

Questo testo presenta le architetture e le linee di tendenza delle basi di dati. Esso nasce da una lunga esperienza di insegnamento in corsi riguardanti le basi di dati, sia in ambito universitario sia in ambito industriale e applicativo, e pertanto si rivolge sia al pubblico degli studenti (in particolare di Ingegneria e di Scienze dell'informazione o Informatica) sia a quello dei professionisti (utenti e progettisti di applicazioni).

## Contenuti

Questo volume è la continuazione del volume “Basi di dati: modelli e linguaggi di interrogazione”, spesso riferito nel testo come il “primo volume.” Il libro si articola in tre parti:

1. *Tecnologia delle basi di dati.* Vengono descritte le tecnologie per garantire la gestione efficiente, affidabile e concorrente dei DBMS, visti come sistemi autonomi. Il *primo capitolo* descrive la struttura fisica dei dati e l'esecuzione e ottimizzazione delle interrogazioni; il *secondo capitolo* descrive il concetto di transazione e la sua realizzazione tramite controllo di affidabilità e di concorrenza.
2. *Evoluzione dei modelli e dei linguaggi per basi di dati.* Vengono illustrate le principali linee di tendenza che caratterizzano l'attuale evoluzione delle basi di dati. Il *terzo capitolo* illustra le basi di dati ad oggetti (object-oriented) e i sistemi misti relazionali e a oggetti (object-relational); il *quarto capitolo* illustra la gestione e interrogazione di dati persistenti in XML; il *quinto capitolo* descrive le basi di dati attive e la loro evoluzione verso i sistemi di gestione di stream di dati. In questi capitoli vengono presentate le estensioni di SQL-3 relative a oggetti e a regole attive e viene introdotto il linguaggio XQuery.
3. *Architetture evolute per basi di dati.* Vengono descritte le architetture dei sistemi informativi moderni. Il *sesto capitolo* descrive il modo in cui la tecnologia delle basi di dati si integra con quella delle reti di trasmissione formando sistemi distribuiti e illustrando le architetture client/server, distribuite, parallele e replicate. Il *settimo capitolo* descrive le interazioni tra la tecnologia di gestione dei dati e quella del World Wide Web, illustrando le varie modalità di generazione di pagine i cui contenuti vengono estratti dinamicamente dalle basi di dati. Infine, l'*ottavo capitolo* illustra le architetture per l'analisi dei dati e per il supporto alle decisioni.

Ciascun capitolo è corredato di esempi ed esercizi, nonché di una nota bibliografica che indica le fonti di 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 abbastanza completo, in un tipico secondo corso di basi di dati, corrispondente cioè a un modulo da cinque o sei crediti (circa trenta ore di lezione e venti di esercitazione). Un corso di approfondimento degli aspetti tecnologici delle basi di dati è offerto, sotto varie denominazioni, in quasi tutti i curricula di Ingegneria informatica e di Scienze dell'informazione o Informatica, al primo o al secondo livello.

Al corso è opportuno associare un'attività pratica di approfondimento che può affrontare un qualunque argomento trattato nel presente volume, coordinandosi con l'attività di laboratorio prevista in un primo corso di basi di dati. Ad esempio, prendendo come punto di partenza una base di dati relazionale realizzata in tale ambito, è possibile realizzare una sua versione in ambito client/server o distribuito, oppure pubblicarla su Web, oppure utilizzare per alcuni contenuti servizi Web presenti in Internet, oppure farne una versione attiva aggiungendo a essa opportuni trigger. Altri progetti specifici possono riguardare le basi di dati a oggetti e i data warehouse.

La coppia costituita dal presente volume e dal testo "Basi di dati: Modelli e linguaggi di interrogazione" copre in modo completo un corso universitario di circa dieci crediti, corrispondente a sessanta ore di lezione e a quaranta ore di esercitazione.

## 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 il corso di Basi di dati presso la Facoltà di Ingegneria dell'Università di Roma "La Sapienza" (e in precedenza presso l'Università di Toronto) e ora tiene i corsi di Basi di dati e Tecnologia delle basi di dati presso la Facoltà di Ingegneria dell'Università Roma Tre. Stefano Ceri tiene i corsi di Basi di dati 1 e 2 presso la Facoltà di Ingegneria del Politecnico, nelle sedi di Milano e di Como (Laurea OnLine), e presso la Facoltà di Informatica e di Scienze della Comunicazione dell'Università della Svizzera Italiana a Lugano. 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 di Sistemi informativi presso la sede di Como della Facoltà di Ingegneria del Politecnico. Stefano Paraboschi ha tenuto i corsi di Basi di dati presso le sedi di Milano e Como della Facoltà di Ingegneria del Politecnico di Milano e ora tiene i corsi di Sistemi informativi presso la Facoltà

di Ingegneria dell'Università di Bergamo. Riccardo Torlone tiene i corsi di Basi di dati e Complementi di basi di dati presso la Facoltà di Ingegneria dell'Università Roma Tre.

