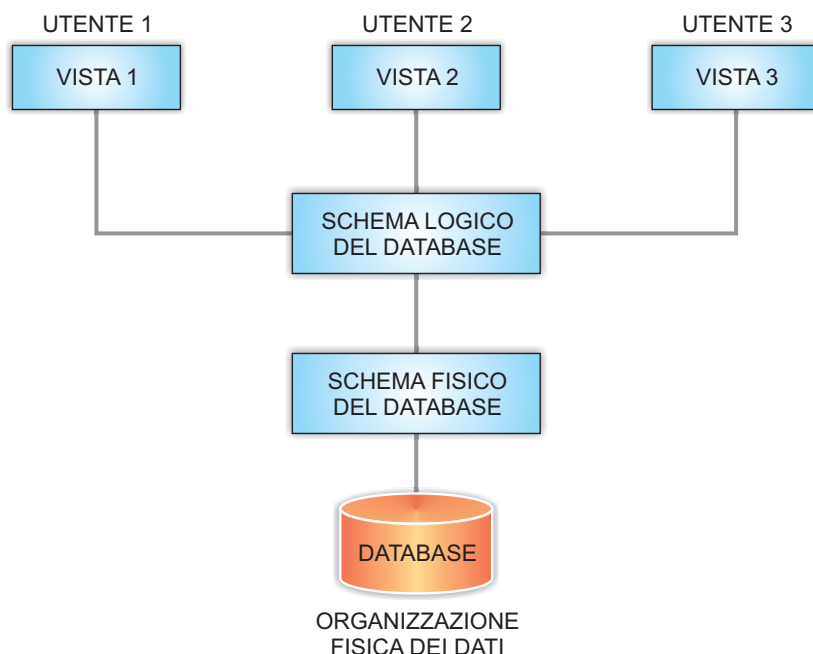




5. L'architettura a 3 livelli dei sistemi per database

L'architettura a tre livelli dei sistemi per database

I moderni DBMS seguono, di fatto, l'impostazione concettuale dell'*architettura a tre livelli ANSI-SPARC*, proposta nel 1975 da un comitato (*Standard Planning and Requirements Committee*) dell'ANSI (*American National Standards Institute*). Secondo questa impostazione i dati sono descritti secondo tre differenti livelli di astrazione, mediante opportuni schemi. I tre livelli sono indicati come il **livello esterno**, il **livello logico** e il **livello interno**.



Il *livello esterno* rappresenta la visione del database da parte dell'utente. Questo livello descrive quella parte del database che è importante per il singolo utente o per un gruppo di utenti. In generale in un database ci sono tanti schemi esterni, quante sono le classi di utente con differenti esigenze e visione dei dati.

Il *livello logico* rappresenta la visione complessiva del database dal punto di vista logico, indipendentemente dalle modalità e dalle tecniche di memorizzazione dei dati. La descrizione del database avviene mediante il modello di database impiegato. Naturalmente esiste un solo schema logico di database.

Il *livello interno* coincide con la rappresentazione fisica del database nel computer. La descrizione di questo livello fornisce informazioni in merito alla realizzazione concreta del livello logico.

Il livello logico è la cerniera tra ciò che vedono gli utenti e quello che viene effettivamente memorizzato nel computer. Il livello logico deve descrivere, con le convenzioni del modello di database scelto, entità, attributi, associazioni e vincoli.

Il livello interno entra nel merito dei dettagli implementativi delle forme di memorizzazione. Per esempio, una tabella di un database relazionale può essere rappresentata con un file ad accesso diretto e un struttura a indice sequenziale (oppure con una struttura degli indici ad albero) per realizzare la sua chiave primaria. Nel caso di un database gerarchico o reticolare è di competenza del livello interno il modo di realizzare, mediante puntatori, gli archi che collegano i nodi.

La trattazione svolta sull'organizzazione degli archivi è un esempio dei problemi di cui si occupa il livello interno. Il livello interno si interfaccia poi con il sistema operativo per la collocazione vera e propria dei blocchi di dati sui dischi, allo scopo di realizzare il livello fisico del database.

Per capire meglio il significato di livello esterno procediamo con un esempio. Consideriamo un database relazionale composto dalle tabelle *Studenti* e *Facoltà*.

Le due tabelle contengono informazioni di tipo anagrafico su alcuni studenti universitari e informazioni sulle facoltà che compongono l'università alla quale sono iscritti.

