Version 5.2

SelectData Integration (DI) Installation Options

Copyright Page

This document supports Pentaho Business Analytics Suite 5.2 GA and Pentaho Data Integration 5.2 GA, documentation revision October 7, 2014, copyright © 2014 Pentaho Corporation. No part may be reprinted without written permission from Pentaho Corporation. All trademarks are the property of their respective owners.

Help and Support Resources

To view the most up-to-date help content, visit https://help.pentaho.com.

If you do not find answers to your questions here, please contact your Pentaho technical support representative.

Support-related questions should be submitted through the Pentaho Customer Support Portal at http://support.pentaho.com.

For information about how to purchase support or enable an additional named support contact, please contact your sales representative, or send an email to sales@pentaho.com.

For information about instructor-led training, visit http://www.pentaho.com/training.

Liability Limits and Warranty Disclaimer

The author(s) of this document have used their best efforts in preparing the content and the programs contained in it. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, express or implied, with regard to these programs or the documentation contained in this book.

The author(s) and Pentaho shall not be liable in the event of incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of the programs, associated instructions, and/or claims.

Trademarks

The trademarks, logos, and service marks ("Marks") displayed on this website are the property of Pentaho Corporation or third party owners of such Marks. You are not permitted to use, copy, or imitate the Mark, in whole or in part, without the prior written consent of Pentaho Corporation or such third party. Trademarks of Pentaho Corporation include, but are not limited, to "Pentaho", its products, services and the Pentaho logo.

Trademarked names may appear throughout this website. Rather than list the names and entities that own the trademarks or inserting a trademark symbol with each mention of the trademarked name, Pentaho Corporation states that it is using the names for editorial purposes only and to the benefit of the trademark owner, with no intention of infringing upon that trademark.

Third-Party Open Source Software

For a listing of open source software used by each Pentaho component, navigate to the folder that contains the Pentaho component. Within that folder, locate a folder named licenses. The licenses folder contains HTML.files that list the names of open source software, their licenses, and required attributions.

Contact Us

Global Headquarters Pentaho Corporation Citadel International, Suite 460

5950 Hazeltine National Drive Orlando, FL 32822

Phone: +1 407 812-OPEN (6736)

Fax: +1 407 517-4575

http://www.pentaho.com

Sales Inquiries: sales@pentaho.com

Introduction

There are several methods to install Pentaho Data Integration (DI). This article helps you make decisions about which method is best for you. If you want to upgrade your current version of the DI components to the newest version, see Upgrade DI Systems.

Prerequisites

There are no prerequisites for this section.

No special expertise is required to understand the concepts presented in this article.

About DI Components

This table lists the components and plugins that come with DI.

Table 1. DI Components Table

Component Type	Name	Description
Server	Data Integration Server	The DI Server is a core component that executes data integration jobs and transformations using the Pentaho Data Integration Engine. It allows you to manage users and roles (default security) or integrate security to your existing security provider such as LDAP or Active Directory. The DI Server provides the ability to centrally store and manage your data integration jobs and transformations. It also provides the services allowing you to schedule and monitor scheduled activities, which are accessed from Spoon.
Design Tool	Spoon	Graphical design environment for modeling Jobs and Transformations. Transformations are metadata models describing the flow of data. Jobs are workflow-like models for coordinating resources, execution and dependencies of data integration activities. It also provides a user interface to the DI Server for managing security and schedules.
Design Tool	Kitchen	Command line tool for executing jobs modeled in Spoon. You can use Kitchen to run jobs at a specific time. For instance, you can use Kitchen to write a cron job that will generate a daily sales report at 3:00 a.m. each morning.
Design Tool	Pan	Command line tool for executing transformations modeled in Spoon. This is similar to Kitchen, but where Kitchen executes jobs, Pan executes transformations.
Design Tool	Carte	Carte is a web server that enables remote execution of jobs and transformations.
Plugin	Instaview	Instaview is software that allows you to use templates to manage the complexities of data access and preparation. Instaview automatically generates transform and metadata models, executes them, and allows you to visualize the results.

Component Type	Name	Description
Plugin	Agile BI	Provides a set of tools that allows developers to collaborate on BI project prototyping. It provides Modeling and Visualization perspectives.

Additionally there are plugins that are developed by the community and therefore, are not documented here.

Decide Where to Install Components

If you are evaluating DI components, we strongly recommend that you install all components on one computer. This enables you to take full advantage of the tutorials and samples we provide to help you evaluate and learn about our products and tools. If you plan to use DI components for development or production, consider distributing components across server and workstation computers. One popular configuration is to install the server on one computer and the design tool on client workstations. DI Plugins that are needed at design and runtime should be installed on the server and on client workstations.

Select Installation Method

To determine the best installation method for you, answer several questions.

- Why do you want to install DI components? Do you want to evaluate it or develop solutions? Do you want to run it in a production environment?

- Do you want to house the DI Repository on the PostgreSQL database that Pentaho supplies? Or would you rather house the DI Repository on another database that you supply such as Oracle or MySQL? The DI Repository stores information related to transformations, jobs, and security. The DI Repository also stores information related to transformations and jobs, such as database connection information and schedules. PDI is configured so that the DI Repository runs on PostgreSQL. The database that houses the DI Repository is different than the database that houses your data.

