



Install All DI Components



This document supports Pentaho Business Analytics Suite 5.0 GA and Pentaho Data Integration 5.0 GA, documentation revision September 6, 2013, copyright © 2013 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

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

Pentaho (TM) and the Pentaho logo are registered trademarks of Pentaho Corporation. All other trademarks are the property of their respective owners. Trademarked names may appear throughout this document. Rather than list the names and entities that own the trademarks or insert a trademark symbol with each mention of the trademarked name, Pentaho 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 340
5950 Hazelhine 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

Contents

Introduction to DI Wizard Installation.....	4
Wizard Installation Process Overview.....	5
Prepare Environment.....	6
Create User Account.....	6
Windows User Account.....	6
Linux User Account.....	6
Download Installation Package.....	7
Memory Settings for PostgreSQL.....	7
Install Pentaho Data Integration.....	8
Install Pentaho Data Integration.....	8
The Pentaho Directory Structure for Pentaho Data Integration.....	8
Uninstall Pentaho Data Integration.....	9
Next Steps.....	10

Introduction to DI Wizard Installation

This guide explains how to use the Wizard to install the Pentaho Data Integration suite. After installation, you can use the suite to create and schedule extract, transform, and load processes, model your data, and perform processing related to big data sources. The Wizard can be used to install the Data Integration server and Spoon design tool, as well as the Kitchen, Pan, and Carte utilities along with standard plugins such as Instaview, Big Data, and MongoDB.

Prerequisites

If you are not sure which installation method you want to use, read [Select DI Installation Option](#) before you begin.

Expertise

The topics in this section are written for IT administrators, evaluators, and analysts who have who have the appropriate permissions and accesses required to install the Pentaho Data Integration suite on servers or workstations.

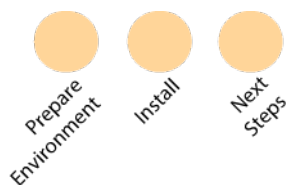
Tools

You must supply a server or workstation that meets the hardware requirements indicated in the [support technologies](#) section, as well as a supported operating system, JRE or JDK, and DI Repository database.

Login Credentials

All of the tasks in this section require that you have the appropriate permissions and accesses required to install software on servers and workstations.

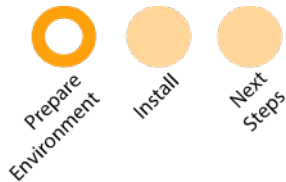
Wizard Installation Process Overview



The Wizard Installation process consists of three steps.

- The [Prepare Environment](#) step explains how to download the installation package. It also provides instruction on how to create a pentaho account.
- The [Install](#) step explains how to use the wizard to install the software.
- [Next Steps](#) indicates what to do after you install the software.

Prepare Environment



To prepare the environment, you must download the installation package and, if you are installing Pentaho on a Linux system, you must create a Pentaho user account. Also, make sure that your memory settings for PostgreSQL are correct.

- [Create Pentaho User Account](#)
- [Download and Unpack Installation Package](#)
- [Check memory settings for PostgreSQL](#)

Create User Account

Windows User Account

If you plan to install Pentaho Data Integration on a server that runs Windows, you do not need to create a special user account on the server. However, you should use an account that has administrator privileges to complete the tasks in this guide.

Linux User Account

If you plan to install Pentaho Data Integration in a Linux environment you must create a user account named **pentaho** on the server computer. By default, license information for Pentaho products is stored in the home directory for this account. Use this account to perform installation tasks that do not require root access. Also, use this account to run start and stop server scripts.

1. Open a **Terminal** window on the server. If you plan to install the software on a remote computer, establish an OpenSSH session to the remote server.
2. In the **Terminal** window, log in as the **root** user by typing this command.

```
su root
```

3. When prompted, type the password in the **Terminal** window.
4. In the **Terminal** window, create a new user account called **pentaho**, along with the **pentaho** home directory, by typing this line.

```
sudo useradd -s /bin/bash -m pentaho
```

Note: /bin/bash indicates that the user account should be created using the Bash shell. In many Linux distributions, the default new user shell is /bin/sh or some equivalent, such as Dash, that might not use the ~/.bashrc configuration file by default. If you don't have or want to use Bash, adjust the instructions throughout this guide accordingly.

5. In the **Terminal** window, assign a password for the **pentaho** user by typing this line.

```
sudo passwd pentaho
```

6. Verify that you can log in using the newly-created **pentaho** user account.
 - a) In the **Terminal** window, log in by typing this command.

```
su pentaho -
```

- b) Type the password for the **pentaho** user account if you are prompted.
- c) Use the **Terminal** window to navigate to the **pentaho** directory to verify that it has been created. By default, it is in the /home directory.
- d) Close the **Terminal** window.

