1. Record Nr.	UNINA9910453487203321
Autore	Meadows Alex
Titolo	Pentaho data integration cookbook : over 100 recipes for building open source ETL solutions with Pentaho data integration / / Alex Meadows, Adrian Sergio Pulvirenti, Maria Carina Roldan
Pubbl/distr/stampa	Birmingham : , : Packt Publishing, , 2013
ISBN	1-78328-068-9
Edizione	[Second edition.]
Descrizione fisica	1 online resource (462 p.)
Altri autori (Persone)	PulvirentiAdrian Sergio RoldaMaria Carina
Disciplina	005.46
Soggetti	Data warehousing Database management - Computer programs Open source software Data structures (Computer science) Data integration (Computer science) Electronic books.
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	First published, June 2011.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	""Cover""; ""Copyright""; ""Credits""; ""About the Author""; ""About the Reviewers""; ""www.PacktPub.com""; ""Table of Contents""; ""Preface""; ""Chapter 1: Working with Databases""; ""Introduction""; ""Connecting to a database""; ""Getting data from a database""; ""Getting data from a database by providing parameters""; ""Getting data from a database by running a query built at runtime""; ""Inserting or updating rows in a table""; ""Inserting new rows where a simple primary key has to be generated""; ""Inserting new rows where the primary key has to be generated based on stored values"" ""Deleting data from a table""""Creating or altering a database table from PDI (design time)""; ""Creating or altering a database table from PDI (runtime)""; ""Inserting, deleting, or updating a table depending on a field""; ""Changing the database connection at runtime""; ""Loading a parent-child table""; ""Building SQL queries via database metadata""; ""Performing repetitive database design tasks from PDI""; ""Chapter 2: Reading and Writing Files""; ""Introduction""; ""Reading a simple file"";

""Reading several files at the same time""; ""Reading semi-structured files""

""Reading files having one field per row""""Reading files with some fields occupying two or more rows""; ""Writing a simple file""; ""Writing a semi-structured file""; ""Providing the name of a file (for reading or writing) dynamically""; ""Using the name of a file (or part of it) as a field""; ""Reading an Excel file""; ""Getting the value of specific cells in an Excel file""; ""Writing an Excel file with several sheets""; ""Writing an Excel file with a dynamic number of sheets""; ""Reading data from an AWS S3 Instance""; ""Chapter 3: Working with Big Data and Cloud Sources""

""Introduction""""Loading data into Salesforce.com""; ""Getting data from Salesforce.com""; ""Loading data into Hadoop""; ""Getting data from Hadoop""; ""Loading data into HBase""; ""Getting data from HBase""; ""Loading data into MongoDB""; ""Getting data from MongoDB""; ""Chapter 4: Manipulating XML Structures"";

""Introduction""; ""Reading simple XML files""; ""Specifying fields by using Path notation""; ""Validating well-formed XML files""; ""Validating an XML file against DTD definitions""; ""Validating an XML file against an XSD schema""; ""Generating a simple XML document""

""Generating complex XML structures""""Generating an HTML page using XML and XSL transformations""; ""Reading an RSS Feed""; ""Generating an RSS Feed""; ""Chapter 5: File Management""; ""Introduction""; ""Copying or moving one or more files""; ""Deleting one or more files""; ""Getting files from a remote server""; ""Putting files on a remote server""; ""Copying or moving a custom list of files""; ""Deleting a custom list of files""; ""Comparing files and folders""; ""Working with ZIP files""; ""Encrypting and decrypting files""; ""Chapter 6: Looking for Data""; ""Introduction""

""Looking for values in a database table""

| Sommario/riassunto | Pentaho Data Integration Cookbook Second Edition is written in a cookbook format, presenting examples in the style of recipes.This allows you to go directly to your topic of interest, or follow topics throughout a chapter to gain a thorough in-depth knowledge.Pentaho Data Integration Cookbook Second Edition is designed for developers who are familiar with the basics of Kettle but who wish to move up to the next level.It is also aimed at advanced users that want to learn how to use the new features of PDI as well as and best practices for working with Kettle. |