Alla concezione di questo testo hanno contribuito, direttamente o indirettamente, anche attraverso discussioni sui contenuti didattici dei corsi, numerosi colleghi e collaboratori, fra cui Maristella Agosti, Giorgio Ausiello, Elena Baralis, Giovanni Barone, Carlo Batini, Aldo Bongio, Giampio Bracchi, Daniele Braga, Stefano Butti, Luca Cabibbo, Alessandro Campi, Sara Comai, Giuseppe Di Battista, Angelo Foglietta, Maurizio Lenzerini, Maristella Matera, Gianni Mecca, Paolo Merialdo, Paolo Missier, Barbara Pernici, Silvio Salza, 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. Siamo ancor più debitori nei confronti dei nostri studenti, che hanno sopportato per anni corsi svolti con dispense parziali e incoerenti, contribuendo spesso al loro miglioramento.

## Note di edizione

La presente edizione è una versione riveduta e ampliata della prima edizione, apparsa nel 2003, che a sua volta aveva esteso i contenuti presenti nelle parti terza e quarta del volume “Basi di dati: linguaggi, sistemi e architetture”, pubblicato nel 1999 contemporaneamente al volume “Database Systems: Concepts, Languages and Architectures” (McGraw-Hill, 1999). È fondamentale infatti rinnovare frequentemente le edizioni del volume per tenere i suoi contenuti allineati con i progressi tecnologici.

Rispetto alla edizione del 2003, le modifiche principali sono così riassunte:

- L'organizzazione in parti del libro è stata ripensata, antepoendo la trattazione delle parti strettamente tecnologiche relative ai database server agli aspetti di evoluzione dei modelli e dei linguaggi, e trattando per ultime le architetture, cioè il modo di comporre database server nei vari contesti applicativi.
- La prima parte, che è più stabile, ha subito essenzialmente una messa a punto.
- Nella seconda parte del libro, tenuto conto dell'evoluzione del settore, abbiamo un po' ridotto l'importanza della trattazione dei sistemi puramente “object-oriented” e invece dato più spazio ai sistemi “object-relational” e allo standard SQL-3. Abbiamo esteso il capitolo dedicato alla trattazione della gestione dei dati in XML e specificamente del linguaggio XQuery, che sta diventando sempre più importante e diffuso. Nel capitolo relativo alle basi di dati attive, abbiamo dato più spazio allo standard SQL-3 e abbiamo dedicato una breve trattazione ai sistemi per la gestione di flussi di dati (stream).
- La terza parte del libro, dedicata alle architetture, ha subito modifiche più o meno significative a seconda del differente grado di maturità dei contesti che abbiamo preso in considerazione. Il capitolo sui sistemi distribuiti, un contesto più stabile, ha subito piccoli aggiustamenti; ad esempio, abbiamo descritto il protocollo di commit Paxos, di recente definizione. In modo analogo, abbiamo

apportato limitate variazioni al capitolo sulla gestione dei dati per il supporto delle decisioni; ad esempio, abbiamo ampliato la sezione dedicata al “data mining”, inquadrando l’argomento in modo più organico. Invece, il capitolo sull’uso delle basi di dati nell’ambito del Word Wide Web ha subito una notevole ristrutturazione, dovuta all’emergere di nuove tecnologie che si sono, nel breve volgere di quattro anni, sostituite alle tecnologie precedenti. Abbiamo quindi aggiunto una sezione relativa alla architettura MVC (*model-view-controller*), una relativa all’emergere di nuovi concetti legati alla programmazione a oggetti in Java, e una ampia trattazione legata ai servizi Web e al loro uso nel contesto della gestione dei dati.

## **Materiale aggiuntivo**

Materiale didattico di supporto a questo volume e al precedente è disponibile sul sito:

`http://www.ateneonline.it/atzeni`

In particolare, sono disponibili lucidi che coprono tutti gli argomenti, ulteriori esercizi, nonché le soluzioni di alcuni esercizi. Per segnalazioni di errori e altri suggerimenti, gli Autori possono essere contattati attraverso il sito stesso.