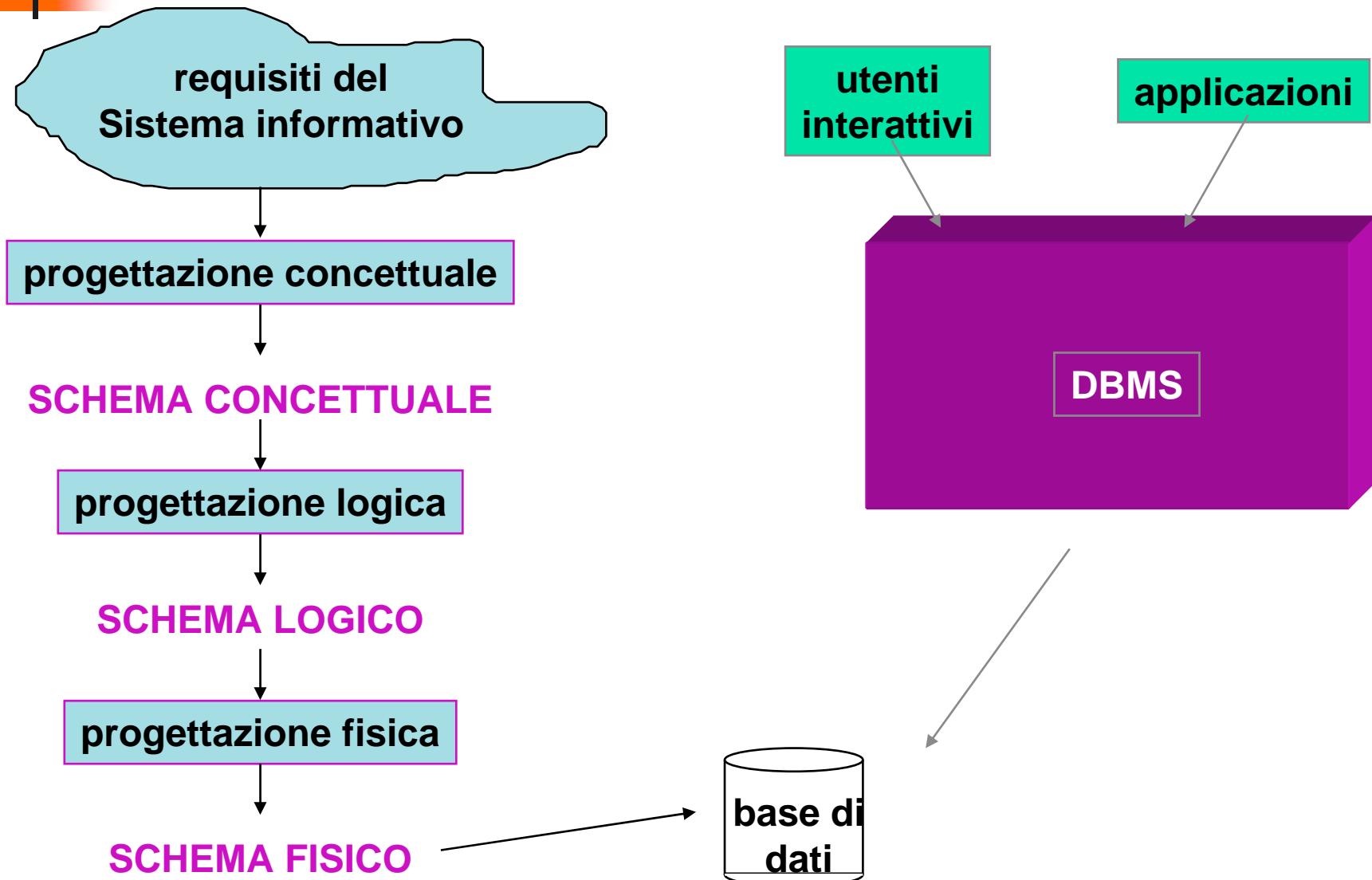
- Per progettare una base di dati (ma non solo) di buona qualità è opportuno seguire una

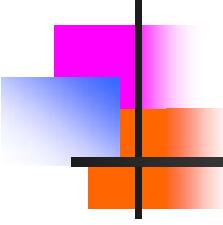
Metodologia di progettazione

che:

- definisca le fasi in cui l'attività di progettazione si articola
- fornisca dei criteri per scegliere tra diverse alternative
- sia supportata da dei modelli di rappresentazione
- sia di applicabilità generale e facile da utilizzare

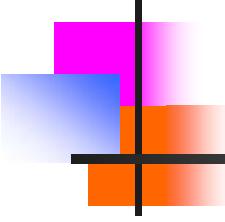
Progettazione della base di dati





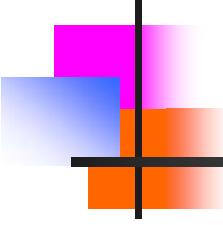
Fasi di progettazione

- La metodologia introdotta prevede 3 fasi:
 - progettazione concettuale
 - progettazione logica
 - progettazione fisica
- La fase di raccolta e analisi dei requisiti in pratica viene ad essere svolta congiuntamente a quella di progettazione concettuale
- Ognuna delle fasi si basa su un **modello**, che permette di generare una rappresentazione formale (**schema**) della base di dati ad un dato livello di astrazione (concettuale, logico e fisico):
 - Schema concettuale
 - Schema logico
 - Schema fisico



