



# RedPoint Data Management Release Notes

## Version 8.1.1

RedPoint Global | 36 Washington Street  
Wellesley Hills, MA 02481  
+1 781 725 0250 | [www.redpointglobal.com](http://www.redpointglobal.com)

# Contents

|   |   |
|---|---|
| Prerequisites and requirements.....   | 1 |
| Features.....   | 2 |
| Data environment-related improvements .....   | 2 |
| Macros .....  | 2 |
| Operating system support .....  | 2 |
| User experience and supportability.....   | 2 |
| Web services.....   | 2 |
| Other features .....  | 2 |
| Fixes .....   | 3 |
| Input/Output tools.....   | 3 |
| Other tools and macros .....  | 3 |
| Hadoop.....   | 3 |
| General.....  | 3 |
| Known issues .....  | 4 |
| 4-byte record ID option removed .....   | 4 |
| SQL scripts containing newlines or semicolons cause problems with Oracle JDBC driver .....        | 4 |
| JARs referenced in classpath aren't shipped to Hadoop .....                                       | 4 |
| Web service ignores Maximum retries setting.....  | 5 |
| Machine Learning Predictor tool: Model Input and Model Output lists are empty.....                | 5 |
| Projects and macros may show spurious "Multiple input connections to upstream tool" warning ..... | 5 |
| Deploying repository data connection creates invisible duplicate RDBMS Input tool.....            | 6 |

# RedPoint Data Management Version 8.1.1 Release Notes

## Prerequisites and requirements

This release of RedPoint Data Management offers the following features and improvements. Before installing it, you should be aware of some prerequisites and requirements:

- You must have the Java Runtime Environment (JRE) Version 8 (64-bit) installed on your server machines to run RedPoint Data Management Version 8. If you have an older version of the Java JRE installed, visit <http://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html> to upgrade.
- If you are installing the Data Management client on a different machine than the Data Management services, you may need to open the firewall ports associated with those processes so that the client can access the services. Data Management Version 8 uses different port numbers than previous versions, so that you can continue to run multiple versions side-by-side.

| Process          | Default ports |
|------------------|---------------|
| Site Server      | 20410         |
| Execution Server | 20411-20416   |
| Web Proxy Server | 20417         |

If any ports in the range 20410-20417 are already in use, you can change these defaults during installation:

- ➔ Click **Show advanced site server settings**, and specify an available **Base Port**. (To review a list of active connections, press the Windows logo key + R and then type `cmd`. Type `netstat -a -n` at the command line.)
- If you are upgrading from Data Management Version 7.x, you will be given the option to migrate the repository to Version 8.0. *Note that previously-defined connections to Hadoop clusters cannot be migrated, and must be defined again.*
- If you are upgrading from Data Management Version 7.x and choose to migrate the repository to Version 8.0, both versions will use the same folder for temporary disk space. This can cause problems if you run both versions at the same time. To prevent this, change the temp space path to a different folder:
  1. In the Data Management client, browse to **Repository > Settings > Site**, and click the **Resources** tab.
  2. Edit each entry under **Temp Spaces** to ensure that Version 8 of Data Management is using a different **Path** than Version 7.

3. Click **Commit**.
4. Restart the Data Management services.

## Features

This release of RedPoint Data Management includes the following improvements:

### *Data environment-related improvements*

- New Input and Output tools that act as simple message producers and consumers for Apache Kafka topics.
- Support for SAP HANA (via JDBC).
- Support for SQL Server 2017.
- Support Azure Data Lake Storage for HDI 3.6.

### *Macros*

- New AO Email Parser tool parses and validates individual email components (local-part, domain) to ensure a successful email campaign or to prepare the data for improved matching.
- UK address matching improvements to handle “Building” address component for customers managing UK addresses.

### *Operating system support*

- Support for Windows Server 2016 Core.

### *User experience and supportability*

- Improved UI experience, with updated look-and-feel for project canvas.
- Project navigator lets you easily change your view of the project canvas.
- New portable diagnostics tool to streamline the troubleshooting process.

### *Web services*

- Support for Blob Output tool in web services enables improved writing of web service payloads to files.
- Support for DELETE, OPTIONS, and HEAD methods in Web Service Call tool.
- Added options for retrying and restarting web services.

