



C. Batini, M. Scannapieco

## **Data and Information Quality**

Dimensions, Principles and Techniques

Series: Data-Centric Systems and Applications

- ▶ **Presents an extensive description of the techniques that constitute the core of data and information quality research**
- ▶ **Combines concrete practical solutions, such as methodologies, benchmarks, and case studies with sound theoretical formalisms**
- ▶ **Includes also necessary foundations from probability theory, statistical data analysis, and machine learning**

This book provides a systematic and comparative description of the vast number of research issues related to the quality of data and information. It does so by delivering a sound, integrated and comprehensive overview of the state of the art and future development of data and information quality in databases and information systems.

To this end, it presents an extensive description of the techniques that constitute the core of data and information quality research, including record linkage (also called object identification), data integration, error localization and correction, and examines the related techniques in a comprehensive and original methodological framework. Quality dimension definitions and adopted models are also analyzed in detail, and differences between the proposed solutions are highlighted and discussed. Furthermore, while systematically describing data and information quality as an autonomous research area, paradigms and influences deriving from other areas, such as probability theory, statistical data analysis, data mining, knowledge representation, and machine learning are also included. Last not least, the book also highlights very practical solutions, such as methodologies, benchmarks for the most effective techniques, case studies, and examples.

The book has been written primarily for researchers in the fields of databases and information management or in natural sciences who are inte

rested in investigating properties of data and information that have an impact on the quality of experiments, processes and on real life. The material presented is also sufficiently self-contained for masters or PhD-level courses, and it covers all the fundamentals and topics without the need for other textbooks. Data and information system administrators and practitioners, who deal with systems exposed to data-quality issues and as a result need a systematization of the field and practical methods in the area, will also benefit from the combination of concrete practical approaches with sound theoretical formalisms.

1st ed. 2015, X, 546 p. 260 illus., 53 illus. in color.

### Printed book

#### **Hardcover**

- ▶ 59,99 € | £44.99 | \$79.99
- ▶ \*64,19 € (D) | 65,99 € (A) | CHF 66.00

### eBook

Available from your library or

- ▶ [springer.com/shop](http://springer.com/shop)

### MyCopy

#### **Printed eBook for just**

- ▶ € | \$ 24.99
- ▶ [springer.com/mycopy](http://springer.com/mycopy)



Order online at [springer.com](http://springer.com) ▶ or for the Americas call (toll free) 1-800-SPRINGER ▶ or email us at: [customerservice@springer.com](mailto:customerservice@springer.com). ▶ For outside the Americas call +49 (0) 6221-345-4301 ▶ or email us at: [customerservice@springer.com](mailto:customerservice@springer.com).

The first € price and the £ and \$ price are net prices, subject to local VAT. Prices indicated with \* include VAT for books; the €(D) includes 7% for Germany, the €(A) includes 10% for Austria. Prices indicated with \*\* include VAT for electronic products; 19% for Germany, 20% for Austria. All prices exclusive of carriage charges. Prices and other details are subject to change without notice. All errors and omissions excepted.