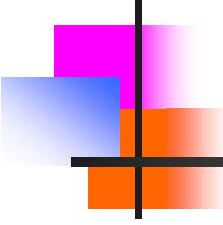
Fase di raccolta e analisi dei requisiti

- È la fase in cui si raccolgono e analizzano le specifiche informali ed eterogenee che i vari utenti danno delle procedure da automatizzare mediante un DBMS
 - **requisiti informativi**: caratteristiche dei dati
 - **requisiti sui processi**: operazioni sui dati
 - **requisiti sui vincoli di integrità**: proprietà dei dati e delle operazioni
- Attività principali:
 - Costruzione glossario dei termini
 - Eliminazione delle ambiguità (sinonimi, omonimi)
 - Raggruppamento dei requisiti “omogenei”
- Fase solo apparentemente semplice, nella realtà è spesso la più complessa perché è difficilmente standardizzabile il processo che porta a **capire cosa gli utenti vogliono!**



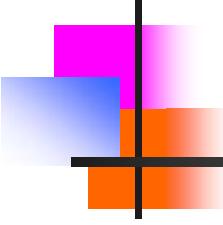
Fase di progettazione concettuale

- A partire dai requisiti informativi viene creato uno **schema concettuale**, cioè una **descrizione formalizzata e integrata** delle esigenze aziendali, espressa in modo **indipendente dal DBMS** adottato
- A tale scopo si adotta un **modello concettuale**, che permette di fornire descrizioni ad alto livello indipendenti dall'implementazione
- Lo schema concettuale è indipendente anche dal **tipo** di DBMS che sarà utilizzato (relazionale, gerarchico, ecc.)



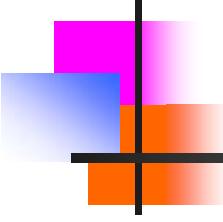
Fase di progettazione logica

- Consiste nella traduzione dello schema concettuale nel modello dei dati del DBMS
- Il risultato è uno **schema logico**, espresso nel DDL del DBMS
- In questa fase si considerano anche aspetti legati a:
 - integrità e consistenza (vincoli)
 - efficienza
- La progettazione logica si articola in due sotto-fasi:
 - ristrutturazione dello schema concettuale
 - traduzione verso il modello logico



Fase di progettazione fisica

- In questa ultima fase si operano scelte spesso strettamente dipendenti dallo specifico DBMS utilizzato
 - Ad esempio, lo stesso schema logico può essere fisicamente rappresentato in modo diverso in DB2 e in Oracle, al fine di meglio sfruttare le caratteristiche dei due DBMS
- Il risultato è lo **schema fisico**, che descrive le strutture di memorizzazione e accesso ai dati (tablespace, clustering, indici, ecc.)



Modelli dei dati: logici vs concettuali

Un modello dei dati è una collezione di concetti che vengono utilizzati per descrivere i dati, le loro associazioni, e i vincoli che questi devono rispettare

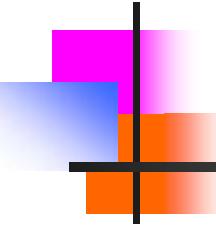
- Un ruolo di primaria importanza nella definizione di un modello dei dati è svolto dai **meccanismi che possono essere usati per strutturare i dati** (cfr. i costruttori di tipo in un linguaggio di programmazione)

Modelli logici: utilizzati nei DBMS esistenti per l'organizzazione dei dati

- utilizzati dai programmi, indipendenti dalle strutture fisiche

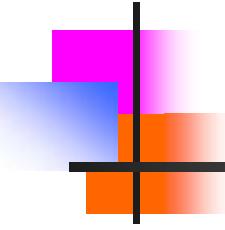
Modelli concettuali: permettono di rappresentare i dati in modo indipendente da ogni sistema

- cercano di descrivere i concetti del mondo reale
 - sono utilizzati nelle fasi preliminari di progettazione
-
- Il più noto è il modello **Entity-Relationship**



Alcuni modelli concettuali

- Nel tempo sono stati proposti numerosi modelli concettuali per la progettazione di basi di dati
 - modelli semantici, RM/T, ... [inizio anni '70]
 - Entity-Relationship (E/R) [entità-associazione] [Chen 1976]
 - IDEF1X [standard adottato dagli uffici governativi USA]
 - UML (Universal Modelling Language) [1999]



Riassumiamo:

- La progettazione di un sistema informativo è guidata dai dati, e si avvale di una metodologia che consta di diverse fasi
- Ogni fase produce uno schema, facendo uso di uno specifico modello
- Per la progettazione concettuale si fa uso di un modello concettuale che, astraendo da aspetti specifici dei DBMS e dalla rappresentazione concreta dei dati, costituisce un valido compromesso tra ciò che si dovrà realizzare e la realtà che si deve modellare