

# Prefazione

---

Questo testo presenta i concetti fondamentali sulle basi di dati, sui linguaggi di interrogazione e di gestione, sulle tecniche e sui metodi di progettazione, e sulle tecnologie e architetture per la gestione dei dati. Esso nasce da una lunga esperienza di insegnamento in corsi riguardanti le basi di dati, in ambito sia universitario sia industriale e applicativo, e pertanto si rivolge al pubblico degli studenti (in particolare di Ingegneria e di Scienze dell'informazione o Informatica) e a quello dei professionisti (utenti e progettisti di applicazioni).

## Contenuti

Il libro si articola in quattro parti:

*Basi di dati relazionali, modello e linguaggi.* Vengono presentate le caratteristiche fondamentali delle basi di dati che risultano di interesse per gli utenti e i programmatori. In particolare, si illustrano il modello relazionale e i relativi linguaggi, in modo preciso e concreto, con riferimento sia alle definizioni formali (del modello, dell'algebra e del calcolo) sia ai sistemi esistenti (con riferimento soprattutto al linguaggio SQL).

*Progettazione di basi di dati.* Viene illustrato ed esemplificato il processo di progettazione concettuale, logica e fisica delle basi di dati relazionali, che permette, partendo dai requisiti di utente, di arrivare a produrre strutture di basi di dati di buona qualità. Vengono poi discusse le tecniche principali per l'utilizzo delle basi di dati nelle applicazioni.

*Tecnologia delle basi di dati.* Vengono descritte le caratteristiche interne dei sistemi di basi di dati in rapporto all'architettura hardware e software del sistema informativo, in modo da comprenderne il funzionamento e sfruttarne appieno le potenzialità.

*Evoluzione dei linguaggi, dei modelli e dei sistemi per basi di dati.* Vengono illustrate le moderne varianti rispetto al modello e al linguaggio relazionale, focalizzandosi sulla gestione di dati XML (e sul linguaggio XQuery) e sulle basi di dati semantiche (e sul linguaggio Sparql). Vengono poi illustrate le principali architetture dei sistemi informativi moderni, focalizzandosi sulla distribuzione e sul parallelismo dei dati, sulla integrazione con il World Wide Web, sui sistemi per l'analisi dei dati e il supporto alle decisioni, e infine sui sistemi per la gestione dei *big data*, cioè di dati caratterizzati da semplice struttura ed elevati volumi.

Ciascun capitolo è corredato di numerosi esempi ed esercizi, nonché di una nota bibliografica che indica le fonti per possibili approfondimenti, elencate poi globalmente alla fine del volume.

## Utilizzo didattico

Nell'esperienza degli Autori, gli argomenti trattati in questo volume vengono svolti in modo completo in un tipico corso di basi di dati, da 10-12 crediti (circa 60 ore di lezione e 40 di esercitazione). A essi è opportuno associare un'ampia attività pratica; in particolare, per quanto riguarda le prime due parti del testo si consiglia lo svolgimento di un progetto di un sistema informativo che includa una base di dati, mentre per quanto riguarda le successive parti sono possibili vari approfondimenti progettuali legati all'uso di un linguaggio evoluto o di una particolare tecnologia o architettura. Informazioni utili circa l'organizzazione di alcuni DBMS relazionali sono presenti nei siti in appoggio al presente volume, descritti più avanti.

## Esperienze e ringraziamenti

L'organizzazione di questo testo e i suoi contenuti riflettono l'esperienza didattica degli Autori, che hanno tenuto per molti anni il corso universitario di Basi di dati e hanno svolto in altri contesti corsi sugli stessi temi. In particolare, Paolo Atzeni ha svolto in passato il corso di Basi di dati presso la facoltà di Ingegneria dell'Università di Roma "La Sapienza" e prima ancora presso l'Università di Toronto, tiene corsi di Basi di dati I e Basi di dati II presso il Dipartimento di Ingegneria dell'Università Roma Tre. Stefano Ceri tiene i corsi di Basi di dati e Data management for the Web presso la Scuola di Ingegneria Industriale e dell'Informazione del Politecnico di Milano. Ha inoltre tenuto varie edizioni del corso Principles of Distributed Databases presso l'Università di Stanford. Piero Fraternali tiene i corsi di Basi di dati e Web e Advanced Web Technologies presso la Scuola di Ingegneria Industriale e dell'Informazione del Politecnico di Milano. Stefano Paraboschi tiene corsi di Basi di dati e Web e Tecnologie per basi di dati presso la Scuola di Ingegneria dell'Università di Bergamo e presso la Scuola di Ingegneria Industriale e dell'Informazione del Politecnico di Milano. Riccardo Torlone tiene corsi di Basi di dati e Big Data presso il Dipartimento di Ingegneria dell'Università Roma Tre.