Studenti

Matricola	Nome	Cognome	Indirizzo	CodFac
2340	Nino	Verdi	Milano	Ing
2360	Enrico	Bianchi	Venezia	Eco
2361	Francesco	Spreafico	Bologna	Eco
2362	Anita	Rossigni	Ferrara	Let
2363	Clara	Riseri	Palermo	Let
2367	Gualtiero	Ceronte	Modena	Eco
2371	Giuseppe	Magnani	Padova	Lin
2373	Giuseppe	Galeotti	Milano	Ing
2374	Fabio	Lorenzoni	Cagliari	Ing
2589	Giusy	De Stefani	Napoli	Giu
2720	Franco	Rossi	Bergamo	Giu
3210	Silvia	Dini	Bari	Lin
3540	Marzia	Merli	Torino	Giu

Facoltà

CodFac	NomeFacoltà	Indirizzo
Eco	Economia	Via DeGasperi
Giu	Giurisprudenza	Piazza Martini
Ing	Ingegneria	Via Carducci
Let	Lettere Moderne	Via Monte Rosa
Lin	Lingue Straniere	Via Garibaldi

Gli uffici delle diverse facoltà sono interessati alle informazioni anagrafiche dei soli studenti della facoltà stessa. Di conseguenza la visione dei dati che hanno gli impiegati delle facoltà di Ingegneria ed Economia è differente. I primi desiderano conoscere i soli valori di *Matricola*, *Nome*, *Cognome* e *Indirizzo* dei record di *Studenti* per i quali il campo *CodFac* vale *Ing*, mentre i secondi desiderano le medesime informazioni, ma per record di *Studenti* che hanno il valore *Eco* nel campo *CodFac*.

Agli impiegati della facoltà di Economia e Ingegneria vengono di conseguenza fornite differenti visioni esterne della tabella *Studenti*, indicate in figura con i nomi di *StudentiEconomia* e *StudentiIngegneria*.

StudentiIngegneria

Matricola	Nome	Cognome	Indirizzo
2340	Nino	Verdi	Milano
2373	Giuseppe	Galeotti	Milano
2374	Fabio	Lorenzoni	Cagliari

StudentiEconomia

Matricola	Nome	Cognome	Indirizzo
2360	Enrico	Bianchi	Venezia
2361	Francesco	Spreafico	Bologna
2367	Gualtiero	Ceronte	Modena

Le tabelle in figura sono tabelle virtuali che prendono il nome di **viste logiche** e sono un esempio di quello che si intende per livello esterno dei database. Naturalmente esisteranno anche viste logiche di nome *StudentiGiurisprudenza*, *StudentiLettere* e *StudentiLingue*.

In relazione a questo semplice database possiamo dire che:

- il livello esterno è composto dalle cinque differenti visioni dei dati che hanno gli impiegati delle cinque facoltà;
- il livello logico è costituito dalla coppia di tabelle *Facoltà* e *Studenti*;
- il livello interno descrive il modo con il quale sono realizzate le tabelle *Studenti* e *Facoltà* e dipende dal DBMS scelto;
- al di sotto del livello interno ci sono i bit, o meglio i blocchi di byte, su disco e sono di competenza del sistema operativo.

Il DBMS mantiene al proprio interno la descrizione delle viste logiche, ovvero degli schemi esterni, dello schema logico e dello schema interno del database. In base a queste informazioni è in grado di attuare la corrispondenza (*mapping*) tra le viste logiche e lo schema logico del database e tra lo schema logico e il livello interno.

Supponiamo per esempio che un impiegato di Ingegneria desideri accedere tramite un'apposita interrogazione alle informazioni anagrafiche di uno studente di nome *Nino Verdi*:

- il DBMS traduce la richiesta nella ricerca dei record di *Studenti* con quel nome e cognome, ma con il campo *CodFac* di valore *Ing*;
- in base alle informazioni sull'organizzazione degli archivi contenute nello schema interno del database, il DBMS riesce a identificare le righe che contengono le informazioni cercate. Il risultato di questa operazione è del tipo: la riga cercata sta all'interno del file nel record di posto 3525;

- l'abbinamento tra numero di record e numero di blocco del file e la successiva ricerca su disco, vengono attuate dal DBMS e dal sistema operativo con modalità che dipendono dal DBMS usato e dallo specifico sistema operativo.

L'architettura a tre livelli dei database realizza meccanismi di astrazione dei dati e assicura l'*indipendenza dei dati*. Con questo termine si vuole indicare che i livelli superiori non sono influenzati, entro certi limiti, dai cambiamenti che avvengono nei livelli inferiori dell'architettura dei dati. Si identificano due livelli di indipendenza dei dati: l'**indipendenza logica** e l'**indipendenza fisica**.

Con *indipendenza logica dei dati* si fa riferimento alla capacità dello schema esterno di non essere influenzato dai cambiamenti apportati allo schema logico. Per esempio, l'aggiunta di campi alla tabella *Studenti* non ha alcun effetto sulla vista logica *StudentiEconomia* (e sulle altre quattro); le applicazioni basate su queste tabelle virtuali non ne risentiranno e non sarà necessario modificarle.

Con *indipendenza fisica dei dati* si fa riferimento alla capacità dello schema logico di non essere influenzato da cambiamenti apportati allo schema interno dei dati. Lo spostamento di una tabella da un dispositivo di memorizzazione a un altro, o la differente organizzazione degli archivi, non influenza lo schema logico e di conseguenza gli schemi esterni, rendendo il cambiamento trasparente per gli utenti.

L'architettura a tre livelli presenta l'ulteriore vantaggio di semplificare la visione del database da parte degli utenti in quanto essi vedono solo le informazioni alle quali sono interessati, secondo la loro visione del mondo reale, limitando le possibilità di errore e permettendo di affrontare i problemi di riservatezza in modo puntuale.