### *Other features*

- Regex Match Table tool matches data against a user-defined table of regular expressions, enabling analysis and identification of undefined data content (for example phone numbers) based on patterns.
- Acxiom Data Services Match tool provides direct web access to Acxiom third party data products and entity resolution services to customers with Acxiom Data Services accounts.
- Updated Java SDK sample pack.
- Support for Windows authentication with cross-domain security with SQL Server.

## Fixes

This release of RedPoint Data Management addresses the following issues:

### *Input/Output tools*

- Project could be lost when certain UI actions were performed in the Salesforce Output tool.
- Date values in DBF files could be advanced one day on output to Data Viewer and CSV.
- The JSON Output tool did not show **Analyze** tab when configured for a repository schema.
- RDBMS tools' **Dynamic query** did not work with Unicode input.

### *Other tools and macros*

- The RDBMS Execute tool issued its commands twice.
- The Pair To Group tool could produce incorrect sort orders when used in used in a web service.
- AO Match Test did not apply **Match threshold** setting correctly.
- AO Classifier did not recognize some long secondary names or building/apartment numbers.
- AO Consumer Match Hierchical did not specify default value (**Match residents**) for **Choose match levels** setting.
- Unicode values were converted to Text by macro text controls, resulting in ??? characters for unconvertable code points.

### *Hadoop*

- The Hadoop Application Master treated task pre-emption the same as task failure.
- In a Hadoop parallel section with two inputs, partitioning on two keys could create slightly different data splits on the two inputs, resulting in data mismatch errors.

### *General*

- **Notify connected clients** command issued during upgrade installation didn't work.
- The **Disable all Data Viewers** command was unavailable.
- Data Management date-time fields could convert to JDBC timestamps incorrectly.
- Property pane checkboxes were still selectable even when project was set to read-only.
- Logs that accumulate over time were not purged.

## Known issues

### *4-byte record ID option removed*

Certain Data Management tools generate “record IDs” to sequentially number records. In previous versions of Data Management, the default record ID size was 4 bytes (with an available Site Settings option for 8-byte record IDs). Starting with version 8.1.1, 8-byte record IDs are standard with no 4-byte option. This change accommodates users who have accidentally overflowed 4-byte record numbering during production.

### Potential issues

If you had previously left this option at its 4-byte default, your project behavior may change slightly. Normally this will make no difference. However, there are a few cases that can be sensitive to the record ID size.

- **Flat File Output** and **Flat File Streaming Transform** tools: the fixed-size output field will be widened.
- **RDBMS Output tool**: the database field may be a larger numeric type if you create DDL or a new table based on the upstream connection schema.
- **Extract Layout** tool will produce a different field size in the report output.
- **\$FieldSize()** function will produce a result of 8 instead of 4.

### Workaround

If you have existing projects that are adversely affected by this change, add an upstream **Change Field Types** ahead of the affected tool(s) to change the size of the record ID column back to 4 bytes.

### *SQL scripts containing newlines or semicolons cause problems with Oracle JDBC driver*

When using JDBC to execute scripts, Oracle's JDBC driver usually requires that statements are *not* terminated with a semicolon, and that long commands (such as a CREATE TABLE script with many column descriptions) contain no internal newlines. This applies to preload and postload scripts in the RDBMS Output tool, as well as scripts in the RDBMS Execute and RDBMS In-Database Processor tools. Data Management enforces these rules programmatically.

However, there are cases where Oracle *requires* embedded newlines and a semicolon after the final command (especially if this command is `END;`).

### Workaround

If you are using Oracle's JDBC driver and must use an embedded newline, precede it with a single backslash, the standard escape character ( \ ). Similarly, if you must use a semicolon at the end of a command, precede it with a single backslash ( \ ), or double the semicolon ( ; ; ).

### *JARs referenced in classpath aren't shipped to Hadoop*

When JARs are added to the classpath using the **Add folders and jars to classpath** option in Repository Site Settings or Haroop Settings, they are added to the classpath, but are not shipped to Hadoop.

### Workaround

Copy the JAR files to the `java_plugins` folder in your Data Management installation location (typically `C:\Program Files\RedPointDM8\java_plugins`). Data Management will ship all JARs in this folder to Hadoop.