Alla concezione e alla revisione di questo testo hanno contribuito, direttamente o indirettamente, anche attraverso discussioni sui contenuti didattici dei corsi o suggerimenti di vario tipo, numerosi colleghi, collaboratori e lettori. Citiamo, fra gli altri, Maristella Agosti, Giorgio Ausiello, Elena Baralis, Giovanni Barone, Carlo Batini, Giampio Bracchi, Daniele Braga, Francesca Bugiotti, Luca Cabibbo, Alessandro Campi, Sara Comai, Sabrina De Capitani, Paolo De Nictolis, Giuseppe Di Battista, Angelo Foglietta, Sara Foresti, Maurizio Lenzerini, Davide Martinenghi, Gianni Mecca, Paolo Merialdo, Barbara Pernici, Giuseppe Psaila, Silvio Salza, Pierangela Samarati, Fabio Schreiber, Giuseppe Sindoni, Elena Tabet e Letizia Tanca. A ciascuno di essi, nonché a coloro che abbiamo dimenticato, va il nostro più sincero ringraziamento.

## Nota di edizione

Questa nuova edizione esce a piú di vent'anni di distanza dalla pubblicazione del volume: "Basi di dati: concetti, linguaggi e architetture", che ha visto la sua seconda edizione e un'edizione internazionale nel 1999. Da allora, abbiamo suddiviso il materiale in due volumi, le cui prime edizioni sono uscite nel 2003 e 2004, le seconde edizioni nel 2006 e 2007, e una terza e quarta edizione del primo volume nel 2009 e 2013. Nel 2014 abbiamo poi deciso un "ritorno all'antico", cioè di riproporre un volume unico, sulla spinta di indicazioni di molti docenti che ritenevano necessario per lo svolgimento di un primo corso di basi di dati alcuni materiali presenti nel secondo volume. Oggi presentiamo, a distanza di quattro anni, una nuova edizione del volume unico.

Il costante lavoro di aggiornamento e integrazione di queste opere tiene conto del continuo progresso dei linguaggi e della tecnologia per la gestione dei dati (che, per esempio, nell'ultimo decennio ha visto l'esplosione nelle applicazioni su Internet e su Cloud e lo sviluppo dei Big Data), e di commenti sull'uso del testo e di suggerimenti sui nuovi requisiti didattici espressi da numerosi colleghi, docenti di corsi di Basi di dati, offerti dalle sedi accademiche sia di Ingegneria sia di Scienze dell'informazione o Informatica, raccolti da McGraw-Hill e analizzati con attenzione dagli autori.

Questo volume introduce, nella sua quarta parte, alcuni aspetti nuovi rispetto alla precedente edizione.

- Abbiamo introdotto un nuovo capitolo sui big data, inquadrando questa tematica come la naturale evoluzione della nostra disciplina. Dopo aver illustrato le motivazioni del recente successo dei big data e la loro relazione con la scienza dei dati (data science), abbiamo illustrato le problematiche legate alle nuove tecnologie, tra cui Hadoop, MapReduce, Spark e Hive; il capitolo si conclude con una breve descrizione dei sistemi NoSQL, ripresa dalla precedente edizione.
- Abbiamo fortemente rivisto il capitolo relativo all'uso dei dati su Web, seguendo una evoluzione tecnologica che vede prevalere l'importanza delle RIA (Rich Internet Applications), la maggior enfasi sui servizi REST, e lo sviluppo di ambienti per la gestione di piattaforme mobili.
- Visto il diminuito rilievo applicativo, il capitolo sulle basi di dati a oggetti, che era presente nella quarta edizione, è stato omissso e reso reperibile sul sito.

Inoltre, tutto il materiale è stato rivisto e aggiornato, e al termine di ciascun capitolo sono stati aggiunti numerosi esercizi.

## Materiale aggiuntivo e piattaforma Connect

Materiale didattico di supporto al libro è disponibile al sito:

[highed.mheducation.com/sites/8838694451](http://highed.mheducation.com/sites/8838694451)  
<http://www.dia.uniroma3.it/librobd>

In particolare, per gli studenti sono disponibili le soluzioni di tutti gli esercizi presenti nel testo, alcuni esercizi aggiuntivi, l'Appendice A dell'edizione del 1996 (relativa al modello reticolare e omessa nelle edizioni successive) e le Appendici A, B e C della terza edizione, relative ai sistemi *Microsoft Access*, *DB2 Universal Database* e *Postgres*. È disponibile a partire dal 2018 il capitolo sulle basi di dati a oggetti della quarta edizione.

Per i docenti che utilizzano il testo sono disponibili, oltre ai materiali presenti nell'area studenti, i lucidi che coprono in modo completo i diciannove capitoli della presente edizione.

Per eventuali segnalazioni di errori e altri suggerimenti, gli Autori possono essere contattati attraverso i siti stessi.

Per questa edizione è stata sviluppata all'interno della piattaforma Connect una question bank digitale contenente circa 500 esercizi, suddivisi secondo i capitoli del libro cartaceo. Tale strumento permette al docente di testare e monitorare il livello di preparazione degli studenti e a questi ultimi di essere maggiormente coinvolti nella didattica, autovalutando il proprio apprendimento già durante il corso. Ogni esercizio è infatti dotato di correzione automatica e del riferimento puntuale al paragrafo di testo del quale è necessario aver acquisito il contenuto per poterlo risolvere.