As you consider these questions, review the Installation Decision Table to determine the best installation method for you.

Table 2. DI Component Installation Decision Table

Explore Considerations	Choose the Best Option For You			
	Install All DI Components	Install with Your Own DI Repository	Use DI Manual Installation	Install Only DI Tools and Plugins
Summary	**This method installs all DI components.** It uses an installation wizard and is the quickest and easiest installation option, and includes	**This method installs all DI components.** This manual method requires that you download and install Pentaho installation files. You can choose the DI Repository	**This method installs all DI components.** The manual method requires that you download and install Pentaho installation files. You can choose your web application	**This method installs the DI Design Tools and Plugins only.** Use this method if you want to install the design tools only, not the DI Server. This method is often

Explore Considerations	Choose the Best Option For You			
	Install All DI Components	**Install with Your Own DI Repository**	**Use DI Manual Installation**	**Install Only DI Tools and Plugins**
	everything you need to run Pentaho DI.	database that you want to use, but you download and install it yourself. Consider using this method if your company has a standard database platform with which you must comply.	server and DI Repository database, but you must download and install them yourself. Consider using this method if your company has a standard database platform or web application server with which you must comply.	used to install design tools and plugins on client workstations.
Installed Components	This method installs these things. • Servers: DI Server • DI Design Tools, Plugins, and Utilities: Spoon, Kitchen, Pan, Cart, Instaview	This method installs these things. • Servers: DI Server • DI Design Tools, Plugins, and Utilities: Spoon, Kitchen, Pan, Cart, Instaview	This method installs these things. • Server: DI Server. • DI Design Tools, Plugins, and Utilities: Spoon, Kitchen, Pan, Cart, Instaview	This method installs these things. • DI Design Tools, Plugins, and Utilities: Spoon, Kitchen, Pan, Cart, Instaview
You Supply	You need to supply a computer that meets Pentaho's operating system and hardware requirements.	You need to supply these things. Each item must meet or exceed the requirements in the supported technologies matrix. • Computer with the supported operating system and hardware configuration. • Oracle Java Runtime Environment (JRE) or Oracle Java Development Kit (JDK). • DI Repository Database (PostgreSQL, MySQL, or Oracle). • DI Repository Database JDBC	You need to supply these things. Each item must meet or exceed the requirements in the supported technologies matrix. • Computer with a supported operating system and hardware configuration. • Oracle Java Runtime Environment (JRE) or Oracle Java Development Kit (JDK) • Repository Database (PostgreSQL, MySQL, or Oracle) You can also use an preexisting	You need to supply a computer that meets Pentaho's operating system and hardware requirements.

Explore Considerations	Choose the Best Option For You			
	Install All DI Components	Install with Your Own DI Repository	Use DI Manual Installation	Install Only DI Tools and Plugins
		Drivers (PosgreSQL, MySQL, or Oracle).	DI Repository platform if it is compatible with the supported technologies matrix. • DI Repository Database JDBC Drivers (PostgreSQL, MySQL, or Oracle) • Web Application Server (JBoss or Tomcat). You can also use an preexisting application server if it is compatible with the supported technologies matrix.	
We Supply	We supply these things. • Installation Package • Oracle JRE • Repository Platform (PostgreSQL) • Repository Platform JDBC Driver • Web Application Server (Tomcat)	We supply these things. • Installation Package • Web Application Server (Tomcat)	We supply the Installation Package.	We supply the Installation Package.
Technologies Used	These technologies are used in this installation method. • Tomcat • PostgreSQL	These technologies are used in this installation method. • Tomcat	These technologies are used in this installation method. • Tomcat or JBoss	N/A

Explore Considerations	Choose the Best Option For You			
	[Install All DI Components](#)	[Install with Your Own DI Repository](#)	[Use DI Manual Installation](#)	[Install Only DI Tools and Plugins](#)
		• PostgreSQL, MySQL, or Oracle	• PostgreSQL, MySQL, or Oracle	
Expertise	You need this knowledge to perform this type of installation. • Basic computer knowledge.	You need this knowledge to perform this type of installation. • Knowledge of your networking environment, including database port numbers if they differ from the default and IP addresses. • Permission to access installation directories. • Root or Administrative access.	You need this knowledge to perform this type of installation. • Knowledge of your networking environment, including database port numbers if they differ from the default and IP addresses. • Permission to access installation directories. • Root or Administrative access.	N/A
Time	Approximately 30 minutes.	Approximately 60 - 90 minutes.	Approximately 2 - 3 hours.	Approximately 5 - 10 minutes.
Availability	Download this software from the [Pentaho website](#) or the [support portal](#).	Download this software from the [Support portal](#).	Download this software from the [Pentaho Support Portal](#).	Download this software from the [Support portal](#).
Recommendation	**Use for the Pentaho Trial Download or if you want to Evaluate DI. Suitable for Rapid Development and Production use.**	**Suitable for Rapid Development and Production use.**	**Suitable for Rapid Development and Production use.**	**Suitable for Rapid Development and Production use.**

Next Steps

Now that you are familiar with the different installation methods, you are ready to install the DI components.

- [Install All DI Components](#)
- [Install with Your Own DI Repository](#)

- Use DI Manual Installation
- Install Only DI Tools and Plugins

If you need installation help, contact Pentaho support.