How do you approach answering queries when your data is stored in multiple databases that were designed independently by different people? This is first comprehensive book on data integration and is written by three of the most respected experts in the field. This book provides an extensive introduction to the theory and concepts underlying today's data integration techniques, with detailed instruction for their application using concrete examples throughout to explain the concepts. Data integration is the problem of answering queries that span multiple data sources (e.g., databases, web pages). Data integration problems surface in multiple contexts, including enterprise information integration, query processing on the Web, coordination between government agencies and collaboration between scientists. In some cases, data integration is the key bottleneck to making progress in a field. The authors provide a working knowledge of data integration concepts and techniques, giving you the tools you need to develop a complete and concise package of algorithms and applications. *Offers a range of data
Integration solutions enabling you to focus on what is most relevant to the problem at hand. Enables you to build your own algorithms and implement your own data integration applications. Companion website with numerous project-based exercises and solutions and slides. Links to commercially available software allowing readers to build their own algorithms and implement their own data integration applications. Facebook page for reader input during and after publication.

**Principles Data Integration** Anhai Doan

**Connecting the Data: Data Integration Techniques for Building an Operational Data Store (ODS)**

Business data integration is a complex problem that must be solved when organizations change or enhance their internal structures. The goal of this book is to present a simple yet thorough resource that describes the challenges of business data integration and the solutions to these challenges such as schema integration, illustrated through an Operational Data Store (ODS) case study. This book contains three sections spanning ten chapters. Section I, Foundational Concepts, will provide you wit...

**SAS Data Integration Studio 4.7**

Describes the main tasks that you can perform in SAS Data Integration Studio, including: data access; data integration; metadata management; data cleansing and enrichment; extract, transform, and load (ETL); extract, load, and transform (ELT); and service-oriented architecture (SOA) and message queue integration. Describes the main tasks that you can perform in SAS Data Integration Studio, including: data access; data integration; metadata management; data cleansing and enrichment; extract, tran...

**Data Integration in the Life Sciences**

For several years now, there has been an exponential growth of the amount of life science data (e.g., sequenced complete genomes, 3D structures, DNA chips, mass spectroscopy data), most of which are generated by high-throughput experiments. This exponential corpus of data is stored and made available through a large number of databases and resources over the Web, but unfortunately still with a high degree of semantic heterogeneity and varying levels of quality. These data must be combined together...

**The Integration of Psychological Principles in Policy Development**

The thrust of this book is to educate policymakers, academics, political scientists, and others as to the beneficial use of psychological principles within a multidisciplinary framework to aid in the development of more productive policy. Contributors to this volume offer an integrative, collaborative, synergistic approach to effecting positive change. The book's goal is threefold: to offer a varied collection of solutions and models of applications of psychological principles; to collect a dive...

**Data Integration Blueprint and Modeling: Techniques for a Scalable and Sustainable Architecture (IBM Press)**

Making Data Integration Work: How to Systematically Reduce Cost, Improve Quality, and Enhance Effectiveness Today's enterprises are investing massive resources in data integration. Many possess thousands of point-to-point data integration applications that are costly, undocumented, and difficult to maintain. Data integration now accounts for a major part of the expense and risk of typical data warehousing and business intelligence projects—and, as businesses increasingly rely on analytics...

**Interactive Graphics for Data Analysis: Principles and Examples (Chapman & Hall/CRC Computer Science & Data Analysis)**

Interactive Graphics for Data Analysis: Principles and Examples discusses exploratory data analysis (EDA) and how interactive graphical methods can help gain insights as well as generate new questions and hypotheses from datasets. Fundamentals of Interactive Statistical Graphics The first part of the book summarizes principles and methodology, demonstrating how the different graphical representations of variables of a dataset are effectively used in an interactive setting. The authors introduce t...
Lean Integration: An Integration Factory Approach to Business Agility

Lean Integration is an excellent resource for anyone struggling with the challenges of performing integration for a complex enterprise. Steve J. Dennis, Integration Competency Center Director, Nike Use Lean Techniques to Integrate Enterprise Systems Faster, with Far Less Cost and Risk By some estimates, 40 percent of IT budgets are devoted to integration. However, most organizations still attack integration on a project-by-project basis, causing unnecessary expense, waste, risk, and...

Enterprise Integration: The Essential Guide to Integration Solutions

The books use of real-world case study vignettes really does go to the heart of the subject matter. This stuff is real, it has real applicability to real problems, and, as with most things in life, it shows how it all comes down to real money in the final analysis. This book shows you what your peers are doing to drive costs out of integration projects and to build new applications without re-inventing the entire wheel just a few new spokes and off you go. This is a good book. Read it. Peter Rhys ...

Pentaho Kettle Solutions: Building Open Source ETL Solutions with Pentaho Data Integration

A complete guide to Pentaho Kettle, the Pentaho Data Integration toolset for ETL This practical book is a complete guide to installing, configuring, and managing Pentaho Kettle. If you're a database administrator or developer, you'll first get up to speed on Kettle basics and how to apply Kettle to create ETL solutions before progressing to specialized concepts such as clustering, extensibility, and data vault models. Learn how to design and build every phase of an ETL solution. Shows developers and ...

Data Structures, Algorithms, and Software Principles in C

Using C, this book develops the concepts and theory of data structures and algorithm analysis in a gradual, step-by-step manner, proceeding from concrete examples to abstract principles. Standish covers a wide range of both traditional and contemporary software engineering topics. The text also includes an introduction to object-oriented programming using C++. By introducing recurring themes such as levels of abstraction, recursion, efficiency, representation and trade-offs, the author unifies t...

Related Topics

Biological Data Integration

Pentaho Data Integration Tutorial

Pentaho Data Integration 4 Cookbook

Pentaho Data Integration Server

Pentaho Data Integration Wiki

Enterprise Architecture Integration Principles

Principles Of Gene Microarray Data Analysis

Pentaho Kettle Solutions Building Open Source Etl Solutions With Pentaho Data Integration

Machine Learning And Data Mining Introduction To Principles And Algorithms

The Data Model Resource Book Vol. 3 Universal Patterns For Data Modeling