### *Web service ignores Maximum retries setting*

An enabled web service may continue to attempt to restart after failing, even when the **Maximum retries** value defined in Site Settings has been reached.

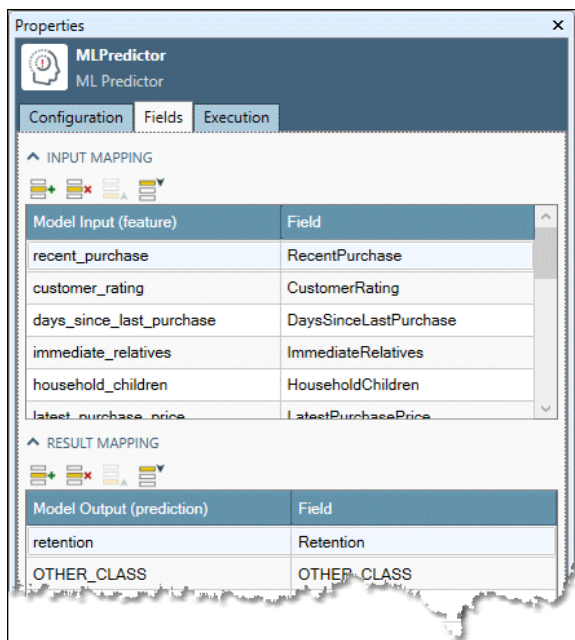
#### Workaround

In the Repository, navigate to Web Services, right-click on the web service, and then click **Disable**.

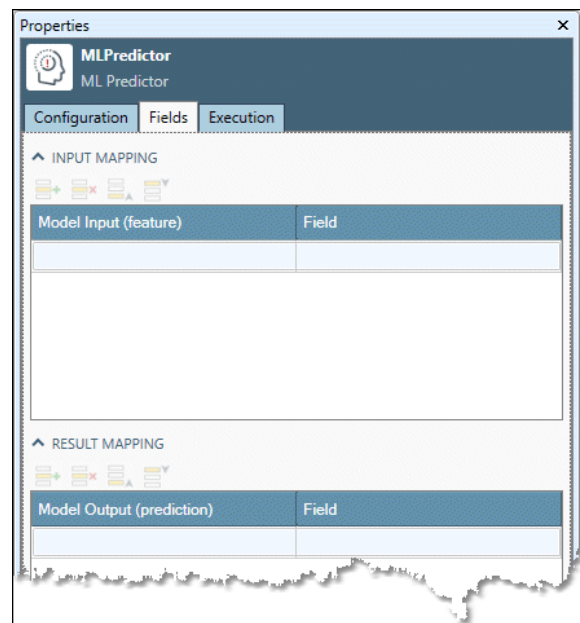
### *Machine Learning Predictor tool: Model Input and Model Output lists are empty*

To configure the Machine Learning Predictor tool, the user specifies a solution file (or *model*) generated by the Machine Learning Trainer, and maps that model's **Model Input (feature)** and **Model Output (prediction)** parameters to input fields. However, if the solution file was generated by an older version of Data Management, the **Model Input (feature)** and **Model Output (prediction)** lists may not display any parameters.

*ML Predictor tool displaying parameters*



*ML Predictor tool with empty parameter lists*



#### Workaround

Re-run the Machine Learning Trainer project to generate an updated solution file.

### *Projects and macros may show spurious "Multiple input connections to upstream tool" warning*

Some projects and macros may incorrectly issue a multiple connection warning when tool connection end points are the same (for example, if you connect multiple different outputs of the same upstream tool into a single Merge tool).

#### Workaround

Ignore the warning.

### *Deploying repository data connection creates invisible duplicate RDBMS Input tool*

If you drag a data connection from the Data Management repository to the project canvas, an extra invisible RDBMS Input tool is created, as shown in the project's Message Viewer.

#### Workarounds

Press F5 to refresh the project, and then select the RDBMS Input tool and delete it. The duplicate tool will remain.

*-or-*

Instead of dragging the data connection from the Repository, deploy an RDBMS Input tool and configure it to use the desired data connection:

1. Select the RDBMS Input tool and then view the **Properties** pane.
2. Click the **Source** tab, and then click **Use data connection**.
3. Select a data connection from the drop-down list.