



# Talend Open Studio for Data Quality

## Release Notes

# 6.3.2

Adapted for v6.3.2. Supersedes previous releases.

Publication date September 7, 2017

## Copyright

This documentation is provided under the terms of the Creative Commons Public License (CCPL).

For more information about what you can and cannot do with this documentation in accordance with the CCPL, please read: <http://creativecommons.org/licenses/by-nc-sa/2.0/>

## Notices

Talend is a trademark of Talend, Inc.

All brands, product names, company names, trademarks and service marks are the properties of their respective owners.

## License Agreement

The software described in this documentation is licensed under the Apache License, Version 2.0 (the "License"); you may not use this software except in compliance with the License. You may obtain a copy of the License at <http://www.apache.org/licenses/LICENSE-2.0.html>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

This product includes software developed at ASM, AntLR, Apache ActiveMQ, Apache Ant, Apache Axiom, Apache Axis, Apache Axis 2, Apache Chemistry, Apache Common Http Client, Apache Common Http Core, Apache Commons, Apache Commons Bcel, Apache Commons Lang, Apache Datafu, Apache Derby Database Engine and Embedded JDBC Driver, Apache Geronimo, Apache HCatalog, Apache Hadoop, Apache Hbase, Apache Hive, Apache HttpClient, Apache HttpComponents Client, Apache JAMES, Apache Log4j, Apache Neethi, Apache POI, Apache Pig, Apache Thrift, Apache Tomcat, Apache Xml-RPC, Apache Zookeeper, CSV Tools, DataNucleus, Doug Lea, Ezmorph, Google's phone number handling library, Guava: Google Core Libraries for Java, H2 Embedded Database and JDBC Driver, HighScale Lib, HsqlDB, JSON, JUnit, Jackson Java JSON-processor, Java API for RESTful Services, Java Universal Network Graph, Jaxb, Jaxen, Jetty, Joda-Time, Json Simple, MapDB, MetaStuff, Paracel JDBC Driver, PostgreSQL JDBC Driver, Protocol Buffers - Google's data interchange format, Resty: A simple HTTP REST client for Java, SL4J: Simple Logging Facade for Java, SQLite JDBC Driver, The Castor Project, The Legion of the Bouncy Castle, Woden, Xalan-J, Xerces2, XmlBeans, XmlSchema Core, atinject. Licensed under their respective license.

---

## Table of Contents

System Requirements .....	1
Data Quality: New Features .....	2
1. Profiling perspective .....	2
Data Quality: Bug Fixes / Change Log .....	3
Data Quality: Known Issues and Known Limitations .....	4
1. Java PSU not supported .....	4
2. Installation on Mac OS X Sierra .....	4
3. Installation on recent versions of Mac OS X with Gatekeeper .....	4
Documentation .....	6
1. Talend Help Center .....	6
2. Videos .....	6
3. New documents .....	6
4. Revised documents .....	6
5. Open issues .....	6

# System Requirements

Users should refer to the *Talend Installation and Upgrade Guide* on the **Talend Help Center** (<http://help.talend.com>) for more information on Installation and System Requirements.

# Data Quality: New Features

## 1. Profiling perspective

- A new option has been added to the column and column set analyses to enable users to run the analysis only with a sample data for which they can define the size. This option is interesting when users are analyzing big data on Hive.
- Match Analysis supports now the matching algorithm: Hamming.
- Tokenization methods have been introduced in Match Analysis to break the analyzed string up into meaningful elements called tokens.

# Data Quality: Bug Fixes / Change Log

A number of improvements within the entire product and significant bug fixes have been made.

See the corresponding Change Log on our bug tracking system for more details on the individual issues:

- <https://jira.talendforge.org/issues/?filter=26294>

# Data Quality: Known Issues and Known Limitations

We encourage you to consult the JIRA bug tracking tool for a full list of open issues and to see which issues will be fixed in the upcoming release:

- <https://jira.talendforge.org/issues/?filter=26295>

## 1. Java PSU not supported

Talend is only supported on Java CPS Versions, not Java PSU.

Some users reported problems using Talend Studio with Java PSU versions (such as jdk8\_u92), but when reverting to a Java CPS Version (such as jdk8\_u91) the problem was resolved.

See also: <http://www.oracle.com/technetwork/java/javase/cpu-psu-explained-2331472.html>.

## 2. Installation on Mac OS X Sierra

By default, Mac OS X Sierra sets all downloaded files to "quarantine". This prevents you running the Studio.

To launch Talend Studio on Mac OS X Sierra, do the following:

1. Unzip the Studio as usual.
2. Open a **Terminal** window on the folder where you unzipped your Studio.
3. Execute either of the following commands.

```
xattr -d com.apple.quarantine *
```

or

```
xattr -c *
```

You can now launch the Studio as usual.

## 3. Installation on recent versions of Mac OS X with Gatekeeper

Users with a recent version of Mac OS X may need to update their Gatekeeper settings to install Talend products. Gatekeeper is available from OS X 10.7.5.

To do so:

1. Click the Apple icon, and then click **System Preferences**.
2. Under **System Preferences**, click **Security & Privacy**.

3. Under **Security & Privacy**, click the **General tab**.
4. Click the lock icon next to **Click the lock to make changes**, and then identify yourself using the user name and password for an account with administrator rights. Only administrators can change the security settings.
5. Under **Allow apps download from**, select the **Anywhere** radio button then confirm your choice in the dialog box that opens.
6. Close all windows, so that your Mac automatically locks the **Security & Privacy** settings for you again.
7. Install Talend Studio (or whatever other executable you want to install) by following the usual installation procedure. Note that, when you double click the **Talend-Studio-macosx-cocoa.app** file, you will see a warning saying that it is an application downloaded from the Internet, but if you click **Open**, your Mac should let you install it without further issues.
8. Once you have completed the installation process, you can repeat the above steps to return your **Security & Privacy** settings to their previous value if you want.



---

# Documentation

## 1. Talend Help Center

Find out more about how to get the most out of your Talend products on the Talend Help Center: <http://help.talend.com>.

## 2. Videos

From within the Studio, you can directly access the videos hosted in Talend Help Center relating to new features.

- For components, related videos are available from the F1 help view of this component.
- For other features, related videos are available in the help view of the wizard for the feature.

Note that not every feature has related videos in Talend Help Center and for this reason, not all the features have this type of access.

## 3. New documents

The following new documents have been added in this release:

- *Talend Data Stewardship User Guide* and *Talend Data Stewardship Getting Started Guide* are two new guides that accompany the new product, *Talend Data Stewardship*.

They are both available in PDF format and online at [https://help.talend.com/search/books?filters=EnrichPlatform~%2522Talend+Data+Stewardship%2522\\*EnrichVersion~%25226.4%2522](https://help.talend.com/search/books?filters=EnrichPlatform~%2522Talend+Data+Stewardship%2522*EnrichVersion~%25226.4%2522).

## 4. Revised documents

In addition to updates to the content across the documentation set, the following specific documentation changes have been made:

- In the *Talend Software Development Life Cycle Best Practices Guide*, major changes have been made to the documentation of the Talend CI Builder, including how to generate sources locally, how to manage POM files, artefact versions, and how to filter executions on specific items.

## 5. Open issues

We encourage you to consult the JIRA bug tracking tool for a full list of open issues:

<https://jira.talendforge.org/issues/?filter=18375>