

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 cinque 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 e delle applicazioni. 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 modelli e dei linguaggi per basi di dati. Vengono illustrate le moderne varianti rispetto al modello e al linguaggio relazionale, focalizzandosi sulle basi di dati a oggetti, multimediali, NoSql, sulla gestione di dati XML (e sul linguaggio XQuery), e sulle basi di dati semantiche (e sul linguaggio Sparql).

Architetture evolute per basi di dati. Vengono illustrate le principali architetture dei sistemi informativi moderni, focalizzandosi sulla distribuzione dei dati, sul parallelismo dei dati in ambiente Cloud, sulla integrazione con il World Wide Web e sui sistemi per l'analisi dei dati e il supporto alle decisioni.

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 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.

Il presente testo include al suo interno, nelle prime tre parti, i contenuti del volume *Basi di dati. Modelli e linguaggi di interrogazione* [91], anch'esso alla sua quarta edizione. Nelle precedenti edizioni, tecnologie, architetture e linee di evoluzioni erano presentate in un volume separato; abbiamo optato per un volume unico rispondendo alle indicazioni in questo senso da parte di vari docenti e lettori.

Per i docenti che nel proprio corso di Basi di dati non trattano tutti gli argomenti inclusi nel testo, l'Editore offre la possibilità di realizzare un libro "personalizzato" (<http://www.ateneonline.it/docenti/create.asp>).

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, Basi di dati II e Sistemi informativi 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 il Dipartimento 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 Bantini, 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 quasi 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. Abbiamo 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. Termina così una operazione di ristrutturazione dei contenuti che ci ha tenuti occupati per tutto il 2013 e che ci ha portato anche a spostare alcune appendici nei materiali disponibili in rete. 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 un'esplosione nelle applicazioni su Internet e su Cloud), 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 e quinta parte, molti aspetti nuovi rispetto alla precedente edizione del volume *Architetture e linee di evoluzione* (del 2007).

- Nel capitolo sulle basi di dati ad oggetti e multimediali abbiamo introdotto una breve descrizione dei sistemi NoSQL, che si stanno diffondendo soprattutto per la gestione di grandi quantità di dati su Web (big data).
- Abbiamo introdotto un nuovo capitolo sulle basi di dati semantiche, approfondendo in particolare la descrizione del linguaggio SPARQL e lo sviluppo dei dati "aperti" (open data).
- Abbiamo separato l'argomento del parallelismo da quello della distribuzione dei dati, dedicando al parallelismo un capitolo autonomo che descrive le nuove architetture per cloud computing e il linguaggio Pig Latin per la gestione dei dati su cloud.
- Abbiamo fortemente rivisto il capitolo relativo all'uso dei dati su Web, introducendo una nuova sezione relativa alle Rich Internet Application che include la descrizione di nuovi standard - tra cui AJAX, JSON e HTML 5.

- Nel capitolo relativo alla analisi dei dati è stato dato ampio spazio alla progettazione dei data warehouse, introducendo una metodologia completa e un esempio di progettazione.

Le prime tre parti del volume sono identiche al volume *Basi di dati. Modelli e linguaggi di interrogazione*, anch'esso giunto alla sua quarta edizione (del 2013). Inoltre, tutto il materiale è stato rivisto e aggiornato, e al termine di ciascun capitolo sono stati aggiunti numerosi esercizi.

Materiale aggiuntivo

Materiale didattico di supporto al libro è disponibile ai siti:

<http://www.ateneonline.it/atzeni>
<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*.

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.