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 database administrators how to use the open-source Pentaho Kettle for enterprise-level ETL processes (Extracting, Transforming, and Loading data)

• Assumes no prior knowledge of Kettle or ETL, and brings beginners thoroughly up to speed at their own pace

• Explains how to get Kettle solutions up and running, then follows the 34 ETL subsystems model, as created by the Kimball Group, to explore the entire ETL lifecycle, including all aspects of data warehousing with Kettle

• Goes beyond routine tasks to explore how to extend Kettle and scale Kettle solutions using a distributed “cloud”

Get the most out of Pentaho Kettle and your data warehousing with this detailed guide—from simple single table data migration to complex multisystem clustered data integration tasks.
ABOUT THE AUTHOR

Matt Casters is Founder of Kettle and works as Chief Data Integration at Pentaho, where he leads Kettle software development. Roland Bouman is an application developer focusing on open source web technology, databases, and business intelligence. Jos van Dongen is an independent business intelligence consultant and well-known author, analyst, and presenter.

To purchase this product, please visit https://www.wiley.com/en-gb/9780470635179