Download Installation Package

Download installation files from the Pentaho website. There are installation files for the Windows, Linux, and Mac platforms. Download the installation file for your operating system.

- **Windows:** `pdi-5.0.0-GA-x64.exe`
- **Linux:** `pdi-5.0.0-GA-x64.bin`
- **Mac:** `pdi-5.0.0-GA-x64.app.tar.gz`

Memory Settings for PostgreSQL

Make sure that your shared memory settings are set as described in <http://www.postgresql.org/docs/devel/static/kernel-resources.html>.



Note: PostgreSQL may fail to start up if the kernel shared memory settings are too low.

Install Pentaho Data Integration

After you have [prepared your environment](#), use the Wizard to install Pentaho Data Integration.

Install Pentaho Data Integration

Use the wizard to install Pentaho tools with the default web application server (Tomcat) and DI Repository (PostgreSQL). Instructions are the same for the Windows, Linux, and Mac platforms.

1. Launch the wizard by double-clicking the installation file or launching it from a **Command Prompt** or **Terminal** window.

Note: If you are using Linux and do not have graphics capabilities, run the installation utility in text mode by invoking the **--mode text** switch. You can also perform the installation on a Linux workstation, then transfer over the `/pentaho/` and `~/pentaho/` directories to your server after installation. The installation utility's graphical mode defaults to the GTK graphics toolkit. If you do not have the requisite GTK libraries on your system, you can run the installer with the **--mode xwindow** switch to use OpenMotif instead.

2. The Pentaho wizard splash screen appears, then the **Welcome** window. Click **Next**.
3. Read the license agreement. To continue, accept the license agreement, then click **Next**.
4. When the **Installation Folder** window appears, accept the default directory or browse to the directory where you want to install Pentaho Data Integration. Click **Next** to continue.

Note: You cannot install into a directory that already exists.

5. When the **PostgreSQL postgres user password** window appears, type the password for the PostgreSQL user. Click **Next** to continue.
6. The **Ready to Install** window appears. Click the **Next** button to install the software.
7. The **Installation in Progress** window appears. The process takes about 15 minutes to complete.
Note: An error occurs if the DI Server port numbers (**9092**, **9080**, **9443**, and **5432**) are not available on your system and if the additional ports that the wizard has tried are not available. If the wizard cannot find an available port, you are prompted to enter port numbers you want to use.
8. The **Installation Complete** window appears. Click the **Open Pentaho Data Integration** and **Begin Instaview** options to view **Spoon**. Click the **Finish** button to end the process.
9. **Spoon** and the **Instaview** perspective are launched.

The Pentaho Directory Structure for Pentaho Data Integration

The wizard installer creates a directory structure for Pentaho Data Integration software. The diagram shows the entire structure. Compare this structure with yours to verify that the installation occurred successfully.

```
/pentaho
../server
../data-integration-server
../../../../data
../../../../licenses
../../../../pentaho-solutions
../../../../tomcat
../design-tools
../data-integration
../common
../documentation
../java
../licenses
../postgresql
../scripts
```

Top-level files:

These files are found in the top-level `/pentaho` directory:

- **ctlscript:** (.sh or .bat) Starts, stops, restarts, and shows the status of Pentaho services.
- **installation-summary.txt:** Contains the information from the summary screen at the end of the installation process.
- **pentaho.log:** The Tomcat log file for the Pentaho Web application.
- **uninstall:** A script that removes what has been installed.

Locations of specific programs

- **DI Server:** /pentaho/server/data-integration-server/
- **Spoon:** /pentaho/design-tools/data-integration/
- **License Installer:** /pentaho/server/enterprise-console/license-installer/

Note: The license installer can also be found in the data-integration directory. All of these license installers perform the same functions, but are available in separate places to account for instances where only specific parts of the Pentaho suite are installed to a particular server or workstation. It does not matter which one you use. You can also install licenses through the User Console if you also install the BA Server.

Uninstall Pentaho Data Integration

If you need to uninstall Pentaho, run the uninstallation utility.

1. Launch the `uninstall.exe` file. Access this file from the **Start** menu in Windows, or you can locate the file in the root folder where you installed the Pentaho software. By default, it is in the `pentaho` directory.
2. A message appears asking whether you want to uninstall the Pentaho software. Click **Yes**.
3. The **Uninstalling Pentaho Integration Enterprise Edition** window appears.
4. When prompted, indicate whether you want to delete the datafiles.
5. A message appears indicating that the uninstallation process is complete. Click **OK**.

Next Steps



Congratulations, you have finished installing Pentaho Data Integration! Now, you are ready to do these things.

- [Configure the DI Server and Design Tool](#)
- Read the [Get Started with DI](#) tutorial.

Note: If you have installed the DI Server and you want to migrate content from the previous version of the servers to this one, make sure that your license keys have been installed, then view the DI Upgrade instructions.